THE EFFECT OF NAMING SYSTEMS
ON THE ACQUISITION OF AND REASONING ABOUT
TIME CONCEPTS

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF
THE UNIVERSITY OF HAWAI‘I IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

LINGUISTICS

AUGUST 2012

By
Nian Liu

Dissertation Committee:

Benjamin Bergen, Chairman
Kamil Deen
William O’Grady
Amy Schafer
 Luca Onnis
Catherine Sophian
© 2012
by
Nian Liu
To Dan X. Hall
ACKNOWLEDGEMENTS

This dissertation would not exist without the tremendous support I have received from the faculty and staff of the Department of Linguistics at the University of Hawai‘i. The department has given me many wonderful opportunities to learn and teach. The years I have spent in this ‘ohana have been a truly joyful and fruitful experience. I especially love the competitive yet cooperative atmosphere here that is fostered by the inspiring faculty and students, and that has encouraged me to take academic challenges while having fun. Also, my research could never have come this far without the support of the NSF’s Doctoral Dissertation Improvement Grant (BCS-1124006). Additionally, many thanks to the following organizations for their contributions: the Arts and Sciences Advisory Council, the Center for Chinese Studies, the UH Foundation, the Graduate Student Organization at the University of Hawai‘i, and the East West Center.

Among the members of my committee, my deepest appreciation goes to my advisor and committee chair, Benjamin K. Bergen. Thank you for your profound knowledge, sharp and open mind, invaluable advice, and constant support. Through your patient and perceptive mentorship, you provided more than just guidance. The days after you left Hawai‘i were hard for all of your students, including me, but the continuous support and timely feedback you offered encouraged me to keep on taking academic challenges. I have always felt lucky to have such a caring and fun advisor. The times you spent with us outside class were also unforgettable; the CLRG-dinner meetings, the beers, bowling, and karaoke nights are all vivid in my memory. If I can one day become a scholar resembling you just a little, I will be very content. By the way, did I ever
confess that it was me and two other perpetrators who defaced your office door and spread the “wanted” posters around the department on that April Fool’s Day? Surprise.

My sincere gratitude also goes to Kamil Deen. Your positive encouragement and unwavering optimism have always inspired me to pursue my goals with confidence. I am grateful for the extensive time and knowledge you have shared with me. Thanks (or not) to your encouragement, I sometimes day-dream about becoming rich by publishing a book on bilingualism. Given that making a fortune by doing linguistic research is difficult, I think you are responsible for the time wasted on my fruitless day-dreaming.

I would like to thank Amy Schafer. I’ve benefited immensely from taking your classes, but the methodologies are not the most important things I learned from you; your integrity and academic rigor have encouraged me to take every detail seriously and shown me the way to be a professional scholar. And I was so close to winning that “murder in Moore Hall” game—only to lose to you. The last thing I expected was to see you waiting outside of the unit mastery classroom—a wonder woman not only in academics—that was my last thought before being “murdered.”

I would like to thank William O’Grady. Despite the incredible amount of work you had as department chair along with your teaching and researching, you have always been ready to give whatever kind of help you can, be it a reference letter, feedback on conference presentations, or preparation for a job interview. Also, many thanks to Luca Onnis, for your valuable feedback and discussion of possible experimental designs, especially when your schedule was really tight. I would also like to thank Catherine Sophian, who has always provided me with new perspectives for viewing this
dissertation’s topic. Your contributions to my dissertation are far greater than I could have asked for from an “outside member.”

My heart-felt thanks also go to other faculty members in the Linguistics Department: to Dr. Robert Blust for his informative and inspiring class; to Dr. David Stampe and Dr. Patricia Donegan for their humorous lectures and fun parties; to Dr. Ann Peters for her guidance through my first two years of doctoral study; to Dr. Ken Rehg for his sharing of fun video clips on phonology; to Dr. Yuko Otsuka for sharing with me her baby-caring experience; to Dr. Michael Forman for the interesting debate competition in class; to Dr. Victoria Anderson for her emotional support; and to Dr. Albert Schütz for his efforts editing my working paper. Without the positive influence from all of you, I would not be the person I am now.

During the journey of my doctoral research, I have also been trained by faculty members from beyond the Department of Linguistics. My sincere thanks to Ying-Che Li, Tao-Chung Yao, Hsin-I Hsieh, and Song Jiang in the EALL department. I’ve learned a lot from your classes, and being given opportunities to teach in the EALL department has helped me both academically and financially. I would also like to thank Virginia Yip, who gave me her profound understanding and support during my dissertation writing at the Chinese University of Hong Kong.

Thanks to all of the department and college staff who have provided behind-the-scenes but substantial support. Jennifer Kanda’s and Nora Lum’s incredible patience and professionalism have helped make everything go smoothly. Thanks to John Kawahara for his help with grant-related business and Laurie Durand for her excellent editing skills.
Thanks also to Kurt Brunner for his assistance in lab technology, which has saved me tremendous time and energy.

This dissertation includes several cross-linguistic studies and would never have come into being without generous help from overseas. I owe significant thanks to Dr. Jurgis Skilters. Without that informative conference lunch conversation and the fruitful trip to the Center for Cognitive Sciences and Semantics (CCSS) in beautiful Riga, I might never have developed a conference paper into a dissertation. I owe a lot to Kate Hardeman and Chuan Lin. Thanks for devoting your time to helping me with sound stimuli recordings. My sincere thanks to my research assistants—Dana Vilistere, Yukyoung Shin, Megan (Joelle) Kirtley, and Jessica Seid—for their time and efforts. I would also like to thank Liang Zhang and Qiaoling Deng, who helped recruit a large number of participants within a short time.

Besides all of the academic support, I am also indebted for emotional support from my many friends who I met in Hawai‘i. Deep thanks to Manami Sato, Heeyeon Dennison, and Jung-Hee Kim. The three of you have shared with me your personal experiences of doctoral study and given me valuable advice as “elder sisters.” Enormous thanks to Satoko Kimura Kiyabu, who is my “younger sister” and has the ability to always make me smile. Kaya-chan is lucky to have a mother like you. Thanks to Carl Polley, Jawee Perla, Kathryn Wheeler, Wenwei Han, Bodo Winter, and Stephanie Kakadelis; our vigorous CLRG meetings were the best part of Friday nights. My thanks to Suann Chen and Jinsun Choe; the summer in Berkeley was as good as the kimchi soup and bibimbap you two made. My heart-felt thanks also go to my fellow linguists in the department for their company and warm encouragement. My thanks to Karen Huang,
Akiemi Glenn, Elaine Lau, and Apay Tang. The ABD-group meeting was a great encouragement. I would also like to thank Zoe Wood, Dorinda Liu, Tsz Him Tsui, Meylysa Tseng, Hiroko Sato, Soyoung Kim, Kanjana Thepborriruk, Antonio Cheung, Onsoon Lee, Hunter Hatfield, Toshi Furukawa, Jun Nomura, John Van Way, Jinsook Kim, Sang-Gu Kang, Yukie Hara, Yumiko Enyo, Zhijun Wen, Shu-Ling Wu, Ying Zhou, Jing Zhang, Hui-ju Chuang, Qiaona Yu, and Vivian (Yiwei) Hsu.

Above all, I am truly grateful to my family, without whom I would never have come this far. I am deeply indebted to my parents and parents-in-law for their vital help in taking care of my daughter during the past one and a half years. My deepest gratitude is reserved for my dear husband and best friend, Hao Zhang, who plays many roles other than husband in my life, such as driver, chef, body guard, technician, therapist, and above all, my soul-mate. Whenever I feel lost or depressed, you are always by my side, letting me know that I am not alone. One day when I was complaining about how hard it is to do research and joking that I just wanted to be a “housewife,” you smiled and said “Well, you already are Hao’s wife.” Do you still remember the eight-character vow I made at our wedding? I would like to say that again now.

Last but not least, as I have moved toward the completion of both my dissertation and the journey of my doctoral study, I have been thinking more and more about Dan. Especially while writing my dissertation at the beautiful campus of the Chinese University of Hong Kong, I could not help thinking of the time we came here for a conference four years ago. And I remember all of the adventures the three of us had on the Hawaiian islands, and in Las Vegas, Europe, and around China. For many years, you were always there providing unconditional support and love. I know how much you
wanted to see the day when I become Dr. Liu, and we already even had plans for the celebration. The day is finally coming, but you won’t be here placing a lei on me. Dan, my dear friend and “big brother,” I dedicate this dissertation to you.
ABSTRACT

Languages encode time-related concepts, like days of the week (DOW) and months of the year (MOY), in different ways. Some use numerical labels (e.g., in Chinese, Monday is “weekday one”; January is “month one”) while others use arbitrary names (e.g., English). This project investigates whether languages that use numerical terms provide an advantage to their speakers, both as children acquiring the temporal terms of the language and as adults reasoning about time, when compared to speakers of languages that use arbitrary symbols to encode time-related terms.

The project first tests whether a numerical naming system facilitates children’s acquisition of time sequences, comparing the behavior of monolingual children who speak Chinese (numerical day- and month-systems), English (non-numerical systems), Latvian (a mix of systems), and Korean (a mix of systems). The results show that (1) Chinese monolingual children comprehend and use time words earlier overall than English monolingual children; (2) in Latvian, which has numerical DOW names but arbitrary MOY names, children acquire day-names earlier than month-names; (3) in Korean, which has numerical MOY names and arbitrary DOW names, children acquire month-names earlier than day-names.

Second, two behavioral experiments test whether Chinese- and English-speaking adults’ time calculation abilities are influenced by their languages’ naming systems. The results are consistent with the view that Chinese and English speakers are using different strategies in online processing of calendrical questions.
This dissertation research offers a new piece of evidence for the pervasive influence of language on thought, in the specific domain of cognition of time, by showing that the way calendars are coded can have a substantial effect on the employment of strategies in non-linguistic problem-solving processes.
# TABLE OF CONTENTS

Acknowledgments.............................................................................................................. iv
Abstract............................................................................................................................ x
List of Tables.................................................................................................................. xiv
List of Figures................................................................................................................... xvi

Chapter 1:
Introduction...................................................................................................................... 1

Chapter 2: Acquisition studies.......................................................................................... 11

2.1 Experiment 1: Test with Chinese and English monolingual children ............................................................... 13
   2.1.1 Participants............................................................................................................ 13
   2.1.2 Materials and Design.......................................................................................... 14
      2.1.2.1 Pretest materials and design......................................................................... 14
      2.1.2.2 Main experiment materials and design......................................................... 15
      2.1.2.3 Posttest materials and design...................................................................... 16
   2.1.3 Procedure............................................................................................................. 17
      2.1.3.1 Pretest procedure.......................................................................................... 17
      2.1.3.2 Main experiment procedure........................................................................ 18
      2.1.3.3 Posttest procedure....................................................................................... 19
   2.1.4 Results................................................................................................................ 19
      2.1.4.1 Results from the main experiment............................................................... 19
      2.1.4.2 Results from the pretest and posttest........................................................ 21
   2.1.5 Discussion............................................................................................................ 23

2.2 Experiment 2: DOW and MOY tests with Latvian-speaking children....................... 25
   2.2.1 Participants............................................................................................................ 25
   2.2.2 Materials and Design.......................................................................................... 26
   2.2.3 Procedure............................................................................................................. 26
   2.2.4 Results................................................................................................................ 27
      2.2.4.1 Results from the main experiment............................................................... 27
      2.2.4.2 Results from the pretest and posttest........................................................ 28
   2.2.5 Discussion............................................................................................................ 28

2.3 Experiment 3: DOW and MOY tests with Korean-speaking children....................... 29
   2.3.1 Participants............................................................................................................ 30
   2.3.2 Materials, Design and Procedure....................................................................... 31
   2.3.3 Results................................................................................................................ 31
      2.3.3.1 Results from the main experiment............................................................... 31
      2.3.3.2 Results from the pretest and posttest........................................................ 32
   2.3.4 Discussion............................................................................................................ 33

2.4 General discussion of Experiments 1, 2, and 3......................................................... 33

Chapter 3: Behavioral study with adults.......................................................................... 40
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Names of days of the week in English and Chinese</td>
</tr>
<tr>
<td>1-2</td>
<td>Names of months of the year in English and Chinese</td>
</tr>
<tr>
<td>2-1</td>
<td>Description of complexity levels and sample questions in the DOW test (five questions per level for a total of 25)</td>
</tr>
<tr>
<td>2-2</td>
<td>Mean score in each language and age group with Chinese- and English-speaking children</td>
</tr>
<tr>
<td>2-3</td>
<td>Mean scores in different levels of the DOW test for English and Chinese groups</td>
</tr>
<tr>
<td>2-4</td>
<td>Mean scores in posttest for the English and Chinese language groups</td>
</tr>
<tr>
<td>2-5</td>
<td>Latvian numerical DOW and arbitrary MOY terms</td>
</tr>
<tr>
<td>2-6</td>
<td>Mean scores of DOW (numerical) and MOY (arbitrary) tests in each age group with Latvian-speaking children</td>
</tr>
<tr>
<td>2-7</td>
<td>Korean arbitrary DOW and numerical MOY terms</td>
</tr>
<tr>
<td>2-8</td>
<td>Mean scores of DOW (numerical) and MOY (arbitrary) tests in each age group with Korean-speaking children</td>
</tr>
<tr>
<td>2-9</td>
<td>Chinese ordinal and cardinal numbers compared with DOW terms</td>
</tr>
<tr>
<td>2-10</td>
<td>Latvian ordinal and cardinal numbers compared with DOW terms</td>
</tr>
<tr>
<td>2-11</td>
<td>Mean scores in posttest for the Latvian and Korean language groups</td>
</tr>
<tr>
<td>3-1</td>
<td>Mean RT and SD in each condition of the MOY task</td>
</tr>
<tr>
<td>3-2</td>
<td>Mean RT and SD in each condition of the DOW task</td>
</tr>
<tr>
<td>3-3</td>
<td>Mean RT and SD in Sunday and Non-Sunday conditions for Chinese and English speakers</td>
</tr>
<tr>
<td>3-4</td>
<td>Summary of results in error and reaction time analysis</td>
</tr>
<tr>
<td>4-1</td>
<td>The ten Celestial Stems as a Chinese ordinal system</td>
</tr>
<tr>
<td>4-2</td>
<td>Mean RT and SD in each condition for the Chinese speakers</td>
</tr>
<tr>
<td>Page</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
</tr>
<tr>
<td>4-3</td>
<td>Chinese speakers’ reaction times by compatibility of Prime Direction and Critical Direction</td>
</tr>
<tr>
<td>4-4</td>
<td>Mean RT and SD in each condition for the English speakers</td>
</tr>
<tr>
<td>4-5</td>
<td>English speakers’ reaction times regarding Compatibility of Prime Direction and Critical Direction</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Mean scores for English and Chinese speakers in DOW test</td>
<td>20</td>
</tr>
<tr>
<td>2-2</td>
<td>Mean scores in DOW (numerical) and MOY (arbitrary) tests of Latvian-speaking children</td>
<td>28</td>
</tr>
<tr>
<td>2-3</td>
<td>Mean scores in DOW (numerical) and MOY (arbitrary) tests of Korean-speaking children</td>
<td>32</td>
</tr>
<tr>
<td>2-4</td>
<td>Mean scores in DOW (numerical) test of Chinese-, English-, Latvian-, and Korean-speaking children</td>
<td>35</td>
</tr>
<tr>
<td>3-1</td>
<td>Accuracy rate on the MOY task for Chinese and English speakers</td>
<td>49</td>
</tr>
<tr>
<td>3-2</td>
<td>Accuracy rate on the DOW task for Chinese and English speakers</td>
<td>51</td>
</tr>
<tr>
<td>3-3</td>
<td>Chinese- and English-speaking participants’ accuracy rate in different Sunday conditions</td>
<td>54</td>
</tr>
<tr>
<td>3-4</td>
<td>Reaction times on the MOY task for Chinese and English speakers</td>
<td>59</td>
</tr>
<tr>
<td>3-5</td>
<td>Reaction times on the DOW task for Chinese and English speakers</td>
<td>62</td>
</tr>
<tr>
<td>3-6</td>
<td>Reaction times in Sunday and Non-Sunday conditions for Chinese and English speakers</td>
<td>65</td>
</tr>
<tr>
<td>3-7</td>
<td>Reaction times of Chinese (left) and English (right) in Within Week calculations</td>
<td>71</td>
</tr>
<tr>
<td>3-8</td>
<td>Reaction times of Chinese (left) and English (right) in Across Week calculations</td>
<td>72</td>
</tr>
<tr>
<td>3-9</td>
<td>Reaction times of Chinese (left) and English (right) in Within Year calculations</td>
<td>74</td>
</tr>
<tr>
<td>3-10</td>
<td>Reaction times of Chinese (left) and English (right) in Across Year calculations</td>
<td>76</td>
</tr>
<tr>
<td>3-11</td>
<td>Accuracy of Chinese (left) and English (right) in Within (upper) and Across (lower) Week calculations</td>
<td>79</td>
</tr>
<tr>
<td>3-12</td>
<td>Accuracy of Chinese (left) and English (right) in Within (upper) and Across (lower) Year calculations</td>
<td>81</td>
</tr>
<tr>
<td>3-13</td>
<td>Reaction times for Forward and Backward questions of the Chinese speakers</td>
<td>86</td>
</tr>
</tbody>
</table>
3-14 Reaction times for Forward and Backward questions of the English speakers…86
4-1 Accuracy rates for the Chinese speakers in different Prime conditions……..100
4-2 Accuracy rates for the English speakers in different Prime conditions………101
4-3 Mean response time in each Prime condition for the Chinese speakers………106
4-4 Chinese speakers’ reaction times when Prime Direction and Critical Direction are Compatible (left) or Incompatible (right)…………………………………………………………..109
4-5 Mean response time in each Prime condition for the English speakers………111
4-6 English speakers’ reaction times when Prime Direction and Critical Direction are Compatible (left) or Incompatible (right)……………………………………………………………112
4-7 English speakers’ reaction times in List (left) and Arithmetic (right) conditions for compatibility of Prime Direction and Critical Direction……………………………………114
4-8 Chinese speakers’ reaction times in different Prime conditions for Short (left) and Long (right) Distance calculations………………………………………………………………115
4-9 English speakers’ reaction times in different Prime conditions for Short (left) and Long (right) Distance calculations………………………………………………………………117
4-10 English speakers’ reaction times for Short (left) and Long (right) Distance calculations by Boundary Type……………………………………………………………………118
4-11 Chinese speakers’ reaction times by Distance in each Boundary and Direction condition…………………………………………………………………………………..119
4-12 English speakers’ reaction times by Distance in each Boundary and Direction condition…………………………………………………………………………………..120
The concept of time is complex; given that we cannot see it or touch it, it is a remarkable feat of human cognition that over the course of our development, we come to be able to not only talk but also reason about it. That reasoning includes calculating the temporal distance between one event and another, representing sequences of events, and relating specific events to cyclical temporal frameworks like weeks or years. What exactly are the cognitive mechanisms that people use to perform these types of temporal reasoning, and how do they develop?

The prevailing view on the matter comes from developmental work by Friedman (1990), who presents a model of mental representations of time in which an initial list-based representation (memorized sequences of days or months) is gradually supplemented by an analog spatial representation that incorporates information about the distances between different time units. On this view, children first learn a list of names and remember the ordering; at this stage, reciting the list is necessary for them to determine the distance between days or months. As they grow older, a new strategy of using the visual image of a calendar may be adopted, which permits them to figure out sequential relationships between days and months and use spatial distances to determine the relative distances between them. This process is thought to be based on the acquisition of counting (Williams, 2008); children mainly count by touching objects one by one while reciting the names, and the total count is the last name recited. Later they
can use fingers as proxies for objects being counted, and the total count is determined by the number of fingers raised.

But this proposal, that children develop the concepts of days and months through a two-step process of first remembering a list of unrelated names, and then gradually figuring out the sequential relationship by their order of appearance in the list, may not apply—at least not in the same way—to children developing in different linguistic and cultural environments. The two-stage view is based on the English-language version of the Gregorian calendar, in which the names for days and months are—as far as a child learning them is concerned—semantically opaque. But some languages use more numerically transparent calendar systems. It could be that these different ways to name time concepts allow children to use different learning strategies, which lead to different developmental trajectories. What’s more, adults might implement different mental representations of time concepts and different methods for performing temporal reasoning.

Two examples of numerically transparent calendar naming systems are the Chinese days of the week (DOW) and months of the year (MOY) systems. Unlike English, in which DOW and MOY terms are derived from planetary or mythical terms (Zerubavel, 1985), Chinese uses numbers in the names of months and days. For DOW, this system begins with Monday as the first day of the week. Thus, xīngqī-yī, \(^1\) “weekday one,” is Monday, xīngqī-èr, “weekday two,” is Tuesday, etc. The only exception is Sunday, the last day of the week, which is xīngqī-rī in Mandarin Chinese, which

\(^1\) Literally “star-period one.”
translates to “weekday of sun.” Similarly, months in Chinese are numbered from one to 12, with “one month” denoting January, “two month” February, and so on.

<table>
<thead>
<tr>
<th>English (planetary)</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
<th>Sunday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moon</td>
<td>Mars (Tiw)</td>
<td>Mercury (Woden)</td>
<td>Jupiter (Thor)</td>
<td>Venus (Fria)</td>
<td>Saturn</td>
<td>Sun</td>
<td></td>
</tr>
<tr>
<td>Chinese (numerical)</td>
<td>星期一 (xīngqī-yī)</td>
<td>星期二 (xīngqī-èr)</td>
<td>星期三 (xīngqī-sān)</td>
<td>星期四 (xīngqī-sì)</td>
<td>星期五 (xīngqī-wǔ)</td>
<td>星期六 (xīngqī-liù)</td>
<td>星期日 (xīngqī-rì)</td>
</tr>
<tr>
<td>Weekday 1</td>
<td>Weekday 2</td>
<td>Weekday 3</td>
<td>Weekday 4</td>
<td>Weekday 5</td>
<td>Weekday 6</td>
<td>Weekday sun</td>
<td></td>
</tr>
</tbody>
</table>

Table 1-2: Names of months of the year in English and Chinese

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Janus</td>
<td>Februa</td>
<td>Mars</td>
<td>Aphros</td>
<td>Maia</td>
<td>Juno</td>
<td>Julius</td>
<td>Augustus</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Chinese (numerical)</td>
<td>一月</td>
<td>二月</td>
<td>三月</td>
<td>四月</td>
<td>五月</td>
<td>六月</td>
<td>七月</td>
<td>八月</td>
<td>九月</td>
<td>十月</td>
<td>十一月</td>
<td>十二月</td>
</tr>
<tr>
<td>1mo.</td>
<td>2mo.</td>
<td>3mo.</td>
<td>4mo.</td>
<td>5mo.</td>
<td>6mo.</td>
<td>7mo.</td>
<td>8mo.</td>
<td>9mo.</td>
<td>10mo.</td>
<td>11mo.</td>
<td>12mo.</td>
<td></td>
</tr>
</tbody>
</table>

Children learn basic counting as early as age 2 or 3, before they learn DOW or MOY (Brandt, 1996). This has two potential ramifications for the acquisition of time concepts. First, English time words are—to the learner—arbitrary symbols that they have to commit to memory, whereas the Chinese learner who has already learned small number terms can relate DOW or MOY terms to these other words that they already know. This might facilitate early learning of the DOW and MOY terms themselves. And second, since the Chinese terms are semantically transparent, relating overtly to numbers, any reasoning that the learner knows how to perform with numbers can potentially be

---

2 In English, names for months from September to December are numerical, but their relation to standard English numbers is less transparent than the relationship is in Chinese.
transferred to DOW and MOY reasoning. For instance, if the learner knows the sequence of numbers one through six, then this may facilitate enumerating the sequence of days, (Monday “weekday one” through Saturday “weekday six”) or months. Similarly, once the learner can use basic arithmetic with numbers, this may make temporal arithmetic with semantically transparent DOW and MOY terms easier. By contrast, semantically opaque English DOW and MOY words do not afford any of these potential benefits transferred from knowledge of numbers—once DOW and MOY terms are learned, their sequence must be memorized independently, and learners may or may not ever come to use arithmetic with them.

Can such differences in numerical transparency of DOW and MOY terms across languages affect the age at which children acquire time concepts, and the kinds of cognitive mechanisms they use to reason about time? There has been little work on this issue. There have been a number of studies on numerical cognition itself, showing that differences in number naming systems can affect cognitive development and non-linguistic performance. For example, acquisition studies (Miura, Okamoto, Kim, Steere, & Fayol, 1993; Miura et al., 1994; Miller, Smith, Zhu, & Zhang, 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 with combinations of individual unit cube block representing one and long blocks representing ten, Asian-language-speaking children who learned numerical names congruent with base 10 numeration systems (Fuson, 1990) were much more likely to use the blocks of 10 in
constructing multi-digit numbers than their non-Asian-language-speaking counterparts, who lacked the access to such transparent numerical naming systems. The authors of that study argued 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 verbal number system in French-speaking Belgium is simpler than the one used in France. They reported that second-grade children in France made more errors in Arabic number production than their Belgian counterparts. The effects of naming systems also extend into adulthood and mathematical performance. For instance, one study showed that adult English speakers have difficulty reversing two-digit numbers ending in 1 (e.g., saying “14” when shown “41”), while Chinese speakers showed no such difficulty, 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 research I have conducted investigates whether the same is true for the naming of time concepts (in this case, DOW and MOY). The current study first investigates whether numerically transparent naming systems for time concepts facilitate children’s acquisition of those concepts. Second, I investigate whether adults’ temporal reasoning abilities differ, depending on the transparency of the naming systems that their languages employ for time sequences. Finally, the research tests whether these
hypothesized differences in performance are due to different strategies employed by adults when they perform temporal calculations, again, depending on the transparency of their time concept naming systems.

A useful way of structuring questions about how the characteristics of language may affect cognitive development is offered by Miller and Paredes (1996), who make a three-way distinction among (1) effects on initial acquisition, (2) continuing effects on the use of that system after it has been acquired, and (3) effects on conceptual understanding of the domain that the symbol system represents. In this research, I ask these questions in the following, more specific form. First, how does the use of numbers in temporal words affect children’s acquisition of time structure? Second, do any effects carry over to adult performance of temporal calculations? And if so, what are the differences in mechanism that underlie these differences? It is worth noting that this research does not actually address the third of Miller and Paredes’ (1996) points. That is, I am not looking at whether conceptualizations of time are different in adults, only whether temporal reasoning is different.

The first research question to be explored in this dissertation is whether differences in the naming systems for time concepts affect children’s acquisition of time concepts. Some of the problems children experience in mastering symbol systems stem from the fact that symbolic representations are arbitrary and complex (Deloache, Miller, & Pierroutsakos, 1998). Children’s mastery of such systems may be facilitated, however, by linguistic linkages between symbol and number systems. Pollman (2003) claims that the principles that govern an English-speaking child’s acquisition of numbers are the same as those that drive acquisition of the DOW—instead of mapping meaning-to-form
for each word, children learn by “rote a list of at first meaningless word forms and detect in this list a repetitive structure that enables the learner to produce other items of the list” (Pollman, 2003, p. 3). If it is true that mastery of abstract time sequences such as the DOW is so closely related to mastery of numerical sequences, then it is likely that, when the time sequences make use of numerical terms in their naming system, the previously acquired knowledge of numbers can facilitate early acquisition of the DOW and MOY.

The different naming systems exhibited by Chinese and English, for instance, may result in different acquisition ages for time words. Chinese-speaking children have the benefit of a temporal term system that is based on an already-acquired numeric system. English-speaking children, on the other hand, have no existing system on which to base the acquisition of time concepts. Therefore, Chinese-speaking children should show an earlier acquisition of the DOW and MOY systems than English-speaking children because of their system’s reliance on numerosity for its naming conventions. Because, however, there may be many other differences across these two languages and the cultures they are embedded in, establishing a causal link between the transparency of time language and the age of acquisition of time words will require comparisons across and within other languages as well. The experiments that I designed and conducted for this purpose are described in detail in Chapter 2.

The second research question is whether the naming system effect carries over to adulthood in such a way that it affects performance of time calculations. Do different developmental trajectories caused by differences in the naming of time concepts have lingering effects on adults? On the one hand, adults in many cultures make use of similarly formatted calendars when performing time-related calculations. If, as Friedman
(1990) argues, adults perform time calculations using mental images of calendars, and calendars are similar across cultures, then cross-linguistic differences based on naming system differences might dissipate in adulthood. On the other hand, it could be that the developmental origin of time concepts affects subsequent time calculations, throughout the lifespan.

Previous cross-linguistic studies have shown that there is at least some causal influence from language to non-linguistic cognition and unconscious habitual thought (Kay & Kempton, 1984; Lucy, 1992; Gumperz & Levinson, 1996; Boroditsky, 2001), although the precise implications of these studies continue to be debated in the literature (e.g., Levinson, Kita, Haun, & Rasch, 2002; Li & Gleitman, 2002). 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, 2009).

The transparent numerical structure of the Chinese calendar might facilitate calendar calculation, causing Chinese-speaking adults to outperform their English-speaking counterparts in time calculation tasks, perhaps by exhibiting shorter reaction times and making fewer errors. In general, such a finding would support the hypothesis that linguistic differences can produce non-linguistic consequences, in this case by affecting people’s reasoning about time (Boroditsky, 2000, 2003; Boroditsky & Ramscar, 2002; Matlock, Ramscar, & Boroditsky, 2005; Núñez & Sweetser, 2006; Casasanto & Boroditsky, 2008).

Chapter 2 explores the first research question by describing a set of three acquisition experiments (Experiments 1–3) that investigate whether the use of numerical
systems in the names of time concepts affects how children acquire time concepts. More specifically, Experiment 1 compares the acquisition of days of the week concepts by monolingual child speakers of Chinese (numerical day- and month-systems) and English (non-numerical systems). Experiments 2 and 3 test the performance of child speakers of Latvian (numerical day- and arbitrary month-systems) and Korean (arbitrary day- and numerical month-systems) on two sets of day of the week and month of the year questions. The results from the three experiments show that in a number of calendar-related tasks, Chinese-speaking children outperform English-speaking children in that they acquire these terms earlier and get higher scores on tests that require manipulating these terms. Also, Latvian-speaking children acquire Latvian’s numerically coded day-names earlier than the arbitrary month-names, while the Korean-speaking children show the opposite pattern, acquiring Korean’s numerically coded month-names before the arbitrary day-names.

Chapters 3 and 4 discuss two behavioral experiments with Chinese- and English-speaking adults and investigate whether the naming system effect carries over into adulthood such that it affects how adults reason about time. More specifically, Chapter 3 describes Experiment 4, which tests Chinese and English speakers’ performance in temporal distance calculation tasks in order to check whether there is any difference influenced by their language’s calendrical naming system. Chapter 4 describes a priming study (Experiment 5) that further tests what specific time calculation strategies Chinese- and English-speaking adults are using. The results confirm the prediction that differences in the systems in Chinese and English affect the online processing of these languages’ speakers.
Finally, Chapter 5 discusses the implications of the findings of this study and suggests possible directions for future research.

In sum, this project explores how language is influenced by and influences general cognition, in the specific domain of cognition of time. The results provide evidence that the use of numerical systems in the names of time concepts affects how children acquire time concepts and how adults reason about time.
CHAPTER 2
ACQUISITION STUDIES

The contrasting calendar naming systems described in Chapter 1—the Chinese reliance on numerosity compared to the English use of arbitrary symbols—should be manifested in a difference in the rate of acquisition of the abstract realm of time words associated with the notion of a week. This is because the Chinese-speaking children can easily transfer the knowledge of numbers to their understanding of week terms, while the English-speaking children’s acquisition of week terms may not have this foundation to build on. The ultimate objective of this chapter is to demonstrate this empirically. First, however, I need to establish that there actually is a difference between the Chinese speakers’ and the English speakers’ acquisition rates.

A series of acquisition studies has been conducted to explore whether transparency of time naming systems affects the age at which children acquire time concepts. There has been relatively little work on this issue, the major exception being the research done by Kelly, Miller, Fang, & Feng (1999), who tested Chinese- and English-speaking children’s performance in calendar naming and calculation tasks. Their experiments, however, only included children from the second and fourth grades, which means that many of their subjects (with an average age of 9) were past the age at which the calendar system is acquired (6 to 7 years old, according to Brandt [1996]). To provide a picture of the initial acquisition of time-related terms and concepts, I therefore worked with children starting at a younger age, and I used a more refined method, which is described in this section and in Section 2.1.
Before designing an experiment to test for a difference in acquisition rates between Chinese and English speakers, I determined a means of measuring proficiency. This brings us back to the notion of staged development. It is possible that when acquiring the concept of the days of the week, children first get to know that the week is a time unit with different parts and then get to know the names for the different parts. After that, they may gradually figure out the relationships or order of these parts within a week. At a more advanced stage, they may finally learn to manipulate the different days of the week terms to solve temporal distance problems. Based on these possible stages of learning days of the week terms, I propose that the general ability to understand and use DOW terms can be broken down into five levels of increasing proficiency. I will hereafter refer to these as the five levels of DOW acquisition. They are described in Table 2-1.

Table 2-1: Description of complexity levels and sample questions in the DOW test (five questions per level for a total of 25)

<table>
<thead>
<tr>
<th>Level of Complexity</th>
<th>Description</th>
<th>Sample Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Basic Composition</td>
<td>Knowledge of the week as an entity of time that has parts</td>
<td>How many days are there in a week?</td>
</tr>
<tr>
<td>2 Name Recognition</td>
<td>Knowledge of and the ability to distinguish the specific names of the days of the week</td>
<td>On what day does Winnie the Pooh go swimming?</td>
</tr>
<tr>
<td>3 Adjacency Relationships</td>
<td>Knowledge that days are sequentially related and the ability to solve problems that involve days that occur next to each other</td>
<td>Today is Tuesday and Winnie goes swimming. What will he do tomorrow?</td>
</tr>
<tr>
<td>4 Within-week Distance</td>
<td>Ability to recognize, compute, and verbalize the temporal relationship of days that are not simply adjacent, but are still within the scope of the same target week</td>
<td>Today is Sunday and Winnie eats some honey. On Tuesday he will fly a balloon. How many days must he wait to fly a balloon?</td>
</tr>
<tr>
<td>5 Cross-week Distance</td>
<td>Ability to recognize, compute, and verbalize the temporal relationship of days that cross the boundaries of a conventional 7-day week as configured in the speaker’s native language</td>
<td>Today is Friday and Winnie goes swimming. Next Monday he will ride in a boat. How many days must he wait to ride in a boat?</td>
</tr>
</tbody>
</table>
Employing these five definitions, I designed and conducted Experiment 1, the materials for which consisted of a children’s story and a list of 25 questions (five at each level of complexity). This experiment aimed at assessing at what ages children in the two language populations reach the various proficiency levels.

2.1 Experiment 1: Test with Chinese-monolingual and English-monolingual children

2.1.1 Participants

One hundred children of various ages between 3 and 7 participated in this study. They were divided into two groups according to their native language.

The Chinese group consisted of 50 Chinese-monolingual children at a public kindergarten and a public primary school in Jiangsu Province, China. The 3-year-old participants (7 females, 3 males) ranged in age from 3;00 to 3;12 ($M = 3;05, SD = 0.48$). The 4-year-old participants (6 females, 4 males) ranged in age from 4;01 to 4;07 ($M = 4;04, SD = 0.20$). The 5-year-old participants (9 females, 1 male) ranged in age from 5;00 to 5;05 ($M = 5;02, SD = 0.17$). The 6-year-old participants (5 females, 5 males) ranged in age from 6;00 to 6;03 ($M = 6;01, SD = 0.10$). The 7-year-old participants (5 females, 5 males) ranged in age from 7;03 to 7;07 ($M = 7;06, SD = 0.11$).

The other 50 participants were English-monolingual children living in the state of Hawai’i in the United States. Their native language is English, and they will be referred to as the English group. The 3-year-old participants (6 females, 4 males) ranged in age from 3;00 to 3;03 ($M = 3;01, SD = 0.10$). The 4-year-old participants (5 females, 5 males) ranged in age from 4;03 to 4;08 ($M = 4;06, SD = 0.16$). The 5-year-old participants (3
females, 7 males) ranged in age from 5;00 to 5;05 ($M = 5;02, SD = 0.16$). The 6-year-old participants (6 females, 4 males) ranged in age from 6;02 to 6;10 ($M = 6;06, SD = 0.22$). The 7-year-old participants (6 females, 4 males) ranged in age from 7;00 to 7;11 ($M = 7;06, SD = 0.33$).

### 2.1.2 Materials and Design

The experiment consisted of three parts—a pretest to match the general cognitive ability of the youngest children in the two language groups, the main experimental test, and a posttest to measure the older children’s mathematical ability. The pretest was given only to the 3- and 4-year-old children, because at this age most children have not yet quite acquired arithmetic knowledge, and it is reasonable to test and match their general cognitive ability instead. On the other hand, for children above age 5, the posttest was administered to measure their arithmetic ability. The researcher conducting the experiment spoke the same native language as the language group being tested.

#### 2.1.2.1 Pretest materials and design

For the pretest, two sets of materials were used. One is the set of picture cards described by Zelazo (2006) for the Dimensional Change Card Sort (DCCS), which is “an easily administered and widely used measure of executive function that is suitable for use with participants across a wide range of ages” (p. 297). The standard version of DCCS used as a pretest in this experiment is a switch test in which the children are required to sort a series of bivalent test cards, first according to one dimension (color) for six preswitch trials and then according to another dimension (shape) for another six
postswitch trials. There are 16 cards, each depicting one of two entities (a rabbit or a boat) in one of two colors (red or blue). The children’s performance is marked by a binary standard as pass or fail. When the child is successful with a majority of the trials (5 out of 6) in the post-switch session, he or she is considered as passing the test. “The performance on the DCCS provides an index of the development of executive function” (Zelazo, 2006, p. 297), as it accesses the flexible use of rules to govern behavior.

Employing the DCCS in this experiment allowed me to compare and measure whether children from the two language groups were matched in general cognitive ability and mental flexibility.

2.1.2.2 Main experiment materials and design

The main experiment was conducted using two types of materials: a set of story cards and a list of stimulus questions. The story cards consisted of seven 4.5 inch by 3.5 inch laminated cards, each with a picture of the cartoon character Winnie the Pooh engaged in some activity easily recognizable to a child. Each card was also associated with one of the seven days of the week (see Appendix A for a graphic representation of the actual picture cards). Together, these cards formed a story involving Winnie the Pooh’s activities on each day of the week.

The stimulus questions (presented in Appendix A) consisted of short questions in either Chinese or English. Each question was designed to assess how well participants understood the words used for the names of the DOW and how proficient they were at using them to solve problems involving time. Also, each question was designed to elicit a short (usually one- or two-word) response from the participants. These questions were
further divided into five groups representing the five levels of proficiency regarding time words as outlined in Table 2-1 in Section 2.1.1.

2.1.2.3 Posttest materials and design

The posttest serves as a control test of general arithmetic ability. It is necessary because, if the Chinese group shows any advantage in temporal terms acquisition and distance calculations, it could be the result of better mathematical ability in general. It is necessary, therefore, to add a test controlling for this confounding factor. The materials were a list of 10 questions requiring calculation of two units of time other than days of the week: hours and dates. Both hours and dates are numerical in English and Chinese. If the English speakers perform as well as the Chinese speakers in these tests, then general mathematical competence cannot explain a score difference in the main experiment.

Two sets of five questions with calculations involving hours and dates were generated. Sample questions are: “Winnie the Pooh is going to visit Tigger by ferry. Now it is 5 o’clock and the ferry will leave at 10. How many hours does Winnie the Pooh have to wait?” for the hour calculation, and “Today is the 4th of December. Six days from now Mickey has a soccer contest. What date will that be?” for the date calculation. These 10 questions included Backward/Forward and Within Boundary/Across Boundary calculations. The calculation distance was controlled to use numbers under 12, as in the questions in the main experiment. The children received scores based on their answers in the control test, with a perfect score being 10 points. A complete list of the 10 questions can be found in Appendix B.
2.1.3 Procedure

The experiment was conducted in two ways. The 3- and 4-year-old children were first tested with the DCCS task followed by the main experiment. The other three age groups began with the main experiment, which was followed by the posttest of the date and hour calculations.

2.1.3.1 Pretest procedure

In the pretest, the researcher conducting the experiment placed two sorting trays in front of the participant. The experimenter initiated the pretest by telling the child to participate in a “color game.” The test began with the experimenter first labeling the target cards one by one by both dimensions (e.g., “Here is a red rabbit”/“Here is a blue boat”) and then giving instructions verbally (e.g., “The red rabbit goes to this tray and all of the other red cards go to this tray, too”/“The blue boat goes to this tray and all of the other blue cards go to this tray, too”) while putting two cards face-down into two trays. Basically, the child was asked to place each randomly selected card into a sorting tray according to the color of the entity in the picture card. After six trials of the “color game,” the experimenter proceeded immediately to the postswitch session by telling the child to play a new game, the “shape game.” Similar instructions were given, except that this time it was the shape of the entity on the cards that determined in which sorting tray a picture should be placed. The postswitch session also consisted of six trials. If the child was able to sort five or more out of six correctly in the postswitch task, he or she was considered to have passed the task. It took about five minutes for the children to complete the pretest.
2.1.3.2 Main experiment procedure

The researcher conducting the experiment guided each participant through three experiment phases. The first phase was the story phase, in which the researcher laid out the story cards in front of the participant one card at a time while telling the story of Pooh’s week of activities. The researcher began with the Sunday card (in the case of English speakers) or the Monday card (in the case of Chinese speakers). Each card was revealed as the story unfolded. The researcher, making sure that the child was paying attention, would verbally relate what the activity was by saying, for instance, “On Sunday Winnie the Pooh eats some honey.” The story phase was finished when all seven cards were spread out before the child and each day’s activity had been described in the child’s native language.

In phase two, the researcher left the cards spread out in front of the participant in the order they were introduced and asked 15 questions from the list of questions. These questions were taken from Levels 1, 2, and 3 of the question list and randomly ordered. Thus, while the cards were still visible in DOW order, the researcher would ask, for instance, “On what day does Winnie the Pooh go swimming?” The child was given some leeway in providing an answer to the question. In case a child would not answer immediately or gave an unrelated response, such as “Swimming in the lake,” the researcher would repeat the question if necessary until the child gave an appropriate response or until the researcher decided that the child could not respond to the question, at which time the next question was asked.

Phase three was conducted similarly to phase two with one exception—the questions asked in this phase were taken from the sets of stimulus questions in Levels 4
and 5, both of which require the child to solve time-related problems that involve non-adjacent days. As in phase two, the researcher waited until it was clear that the participant had provided an appropriate answer or could not answer the question.

All participants were rewarded at the end of the experimental procedure with a parchment certificate of completion and some candy or stickers. The whole procedure took about 10 to 20 minutes.

2.1.3.3 Posttest procedure

The procedure for the posttest was the same as for the main experiment: the children viewed a picture card of a cartoon character and answered questions about hours and dates. The order of questions was random, and the children received a score for each correctly answered question.

2.1.4 Results

2.1.4.1 Results from the main experiment

In order to score the overall performance by a participant, one point was assigned for each correct answer in the main experiment. There was a total of 25 questions (five each for the five levels of proficiency), so the maximum possible score was 25. In this way, a given participant’s degree of overall understanding of how to use the DOW and ability to accurately verbalize about time-related problems was reflected in their numeric score: the higher the score, the greater the degree of time-word proficiency.

In data coding, I noticed that there were two possible ways to answer Level 4 and Level 5 questions. For example, the correct answer to how many days there are from Sunday to Tuesday can be two or three, depending on whether or not both the first and
last day are included in the count. Both answers were considered correct in the data coding as long as the way of computing was consistent for each participant. The data from five participants from the Chinese group and three from the English group were excluded from the data analysis because the participants were not concentrating on the experiment and failed to respond to the experimenter. The results (Table 2-2) showed that in all age groups the Chinese-speaking children demonstrated a better understanding of the weekday terms. By age 7, the Chinese-speaking children reached the maximum possible score, suggesting that at least for the tasks at hand, they had attained proficiency. The comparison is clearly presented in Figure 2-1.

Table 2-2: Mean score in each language and age group with Chinese- and English-speaking children

<table>
<thead>
<tr>
<th></th>
<th>3 yrs</th>
<th>4 yrs</th>
<th>5 yrs</th>
<th>6 yrs</th>
<th>7 yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>English</strong></td>
<td>5.0 (N = 10)</td>
<td>8.1 (N = 10)</td>
<td>8.8 (N = 9)</td>
<td>15.5 (N = 10)</td>
<td>21.9 (N = 8)</td>
</tr>
<tr>
<td><strong>Chinese</strong></td>
<td>10.6 (N = 10)</td>
<td>14.7 (N = 10)</td>
<td>18.8 (N = 10)</td>
<td>20.9 (N = 7)</td>
<td>24.6 (N = 8)</td>
</tr>
</tbody>
</table>

Figure 2-1: Mean scores for English and Chinese speakers in DOW test
The data were submitted to a two-way ANOVA, with Language and Age as categorical independent variables and Score as the dependent measure. Language showed a large main effect, \( F(1, 82) = 42.94, p < 0.001 \); there was a significant difference in the scores of the two language groups, indicating that they performed at different levels on the tasks when age is controlled for. The results also showed a large main effect for Age, \( F(4, 82) = 32.685, p < 0.001 \); scores were affected by age group. A further regression test confirmed that Age significantly predicted Score, \( R^2 = 0.48, F(1, 90) = 83.541, p < 0.001 \); with increasing age, the scores are significantly higher. However, there was no interaction between Age and Language, \( F(4, 82) = 1.41, p = 0.24 \), which means that the difference in scores between the language groups is not measurably different at different ages. In sum, this analysis suggests that there is a significant difference between the time-word proficiencies exhibited by the children in the two target populations.

2.1.4.2 Results from the pretest and posttest

All of the 4-year-old participants passed the DCCS task. In the 3-year-old groups, three out of 10 English speakers and two out of 10 Chinese speakers failed the task. This result suggests that the children in the two language groups were at about the same level of general cognitive ability.

The Chinese speakers’ score advantage in these age groups in the main test was overwhelmingly for the questions at Levels 1 to 3, which involve no calculation, and so their acquisition advantage cannot be due to mathematical ability. The scores broken down by level are presented below in Table 2-3.
Table 2-3: Mean scores in different levels of the DOW test for English and Chinese groups

<table>
<thead>
<tr>
<th>Age</th>
<th>Language</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 yr. old</td>
<td>English</td>
<td>2.2</td>
<td>1.3</td>
<td>1.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Chinese</td>
<td>2.9</td>
<td>3.4</td>
<td>3.2</td>
<td>0.9</td>
<td>0.2</td>
</tr>
<tr>
<td>4 yr. old</td>
<td>English</td>
<td>2.5</td>
<td>1.4</td>
<td>2.3</td>
<td>1.1</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>Chinese</td>
<td>3</td>
<td>4.9</td>
<td>4.4</td>
<td>2</td>
<td>0.4</td>
</tr>
<tr>
<td>5 yr. old</td>
<td>English</td>
<td>2.1</td>
<td>2.1</td>
<td>2.9</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>Chinese</td>
<td>3.2</td>
<td>5.0</td>
<td>5.0</td>
<td>4.4</td>
<td>1.2</td>
</tr>
<tr>
<td>6 yr. old</td>
<td>English</td>
<td>2.9</td>
<td>4.4</td>
<td>4.0</td>
<td>3.0</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>Chinese</td>
<td>4.6</td>
<td>4.7</td>
<td>4.7</td>
<td>4.6</td>
<td>2.3</td>
</tr>
<tr>
<td>7 yr. old</td>
<td>English</td>
<td>4.6</td>
<td>5.0</td>
<td>5.0</td>
<td>4.3</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>Chinese</td>
<td>4.6</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
</tr>
</tbody>
</table>

It could be that the Chinese children were better at solving all kinds of problems, or that their better performance resulted from the transparency of number and calendar terms in their native language, which helps them learn the calendar terms earlier and form an understanding of the sequential relationship between the terms. The results of the pretest confirm that it is not general cognitive ability that distinguishes the two language groups’ performance in the DOW test.

The scores from the date and hour calculation posttest are reported in Table 2-4. The Chinese children had higher scores in the date and hour calculations. Score was the dependent measure in a two-way ANOVA with Language and Age as categorical independent variables. The results showed that the difference was not statistically significant as there was no main effect of Language in the ANOVA test, $F(1, 45) = 1.88$, $p$
= 0.18, demonstrating that general mathematical competence is about at the same level for 5- to 7-year-old English- and Chinese-speaking children.

Table 2-4: Mean scores in posttest for the English and Chinese language groups

<table>
<thead>
<tr>
<th></th>
<th>5 yr. old</th>
<th>6 yr. old</th>
<th>7 yr. old</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>4.7 (SD = 1.41)</td>
<td>6.5 (SD = 2.12)</td>
<td>8.1 (SD = 1.57)</td>
</tr>
<tr>
<td>Chinese</td>
<td>5.5 (SD = 2.12)</td>
<td>7.1 (SD = 2.19)</td>
<td>9.0 (SD = 1.30)</td>
</tr>
</tbody>
</table>

In addition, an analysis of covariance was performed to compare the Chinese and English 5-, 6-, and 7-year-olds with posttest scores as covariates. The result reveals that the children’s math ability does not correlate with the test score attained in the main experiment, $F (1,45) = 0.015$, $p = 0.903$. This shows that the Chinese- and English-speaking children performed differently in the date/hour calculations and the DOW questions, and that the difference comes from how the days of the week terms are coded in their native languages, rather than from a general mathematical competence difference.

### 2.1.5 Discussion

These results provide support for the notion that children who learn Chinese as a native language are able to comprehend and use terminology for the DOW earlier than children who learn English as a native language. It is possible that the different systems for naming the DOW are at the root of this language-specific acquisition difference. Chinese-speaking children, like their English-speaking counterparts, acquire the basic numeric sequence terms at an early age (Brown, 1973). And the use of number words to
refer to the DOW in Chinese might form a link between the two acquisition processes that provides the Chinese-speaking child with an advantage in achieving time-word proficiency, through a process that we might call lexical leveraging. This hypothesized leverage would start taking effect by the age of 3, and would explain why Chinese- and English-speaking children diverge as early as that age in their linguistic development regarding the DOW and their ability to solve time-related DOW problems.

Alternatively, however, the difference could be due to a host of cultural factors that were not controlled for in this experiment. The Chinese children did attain higher scores in the posttest, though the difference was not statistically significant. Given the fact that the number of children tested in each age group was small (about 10), the match of the general mathematical competence of the two language groups may need further and stronger proof. Comparison between samples of children drawn from distinct populations runs the risk of encountering unknown confounds—whether those confounds be cultural, linguistic, individual, sampling, or others. It is desirable, therefore, to compare the acquisition of time terms and time concepts within a single language to prove that the acquisition difference found in this experiment is not due to other factors, but to linguistic differences.

For this purpose, I pursued this promising line of research by replicating this experiment, with some modifications, in other languages. I expected that doing so would allow me to tease apart numerical transparency of time words from other cultural factors, as described in Section 2.2 and Section 2.3.
2.2 Experiment 2: DOW and MOY tests with Latvian-speaking children

The first extension looked at Latvian, which is what might be called a “hybrid” language with respect to numerical transparency of time words. Latvian has numerically transparent names for days of the week, similar to Chinese, but it makes use of arbitrary names for months of the year, like English. The days of the week and months of the year terms in Latvian are presented in Table 2-5. If it is the numerical transparency of time terms that affects age of acquisition, and not other cultural factors, then for speakers of Latvian, time-word proficiency should be acquired at a faster, Chinese-like pace for the days of the week, but at a slower, English-like pace for months of the year.

Table 2-5: Latvian numerical DOW and arbitrary MOY terms

<table>
<thead>
<tr>
<th>English</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
<th>Sunday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latvian</td>
<td>Piršdiena</td>
<td>Otrdiena</td>
<td>Trešdiena</td>
<td>Ceturtdiena</td>
<td>Piektidiena</td>
<td>Sestdiena</td>
<td>Svētdiena</td>
</tr>
<tr>
<td>Translation</td>
<td>1st day</td>
<td>2nd day</td>
<td>3rd day</td>
<td>4th day</td>
<td>5th day</td>
<td>6th day</td>
<td>Holy day</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>English</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latvian</td>
<td>Janvāris</td>
<td>Februāris</td>
<td>Marts</td>
<td>Aprīlis</td>
<td>Maijs</td>
<td>Jūnijs</td>
</tr>
<tr>
<td>English</td>
<td>July</td>
<td>August</td>
<td>September</td>
<td>October</td>
<td>November</td>
<td>December</td>
</tr>
<tr>
<td>Latvian</td>
<td>Jūlijs</td>
<td>Augsts</td>
<td>Septembris</td>
<td>Oktobris</td>
<td>Novembris</td>
<td>Decembris</td>
</tr>
</tbody>
</table>

2.2.1 Participants

Fifty Latvian-speaking children participated in the pilot study. The children were randomly selected from a public kindergarten and a public primary school in Riga, the capital of Latvia. The 3-year-old participants (6 females, 4 males) ranged in age from
3;09 to 3;11 (\(M = 3;10, SD = 0.10\)). The 4-year-olds (4 females, 6 males) ranged in age from 4;01 to 4;10 (\(M = 4;06, SD = 0.25\)). The 5-year-old participants (4 females, 6 males) ranged in age from 5;01 to 5;11 (\(M = 5;06, SD = 0.32\)). The 6-year-old participants (6 females, 4 males) ranged in age from 6;00 to 6;11 (\(M = 6;04, SD = 0.35\)). The 7-year-old participants (6 females, 4 males) ranged in age from 7;01 to 7;07 (\(M = 7;04, SD = 0.18\)).

2.2.2 Materials and Design

The same types of materials were used in this experiment as in Experiment 1, but with an additional month of the year test. The months are another important time unit that can be used in a variation of the DOW test. The prediction is that the Latvian-speaking children will score higher in the DOW test than in the MOY test. Each child was tested with two sets of questions, DOW and MOY.

For the MOY test, another set of 12 picture cards, these describing Mickey Mouse’s annual activities, was created, and 25 questions about MOY were added to the question and answer phase. The questions were generated according to the five different difficulty levels described in Table 2-1, using year instead of week as a boundary. A complete list of the questions can be found in Appendix C.

2.2.3 Procedure

The whole procedure was the same as in Experiment 1. All of the children were tested individually by a Latvian-speaking experimenter.

The procedure for the MOY test is the same as for the DOW test. In the main experiment, each child was tested with two sets of questions, DOW and MOY. The
DOW and MOY tests were conducted sequentially, and the testing order was balanced across participants. The results are described in Section 2.2.4.

2.2.4 Results

2.2.4.1 Results from the main experiment

Two participants failed to concentrate and interact with the experimenter, and their data were excluded from the analysis; another four participants failed to complete the experiment, and their data were also excluded. The results were as expected. Table 2-6 and Figure 2-2 show that the mean score for the DOW test is higher than for the MOY test for all age groups.

Table 2-6: Mean scores of DOW (numerical) and MOY (arbitrary) tests in each age group with Latvian-speaking children

<table>
<thead>
<tr>
<th>Latvian</th>
<th>3 yrs (N = 8)</th>
<th>4 yrs (N = 10)</th>
<th>5 yrs (N = 10)</th>
<th>6 yrs (N = 9)</th>
<th>7 yrs (N = 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOW (numerical)</td>
<td>5.4 (SD = 1.9)</td>
<td>6.4 (SD = 5.0)</td>
<td>14.6 (SD = 5.8)</td>
<td>19.0 (SD = 6.3)</td>
<td>20.7 (SD = 2.9)</td>
</tr>
<tr>
<td>MOY (arbitrary)</td>
<td>1.6 (SD = 1.2)</td>
<td>4.9 (SD = 4.8)</td>
<td>11.1 (SD = 5.4)</td>
<td>14.1 (SD = 6.7)</td>
<td>18.1 (SD = 3.8)</td>
</tr>
</tbody>
</table>

A repeated-measures ANOVA, with Age as the categorical between-subject factor and Test Type as the within-subject factor, showed a main effect of Test Type, \( F(1,39) = 45.26, p < 0.001 \), confirming the observation that Latvian-speaking children generally performed better in the DOW test. Age also had a main effect, \( F(4,39) = 18.26, p < 0.001 \), and regressions showed that the scores produced by different age groups were different, for both DOW questions, \( R^2 = 0.60, F(1,42) = 63.08, p < 0.001 \), and MOY questions, \( R^2 = 0.60, F(1,42) = 63.55, p < 0.001 \).
2.2.4.2 Results from the pretest and the posttest

All of the 4-year-old participants passed the DCCS task. In the 3-year-old group, three out of 10 children failed to pass the pretest. The scores from the date and hour calculation posttest for the older children were also recorded. The 5-year-old group has an average score of 4.3 ($SD = 1.7$). The 6- and 7-year-old groups scored 6.2 ($SD = 2.0$) and 7.7 ($SD = 1.8$), respectively.

2.2.5 Discussion

Because the Test Type (DOW and MOY) is a within-subjects factor, it cannot be differences in education, culture, and other such factors that caused the difference between DOW and MOY. Instead, the better mastery of DOW concepts by Latvian-speaking children in each age group may result from the numerical transparency of terms for days of the week, in contrast with the opacity of terms for months of the year.

Figure 2-2: Mean scores in DOW (numerical) and MOY (arbitrary) tests of Latvian-speaking children
Alternatively, however, it could be that children, regardless of their language, learn about days of the week earlier than months of the year due to less exposure to month terms in daily life and/or month concepts being difficult to understand. Experiment 3 was designed to test this hypothesis, and it is described in Section 2.3.

2.3 Experiment 3: DOW and MOY tests with Korean-speaking children

There could be many reasons why Latvian-speaking children perform better on DOW than MOY tests, including frequency differences, differences in simplicity or accessibility of the concepts, and so on. Or it could be that the numerical transparency of time words makes a difference.

Table 2-7: Korean arbitrary DOW and numerical MOY terms

<table>
<thead>
<tr>
<th>English</th>
<th>Sunday</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td>일요일</td>
<td>월요일</td>
<td>화요일</td>
<td>수요일</td>
<td>목요일</td>
<td>금요일</td>
<td>토요일</td>
</tr>
<tr>
<td></td>
<td><em>Ir-yo-il</em></td>
<td><em>Wor-yo-il</em></td>
<td><em>Hwa-yo-il</em></td>
<td><em>Su-yo-il</em></td>
<td><em>Mog-yo-il</em></td>
<td><em>Geum-yo-il</em></td>
<td><em>To-yo-il</em></td>
</tr>
<tr>
<td>Translation</td>
<td>Day of sun</td>
<td>Day of moon</td>
<td>Day of fire</td>
<td>Day of water</td>
<td>Day of wood</td>
<td>Day of metal</td>
<td>Day of earth</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td>일월</td>
<td>이월</td>
<td>삼월</td>
<td>사월</td>
<td>오월</td>
<td>난월</td>
<td>칠월</td>
<td>팔월</td>
<td>구월</td>
<td>십월</td>
<td>십일월</td>
<td>십이월</td>
</tr>
<tr>
<td></td>
<td><em>Il-wol</em></td>
<td><em>I-wol</em></td>
<td><em>Sam-wol</em></td>
<td><em>Sawol</em></td>
<td><em>O-wol</em></td>
<td><em>Yuwol</em></td>
<td><em>Chil-wol</em></td>
<td><em>Pal-wol</em></td>
<td><em>Guwol</em></td>
<td><em>Si-wol</em></td>
<td><em>Sip-il-wol</em></td>
<td><em>Sip-i-wol</em></td>
</tr>
<tr>
<td></td>
<td>1mo.</td>
<td>2mo.</td>
<td>3mo.</td>
<td>4mo.</td>
<td>5mo.</td>
<td>6mo.</td>
<td>7mo.</td>
<td>8mo.</td>
<td>9mo.</td>
<td>10mo.</td>
<td>11mo.</td>
<td>12mo.</td>
</tr>
</tbody>
</table>

To distinguish between these two possibilities, Experiment 3 was designed to replicate Experiment 2 in a language with the opposite pattern for DOW and MOY terms. Korean is another “hybrid” language, like Latvian, but with a reversed pattern of time concept naming. Instead of transparent DOW terms, it has opaque, arbitrarily named
DOW terms. And its MOY terms, rather than being opaque, are transparent and numerically based, as shown in Table 2-7.

If it is something about DOW concepts themselves that makes them easier than MOY terms for Latvian-speaking children to learn, then Korean-speaking children should show acquisition patterns like Latvian—children learning both languages should learn DOW terms before MOY terms. But if it is the numerical transparency or opacity of the terms that influences the age at which they are acquired, then Korean-speaking children’s acquisition should show the opposite pattern of Latvian-speaking children’s—MOY terms should be learned earlier than DOW terms.

2.3.1 Participants

Fifty Korean-speaking children participated in this experiment. Forty were residents in Seoul and Pusan, South Korea, while the other 10 lived in the state of Hawai‘i. The participants recruited in Hawai‘i were all native speakers of Korean and had limited exposure to English in their daily lives. The 3-year-old participants (7 females, 3 males) ranged in age from 3;01 to 3;11 (M = 3; 70, SD = 0.31). The 4-year-olds’ (5 females and 5 males) ages ranged from 4;02 to 4;11 (M = 4;06, SD = 0.24). The 5-year-old participants (3 females, 7 males) ranged in age from 5;02 to 5;09 (M = 5;06, SD = 0.19). The 6-year-old participants (4 females, 6 males) ranged in age from 6;01 to 6;11 (M = 6;04, SD = 0.28). The 7-year-old participants (9 females, 1 male) ranged in age from 7;02 to 7;11 (M = 7;07, SD = 0.24).
2.3.2 Materials, Design, and Procedure

The same set of materials as in Experiment 2, translated, was used to test the Korean-speaking children. The procedure was the same as with Latvian-speaking children. All of the children were tested individually by a Korean-speaking experimenter.

2.3.3 Results

2.3.3.1 Results from the main experiment

No participant was excluded from the analysis. Table 2-8 shows that the mean score for the MOY test is higher than that for the DOW test for each age group. The score difference is clearly shown in Figure 2-3.

A repeated-measures ANOVA, with Age as the categorical within-subject factor and Test Type as the between-subject factor, was conducted. There was a main effect of Test Type, $F(1,45) = 9.656, p = 0.003$, showing that Korean-speaking children generally performed better in the MOY test. Age also had a main effect, $F(1,45) = 45.78, p < 0.001$, and regressions showed that the scores produced by different age groups were different, both for DOW questions, $R^2 = 0.74$, $F(1,48) = 139.66, p < 0.001$, and MOY questions, $R^2 = 0.72$, $F(1,48) = 120.51, p < 0.001$.

Table 2-8: Mean scores of DOW (numerical) and MOY (arbitrary) tests in each age group with Korean-speaking children

<table>
<thead>
<tr>
<th>Korean DOW (arbitrary)</th>
<th>3 yrs ($N = 10$)</th>
<th>4 yrs ($N = 10$)</th>
<th>5 yrs ($N = 10$)</th>
<th>6 yrs ($N = 10$)</th>
<th>7 yrs ($N = 10$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.9 ($SD = 2.0$)</td>
<td>10.5 ($SD = 6.0$)</td>
<td>17.8 ($SD = 3.7$)</td>
<td>19.6 ($SD = 3.3$)</td>
<td>22.8 ($SD = 2.1$)</td>
</tr>
<tr>
<td>Korean MOY (numerical)</td>
<td>6.2 ($SD = 4.1$)</td>
<td>10.6 ($SD = 5.9$)</td>
<td>19.9 ($SD = 4.3$)</td>
<td>21.2 ($SD = 3.2$)</td>
<td>23.9 ($SD = 1.5$)</td>
</tr>
</tbody>
</table>
It is worth noting that the terms for “June” and “October” both have idiosyncratic pronunciations in which the final consonant of the numeral is unpredictably lost (yu-wol instead of yuk-wol for “June”; si-wol instead of sip-wol for “October” as shown in Table 2-7). The sound change is a sign of grammaticalization, and it results in nontransparency between month names and numerals. However, Korean-speaking children seem to have no particular difficulty with questions involving these two terms, as the answers to these questions show about the same accuracy rate as the other questions. This shows that Korean-speaking children can somehow relate the two month names with the corresponding numerals despite the slight discrepancy between pronunciations.

2.3.3.2 Results from the pretest and the posttest

All of the 3-year-old participants passed the DCCS task. One out of ten 4-year-old children failed to pass the DCCS test. The scores from the date and hour calculation posttest for the older children were also recorded. Both the 5- and 6-year-old groups had
an average score of 6.9, but with slightly different standard deviations (1.7 for the 5-year olds and 1.6 for the 6-year olds). The 7-year-old group had an average score of 9.1 ($SD = 1.2$).

### 2.3.4 Discussion

The results show a Korean pattern that is the opposite of the Latvian pattern—MOY terms are learned earlier than DOW terms, despite the possibility that month terms occur with lower frequency and there is less exposure to them in daily life. These results provide a piece of strong evidence supporting the hypothesis made at the beginning of this chapter—that the numerically coded calendar terms facilitate children’s acquisition of the temporal terms. Hence, a conclusion can be drawn that the numerical transparency or opacity of the terms probably influences the age at which they are acquired.

### 2.4 General Discussion of Experiments 1, 2, and 3

The series of studies described in this chapter shows that the use of numbers in temporal words can affect children’s acquisition of time structures. Experiment 1 shows that there is a significant difference in the pace at which Chinese-speaking and English-speaking children acquire the ability to name, manipulate, and perform abstract functions on the days of the week (DOW). This is probably because the Chinese system relies on a numerically based set of time-related terms, while the English system uses entirely arbitrary names for the same concepts. Although attempts were made to match the children’s general cognitive ability and mathematical competence across the language groups, there are enough possible confounding cultural, educational, and sampling
variables to prevent me from drawing a definite conclusion without further experiments to tease apart other possible causes for the differences between the groups.

Experiment 2 tested Latvian-speaking children’s acquisition of Latvian’s numerical (Chinese-like) DOW and arbitrary (English-like) MOY terms. The results show that the numerically coded days of the week terms were acquired earlier than the arbitrarily coded months of the year terms. Moreover, Latvian-speaking children were better at distance calculations with DOW than with MOY, probably because of previous knowledge of basic numbers and counting that enabled them to translate DOW questions into arithmetic questions. However, due to the facts that children may well have higher exposure to DOW terms and that the MOY terms would be more difficult to understand given the children’s limited experience, it is possible that the difference that the experiment found was not a linguistic effect.

The test of Korean-speaking children’s performance in Experiment 3 found a similar difference between the two test types, but with the opposite pattern. Korean has arbitrary (English-like) terms for the days of the week, and numerical (Chinese-like) terms for the months of the year. The results clearly show that the Korean-speaking children’s development of time-word proficiency followed a faster, Chinese-like pace for the months of the year, but a slower, English-like pace for the days of the week, proving that the effects found in Experiment 2 were not the result of differences of word frequency or accessibility of the concepts.

The results from the three experiments together are instructive and provide support for the hypothesis proposed in this chapter that when the time sequences make
use of numerical terms, previously acquired knowledge of numbers can facilitate early acquisition of the DOW and MOY.

Some interesting issues emerge from a comparison of the results of the three experiments. First, Figure 2-4 presents the results of the DOW tests for all four language groups together. The complicated situation can be divided into several pair-comparisons.

First, if Chinese and Latvian are compared, it is clear that the Chinese-speaking children scored higher than the Latvian-speaking children, which raises questions because both of the languages refer to the days of the week in a numerical way, and thus the two groups of children should act in similar ways.

The probable reason for this difference is how the weekday terms are coded in both languages. The Chinese DOW terms, as listed in Table 2-9, make direct use of the numbers (Monday is “weekday one,” Tuesday is “weekday two,” etc.). However, the
Latvian DOW terms employ ordinal numbers instead of cardinal numbers, as shown in Table 2-10, and thus this system may be less transparent than the Chinese system. In particular, the terms for the first two days of the week (Monday and Tuesday) are very different from the numbers used in everyday counting. The rest of the DOW names are more similar to cardinal numbers, but the morphological and phonological differences are greater than those between Chinese DOW names and cardinal numbers. As a result, the Latvian-speaking children may have more difficulty in learning the DOW terms than their Chinese-speaking counterparts.

Table 2-9: Chinese ordinal and cardinal numbers compared with DOW terms

<table>
<thead>
<tr>
<th>Chinese DOW</th>
<th>Ordinal numbers</th>
<th>Ordinal numbers</th>
<th>Ordinal numbers</th>
<th>Ordinal numbers</th>
<th>Ordinal numbers</th>
<th>Ordinal numbers</th>
<th>Ordinal numbers</th>
<th>Ordinal numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>weekday1</td>
<td>第一 (dì-yī) 1st</td>
<td>第二 (dì-èr) 2nd</td>
<td>第三 (dì-sān) 3rd</td>
<td>第四 (dì-sì) 4th</td>
<td>第五 (dì-wǔ) 5th</td>
<td>第六 (dì-liù) 6th</td>
<td>第七 (dì-qī) 7th</td>
<td></td>
</tr>
<tr>
<td>weekday2</td>
<td>一 (yī) 1</td>
<td>二 (èr) 2</td>
<td>三 (sān) 3</td>
<td>四 (sì) 4</td>
<td>五 (wǔ) 5</td>
<td>六 (liù) 6</td>
<td>七 (qī) 7</td>
<td></td>
</tr>
<tr>
<td>weekday3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>weekday4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>weekday5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>weekday6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>weekday7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2-10: Latvian ordinal and cardinal numbers compared with DOW terms

<table>
<thead>
<tr>
<th>Latvian DOW</th>
<th>Ordinal numbers</th>
<th>Ordinal numbers</th>
<th>Ordinal numbers</th>
<th>Ordinal numbers</th>
<th>Ordinal numbers</th>
<th>Ordinal numbers</th>
<th>Ordinal numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st day</td>
<td>Pirmās 1st</td>
<td>Otrās 2nd</td>
<td>Trešās 3rd</td>
<td>Ceturās 4th</td>
<td>Piekās 5th</td>
<td>Sestās 6th</td>
<td>Septītās 7th</td>
</tr>
<tr>
<td>2nd day</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd day</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4th day</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5th day</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6th day</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

36
Second, comparing the performance of English- and Latvian-speaking children in the DOW test, the Latvian-speaking children showed a higher average score, as expected, because the Latvian system’s employment of ordinal numbers enabled them to perform better than the English-speaking children, whose language makes no use of numbers. Although their performances were not as good as the English speakers’ in the 4-year-old and 7-year-old groups, the differences are not statistically significant.

Third, the Korean-speaking children’s good performance in the DOW test may raise a question about the hypothesis proposed in this chapter. They performed much better than the Latvian-speaking children in most age groups despite the fact that the Korean DOW terms are arbitrary and the Latvian ones are numerical. If numerically transparent calendar terms facilitate children’s acquisition of the temporal sequences, then Korean-speaking children should not have higher scores on this test.

The unexpected differences, however, probably result from different levels of mathematical ability attained by the two groups. A comparison of the scores of the Latvian-speaking and Korean-speaking children in the posttest (see Table 2-11) suggests that the Korean-speaking children may have been exposed to more intensive mathematical education, as they have higher scores in date and hour calculations in all of the age groups. An analysis of covariance was performed comparing the Latvian-speaking and Korean-speaking 5-, 6-, and 7-year-olds with posttest scores as covariates. The marginally significant result shows that the two groups of children’s math ability to some degree affects their scores in the DOW test, $F (1, 49) = 2.613, p = 0.075$. 

37
Table 2-11: Mean scores in posttest for the Latvian and Korean language groups

<table>
<thead>
<tr>
<th></th>
<th>5 yr. old</th>
<th>6 yr. old</th>
<th>7 yr. old</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latvian</td>
<td>4.3 ($SD = 1.12$)</td>
<td>6.2 ($SD = 2.32$)</td>
<td>7.8 ($SD = 2.32$)</td>
</tr>
<tr>
<td>Korean</td>
<td>6.9 ($SD = 1.66$)</td>
<td>6.9 ($SD = 1.60$)</td>
<td>9.1 ($SD = 1.20$)</td>
</tr>
</tbody>
</table>

In addition, communicating with educators in the two countries provided further clues to help me explain the unexpected differences. Many countries in East Asia, such as China, Japan, and Korea, start mathematical education at an early age. Children learn basic addition and subtraction with numbers between 1 and 10 starting from age 5. Besides the explicit education received in kindergartens and primary schools, many children are also sent to private institutes for extra tutoring. In fact, “over-educating” is coming to be considered a social issue in East Asia. In the year 2012, for instance, the Bureau of Education in Guizhou, China began to enforce policies aimed at preventing the phenomenon of “the kindergartens becoming primary schools”—an acceleration of education that is forcing children to engage with material meant for older children. In contrast, Latvian-speaking children at the age range being tested generally receive less intensive education in kindergarten and primary schools. More physical activities are involved in the curriculum. It is very probable that the Latvian-speaking 3-year-old group had not been exposed to extensive or explicit mathematical education, and that the Korean-speaking children had.

Such educational differences might be the main reason for the differences observed in the comparison. It is still clear that the effect of numerically coded time terms is probably the main reason for earlier acquisition of the time terms because within
a single language, numerical calendar terms are obviously acquired earlier than the arbitrarily named ones, as shown by both the Korean and Latvian experiments.

To conclude, the experiments described in this chapter provide evidence that calendrical naming systems that differ in their use of numbers may result in different acquisition ages for time words. However, further MOY tests with Chinese- and English-speaking children should be conducted to set a baseline for cross-experiment comparisons. Also, more information on educational and cultural factors needs to be collected for better control of possible confounding variables.
Because the acquisition studies described in the previous chapters have been producing promising preliminary results, I decided to follow up on them with behavioral studies with adults that aim at investigating whether adults’ temporal reasoning abilities differ depending on the transparency of the naming systems that their languages employ for time sequences. The question is whether the naming system effect carries over into adulthood in such a way that it affects performance of time calculations.

Do different developmental trajectories caused by differences in the naming of time concepts have lingering effects on adults? On one hand, adults in many cultures make use of similarly formatted calendars when performing time-related calculations. If, as Friedman (1990) argues, adults perform time calculations using mental images of calendars, and calendars are similar across cultures, then cross-linguistic differences based on naming system differences might dissipate in adulthood. On the other hand, it could be that the developmental origin of time concepts affects subsequent time calculations, throughout the lifespan.

Previous cross-linguistic studies have shown that there is at least some causal influence from language on non-verbal cognition and unconscious habitual thought (e.g., Kay & Kempton, 1984; Lucy, 1992; Gumperz & Levinson, 1996). 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, 2009).
The transparent numerical structure of the Chinese calendar might facilitate calendar calculation, causing Chinese-speaking adults to outperform their English-speaking counterparts in time calculation tasks, perhaps by exhibiting shorter reaction times and making fewer errors. In general, such a finding would support 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 & Sweetser, 2006; Casasanto & Boroditsky, 2008).

The current research tests whether these hypothesized differences in performance are due to different strategies employed by adults when they perform temporal calculations, again, dependent on the transparency of their time concept naming systems. The chapter presents an experiment that tests whether differences in temporal cognition persist in adulthood in native speakers of languages with different numerical time-word transparency. If Chinese and English speakers indeed perform differently in temporal calculation tasks, what might be the cause? The experiment tests strategies in adult English and Chinese speakers’ calculations in days of the week and months of the year tasks.

3.1 Experiment 4

Again, there has been a limited amount of work on this question. Kelly et al.’s (1999) study provides essentially the only systematic 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, which is very much in line with what
one would expect, given the child data I have discussed in the previous chapters.
However, this argument about the mechanisms used was based on participants’ self-
reports of calculation strategies after participating in the experiment. While self-reports
are an excellent seed for hypotheses about underlying cognitive mechanisms, they do not
replace carefully designed experiments that access people’s unconscious cognitive
operations through observation of indirect measures. The studies described in this
chapter build on Kelly et al.’s experimental methods to get a handle on the mechanisms
underlying adults’ performance of temporal reasoning tasks, and how these mechanisms
vary with native language.

Kelly et al. (1999) report 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 Chinese speakers, this directly evokes
simple 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 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, such as list reciting.
Alternatively, however, this same result could be the product of cultural differences, for
instance, in the depth, length, and nature of the math education they received. It might be that the Chinese population that was sampled simply was better at performing mental calculations than their English-speaking counterparts.

It is possible to tease apart these two possibilities with a careful experimental design. If the use of different arithmetic strategies by Chinese speakers is responsible for the performance difference, the advantage may disappear in cross-week or cross-year calculations. That is, Chinese speakers may encounter difficulties in calculating distances in cross-boundary calculation because 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.” English speakers, meanwhile, may not be using this same strategy, and if they are not, then they should exhibit less difference in difficulty when calculating within or across boundaries. A second prediction based on the hypothesis that Chinese speakers use arithmetic more than English speakers is that the calculation of distances involving Sunday (which is called “weekday sun” in Chinese) should be more difficult for Chinese speakers than calculations with any other day in the week, as a number is not used in naming this day (see Table 1-1). This irregularity may cause trouble in applying the arithmetic strategy, thus slowing down the calculations of Chinese speakers, compared to calculations starting relative to other days of the week. Again, by contrast, English speakers should show no increased difficulty when making calculations relative to Sunday.

The experiment therefore tests for differences between Within and Across Boundary calculations in Chinese and English speakers, to ascertain whether the Chinese
speakers are using an arithmetic strategy in their calculations due to their language’s employment of number systems. If they exhibit better performance than English speakers only when boundaries are not crossed, and when Sunday is not involved, then this suggests that it is not differences between the populations in general, but rather differences in the strategies adopted for specific time calculations in particular that cause adult differences in temporal reasoning.

3.1.1 Participants

Thirty-two native speakers of Chinese and 40 native speakers of English participated in the experiment. The Chinese speakers were undergraduate and graduate students recruited from Beijing United University, who ranged in age from 18 to 31 ($M = 21, SD = 4.9$). All of them (22 females and 10 males) were native speakers of Mandarin, although they had studied English for 7 to 15 years ($M = 10.97, SD = 2.12$). The English speakers (19 females and 21 males) were undergraduate students at the University of Hawai‘i, aged from 18 to 29 ($M = 21, SD = 2.5$). Some of these English speakers have language backgrounds that include Ilocano, Tagalog, or German, but none of them have studied more than two years of any language that makes use of numbers in its calendar system, such as Chinese, Japanese, Korean, or Vietnamese. All of the participants either received extra credit in an introductory linguistics class or were paid five dollars or the equivalent amount for their time.

---

3 Because English language is a required subject for all Chinese students from primary school to the second year of university level, it is impossible to find true monolingual participants at this age range. The English courses range from about 4 hours to 6 hours per week and the participants use only Chinese in their daily lives, so their exposure to English is unlikely to have been intensive enough to affect the Chinese participants’ performance in this experiment.
3.1.2 Materials and Design

The stimuli consist of two trial blocks: DOW (days of the week) and MOY (month of the year). Each block has two factors with two levels—Boundary Type (Within/Across Boundary) and Direction (Forward/Backward), which produces four question types for DOW and MOY. Examples of the questions are given in (1) through (4). In addition to these two within-subjects factors, Language is a between-subjects factor. There is also an additional factor, Sunday, used only for the DOW blocks. Half of the DOW questions involve Sunday, to show the effects of crossing a boundary versus the effects of Sunday.

(1a) If today is Wednesday, 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)
(3a) If this month is January, four months from now it’ll be what month? (Within/Forward)
(3b) If this month is November, six months ago was what month? (Within/Backward)
(4a) If this month is September, five months from now it’ll be what month? (Across/Forward)
(4b) If this month is March, five months ago was what month? (Across/Backward)

The questions for the English language group were formulated in this way instead of the usual word order (for example, “If today is Wednesday, what is the day three days from now?” and “If this month is January, it’ll be what month four months from now?”) to match the word order of the Chinese stimuli. In Chinese, the temporal distance has to appear before the temporal preposition (“ago”/“from now”) in a question. This wording
ensured that the Chinese-speaking group would not get the temporal distance information earlier than the English-speaking group.

The calculation distance is also controlled in the experiment. The calculation distances are 1 to 7 for DOW questions (1–4 for Within questions and 2–7 for Across questions) and 2 to 12 for MOY questions (2–10 for Within questions and 2–12 for Across questions) in order to match the cyclical nature of the weeks and months. No distances of 1 are included in the Across Week calculations, because the Chinese and English week systems begins with Monday and Sunday, respectively, meaning that distance 1 questions would be regarded as a Within question for one group and an Across question for the other one. Also, distance 1 was excluded from MOY calculations due to the possible confounding of spatial reasoning for Chinese speakers. “Next month” and “previous month” are expressed as “down month” and “up month” in Chinese, which has been shown to affect Chinese speakers’ reasoning (Boroditsky, 2001). To avoid any unnecessary interaction of spatial reasoning, distance 1 calculations were not tested. There are 32 questions in the DOW block and 48 in the MOY block.

3.1.3 Procedure

All English-speaking participants were tested in a sound booth and all Chinese-speaking participants were tested individually in a small language learning classroom at their university. In the experiment, participants were seated in front of a computer. Participants initiated each trial by pressing the SPACE bar when they were ready, and then a recorded voice read one question to them, which they heard over speakers. They were instructed to speak the answer into a microphone as quickly and accurately as
possible after hearing the question. The microphone was used as a voice key and connected to an E-Prime SR-BOX in order to record reaction times (from the offset of the question to the onset of the answer). A digital recorder was used to record the answers for later coding and scoring.

Each trial began with a screen asking participants to press the SPACE bar to listen to the question. As soon as they spoke, there was a fixation cross in the center of the screen lasting for 1000 milliseconds for the participants to articulate the answer, followed by a pause of 2000 milliseconds before they could press the SPACE bar for the next question. The voice key was tested and adjusted for each individual participant before and after the practice session.

Four practice items preceded the experiment, and the experiment was divided into two parts, one for the DOW block and the other for the MOY block. The order of the blocks’ presentation was counterbalanced across participants. The questions were asked randomly within each of the blocks. There was a short break between the two parts.

3.1.4 Results

All filler syllables, tongue clicks, partial responses, and repeated responses (due to their failure to trigger the voice key) were targeted and excluded. In addition, because the Chinese weekdays all begin with xīngqī (weekday), which is then followed by a number, it is possible for speakers to pause after they say xīngqī and before giving the number. These kinds of responses were also excluded.

No participants were excluded because of outlying RTs or low accuracy. Reaction times and error rates were calculated as dependent measures for comparison. In
the coding of accuracy rates, again, I noticed that there were two possible ways to calculate temporal distances. For example, some participants would consider two days after Monday to be Wednesday, but others would consider it Thursday; two months before April could be considered February or January. Two Chinese participants did the “plus one” calculation in the experiment, and their data were excluded from analysis. No English participants tested showed this calculation pattern.

3.1.4.1 Error analysis

Let’s first take a look at the accuracy data. As discussed, the divergence in the naming systems of the two calendar lists in Chinese may result in Chinese speakers having different representations of the systems and using different strategies in the calculation tasks. A consistent arithmetic strategy would work with MOY questions but not with DOW questions because of the non-numerical Sunday term. And the possible use of different strategies may cause differences in the accuracy with which the Chinese speakers performed the calendar calculation tasks. In contrast, English-speaking participants should not show this difference because there are consistent arbitrary naming systems for both day and month terms in English. For this reason, the DOW and MOY data are examined separately.

**MOY task.** Figure 3-1 shows the accuracy rate for each Boundary Type and Direction condition by language in the MOY task. Individual proportions of correct answers were arcsine transformed to comply with the normal distribution premise. Then two repeated-measures ANOVAs were performed, one each with participants and items as random
factors. In the by participant analysis, the data was analyzed by Language (Chinese/English), Direction (Forward/Backward), and Boundary Type (Within/Across) with Direction and Boundary Type as within-subjects variables. For item analysis, Language is taken as a within-item variable and the other two (Boundary Type and Direction) are between-items variables.

There is a large main effect of Language, $F_{1}(1,68) = 26.846, p < 0.001$, $F_{2}(1, 44) = 54.265, p < 0.001$, which shows that the Chinese speakers consistently made fewer errors than the English speakers. There is a main effect of Boundary Type, $F_{1}(1, 68) = 81.614, p < 0.001; F_{2}(1, 44) = 22.262, p < 0.001$; Within questions have a higher accuracy rate than Across questions. It should not be surprising to find that calculating across week or year boundaries might be more difficult. Also, Forward calculations have a higher accuracy rate than Backward ones, $F_{1}(1, 68) = 103.934, p < 0.001; F_{2}(1,44) = 21.143, p < 0.001$. 
There is also an interaction effect between Boundary Type and Direction, $F_1(1, 29) = 24.587, p < 0.001; F_2(1,44) = 6.886, p = 0.012$. To confirm the interaction effect, separate tests were performed for both Within and Across conditions. For Within questions only, a main effect of Direction was found, $F_1(1, 68) = 16.698, p < 0.001; F_2(1,44) = 7.492, p = 0.012$. For Across questions only, there was also a main effect of Direction, but of much greater magnitude, $F_1(1, 68) = 142.585, p < 0.001; F_2(1, 22) = 14.636, p = 0.001$. In both boundary types, Forward questions have a higher accuracy rate than Backward questions, but the difference was much more obvious in the Across Boundary questions.

In addition, there is 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 indicated by Figure 3-1, the two language groups presented different patterns in the two boundary types. I performed 2 (Boundary Type) X 2 (Direction) repeated-measures ANOVAs separately for the Within and Across Boundary questions in order to get a clearer picture. As predicted, Chinese speakers have significantly higher accuracy rates than English speakers in answering Within Boundary questions, $F_1(1, 68) = 70.088, p < 0.001; F_2(1, 22) = 84.138, p < 0.001$. But this advantage disappears in answering Across Boundary questions, $F_1(1, 68) = 0.146, p = 0.703; F_2(1, 22) = 0.091, p = 0.766$.

**DOW task.** Figure 3-2 shows the accuracy rate for each Boundary Type and Direction condition by language in the DOW task. Again, individual proportions of correct answers were arcsine transformed before any statistical analysis was performed. Another
set of repeated-measures ANOVAs was performed using the same methods as with the MOY task.

![Figure 3-2: Accuracy rate on the DOW task for Chinese and English speakers](image)

As we can see from Figure 3-2, the accuracy rate analysis results of the DOW calculation task present a more complicated picture than those of the MOY task.

First, unlike what was found in the MOY task, Language in the DOW task did not show a significant main effect in the participants analysis, though there was a significant main effect in the items analysis, $F_2(1, 28) = 5.177, p = 0.031$. This indicates that Chinese speakers were not showing a general accuracy advantage over English speakers, but that an advantage exists for only some of the questions. As Figure 3-2 shows, the lack of Language effect was probably driven by the Chinese speakers’ poor performance in the Across Boundary condition. The Chinese speakers were more accurate than the English speakers in answering Within Week questions, $F_1(1, 68) = 11.476, p < 0.001$, $F_2(1, 14) = 21.182, p < 0.001$, but this was not the case in the Across condition.
The analysis also found a main effect of Boundary Type, $F_1(1,68) = 73.018, p < 0.001; F_2(1,28) = 28.414, p < 0.001$, meaning that Within questions have a higher accuracy rate than Across questions, and a main effect of Direction, $F_1(1,68) = 66.237, p < 0.001; F_2(1,28) = 9.041, p = 0.006$, showing that Forward questions have a higher accuracy rate than Backward questions.

There was also an interaction effect between Boundary Type and Direction that was significant by participant and marginally significant by item, $F_1(1,68) = 20.235, p < 0.001; F_2(1,28) = 3.276, p = 0.081$. The contrast of accuracy for Forward and Backward questions was bigger in the Across condition, $F_1(1,68) = 55.772, p < 0.001; F_2(1,14) = 8.532, p = 0.011$, than in the Within condition, $F_1(1,68) = 13.775, p < 0.001; F_2(1,14) = 1.118, p = 0.308$.

Another interaction effect was found between Language and Boundary Type, $F_1(1,68) = 18.537, p < 0.001; F_2(1,28) = 24.596, p < 0.001$. For the Chinese speakers only, there was a large main effect of Boundary Type, $F_1(1,29) = 52.258, p < 0.001; F_2(1,28) = 51.806, p < 0.001$. It is obvious that they did much better in the Within Week calculations. For English speakers, there was also a main effect of Boundary Type, but of much smaller magnitude than that of the Chinese group in the by participant analysis $F_1(1,39) = 14.661, p < 0.001$, and only marginally significant in the by item analysis $F_2(1,28) = 3.353, p = 0.078$. This suggests that the English speakers were more accurate on Within Boundary than Across Boundary questions, but the difference was not as obvious as it was with the Chinese speakers.

There was also a three-way interaction among Language, Boundary Type, and Direction significant in the participants analysis, $F_1(1,68) = 8.96, p = 0.004$. This three-
way interaction suggested looking at the data in a more detailed way, and I did so by first analyzing the two language groups separately. For the Chinese speakers, there was an interaction effect between Boundary and Direction, $F_1(1, 29) = 21.007, p < 0.001, F_2(1, 28) = 5.079, p = 0.032$, as they performed more accurately in Forward questions in the Across Boundary condition, $F_1(1, 29) = 5.125, p < 0.001, F_2(1, 14) = 8.567, p = 0.011$, but their performances were not significantly different for the two directions in the Within condition, probably due to a ceiling effect. For the English speakers, however, no interaction effect between Boundary and Direction was found, which is a different pattern than that of the Chinese language group—English speakers were consistently better in answering Forward questions, whether they were Within Boundary, $F_1(1,39) = 3.888, p < 0.001; F_2(1,14) = 4.108, p = 0.039$ or Across Boundary calculations, $F_1(1,39) = 5.251, p < 0.001, F_2(1,14) = 5.347, p = 0.036$. In other words, Boundary Type did not affect English speakers’ calculations.

**Questions involving Sunday.** Comparing the results from MOY and DOW tasks, it is obvious that the Chinese speakers lose their advantage in the weekday calculations, as the main effect of Language disappears in the latter condition.

As described in Section 3.1.2, the DOW block has an extra factor compared to the MOY block—the Sunday factor. Half of the DOW questions were Sunday-related questions, and the natural expectation is that the Sunday questions were the cause of the Chinese speakers’ low DOW accuracy because the irregular Sunday term (“weekday sun” instead of “weekday seven”) would affect the use of the arithmetic strategy in calculating.
To test this hypothesis, it is necessary to perform some extra statistical tests considering Sunday as a factor.

Looking at the two language groups separately, Chinese speakers had a significantly lower accuracy rate 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.

![Figure 3-3: Chinese- and English-speaking participants’ accuracy rate in different Sunday conditions](image)

The non-numerical term for Sunday used in Chinese thus may affect Chinese speakers’ processing and calculation, as they made more mistakes in answering questions involving Sunday. On the other hand, the English speakers did not show any significant differences in these two question types.
**Kinds of errors made.** The kinds of errors made by the English and Chinese speakers may provide some indication of the strategies being employed by the two groups.

One of the common errors that occurred with a majority of Chinese participants was that they were off by two in the month calculations. For example, when hearing the question “If this month is May, seven months ago it was what month?” some Chinese speakers said “eight month” (August) while the correct answer should be “ten month” (October). Likewise, for the question “If this month is July, eight months from now it’ll be what month?” a frequently found error was “five month” (May) while the correct answer should be “three month” (March). This kind of missed-by-two error occurred widely among participants, and it actually makes up 19.3% of the mistakes in the MOY calculations. In addition, it occurred regardless of difficulty (temporal distance to be calculated), which month was calculated from, or the order in which the question appeared in the experiment. This provides a clue that the Chinese speakers might be using the arithmetic strategy; under the time pressure, they sometimes might have forgotten that the month system is base-twelve instead of base-ten. As a result, they added or deducted the two numbers of the months in the question directly, which led to what I will call base-ten errors. If these errors were not caused by the use of the arithmetic strategy, it would be very difficult to explain them. In daily life, most of the measure systems used by Chinese, such as length, weight, volume, etc., are all metrical, hence base-ten, and it is understandable that under time pressure the Chinese speakers would sometimes make mistakes in choosing the right system. However, it is interesting that this type of error did not widely occur in the Chinese speakers’ days of the week calculations (only one participant made this kind of error at the beginning of the DOW
task, and corrected it quickly). Possibly, because the base-seven system used for the weekdays is unique, it would not easily be mistakenly replaced by a base-ten system. Or it could be that the non-numerical term for Sunday would remind the Chinese participants that the week system was not purely numbered, let alone based on a regular base-ten system.

Another kind of error made by Chinese speakers can be called transferring errors. One Chinese speaker said “thirteen month” when the correct answer would be “one month” (January) and another said “day nine” when the correct answer would be “weekday two” (Tuesday); neither of these are valid day or month names in Chinese. However, these answers would have been correct if the Chinese day and month cycles did not have to be converted into modulo 7 and modulo 12 systems. Therefore, this type of transferring error implies an overly general transfer of an arithmetic strategy.

English speakers’ errors, on the other hand, usually involved missing the correct answer by one. This type of missed-by-one error may be due to the fact that the English speakers were listing the months or days to count them. If this is the case, then it is natural to find an increasing error rate with longer distances, especially when the participants are trying to give answers as quickly as possible, as in the experiment. This is confirmed by comparison of accuracy rates for questions with different temporal distances to be calculated. Two pair-wise comparisons were conducted for month of the year calculations. As discussed in Section 3.1.2, temporal calculation distances were controlled to be 2 to 12 for MOY tasks. In the pair-wise tests, the distances were

---

4 MOY, instead of DOW, questions were analyzed to uncover the effect of temporal calculation distances because the distance in the days of the week task was limited to 1–6, and there were fewer items in the DOW task. The month of the year task provides more possibility of observing any effect of distance.
categorized as high difficulty (7–11) or low difficulty (2–6), which together form a within-subject factor Difficulty. Distance 12 was not included in this comparison because it is not a calculation of 12 in a strict sense. Results were as expected; English speakers were more accurate with shorter calculation distances with both Within, \( t(1, 39) = 1.98, p = 0.05 \) and Across questions, \( t(1, 39) = 2.33, p = 0.025 \), indicating that they were having problems counting a longer distance without making mistakes. In contrast, the Chinese speakers did not show any difference in calculating shorter or longer temporal distances with either the Across or Within questions. This result further proves that the Chinese speakers might be using an arithmetic strategy, as performing addition or subtraction of numbers from 7 to 11 would not be much more difficult than doing so with the smaller numbers from 2 to 6.

This error analysis provides promising results that support the idea that the two populations might be using different strategies in solving time-related questions according to how time terms are coded in the language they speak. The numbered day and month names in Chinese give its speakers an advantage in temporal distance calculations, as Chinese speakers are able to directly map days and months to numerals and thus use arithmetic methods in calculations. The direct mapping helps the Chinese speakers solve the problems with fewer errors. However, this advantage disappears when the calculation is across a week or year boundary, as the Chinese speakers have to convert the answer into modulo 7 or modulo 12 systems. Another factor that lowers Chinese speakers’ accuracy is the involvement of Sunday, whose name violates the transparent number-to-time system and thus affects Chinese speakers’ performance—they have more difficulty calculating temporal distances involving Sunday.
3.1.4.2 Reaction time analysis

The error analysis showed an interesting picture of the different performances of the two populations in temporal calculation tasks. In this section, I add to this picture by analyzing reaction time data as a way to explore possible differences in the groups’ strategies. As suggested by the error analysis, I treat the DOW and MOY tasks separately. I have removed all trials with incorrect responses, which excluded 16.7% of the data. I 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 cleaning.

**MOY task.** The results displayed in Table 3-1, and presented graphically in Figure 3-4, show the reaction times for each Boundary Type and Direction condition by language in the MOY task. As expected, the Chinese speakers had overall shorter reaction times in all conditions, showing that there was no speed-accuracy trade-off with the Chinese speakers—they had higher accuracy rates as well as shorter reaction times.

<table>
<thead>
<tr>
<th>Language</th>
<th>Boundary Type</th>
<th>Direction</th>
<th>Mean RT (ms)</th>
<th>SD (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td>Within</td>
<td>Forward</td>
<td>708</td>
<td>341</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backward</td>
<td>760</td>
<td>313</td>
</tr>
<tr>
<td></td>
<td>Across</td>
<td>Forward</td>
<td>1392</td>
<td>582</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backward</td>
<td>1820</td>
<td>843</td>
</tr>
<tr>
<td>English</td>
<td>Within</td>
<td>Forward</td>
<td>2147</td>
<td>1121</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backward</td>
<td>3554</td>
<td>2319</td>
</tr>
<tr>
<td></td>
<td>Across</td>
<td>Forward</td>
<td>2318</td>
<td>1044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backward</td>
<td>3074</td>
<td>1673</td>
</tr>
</tbody>
</table>
Reaction times for correct responses for the month of the year calculation task were analyzed in 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. Item analysis was also performed, in which Language was taken as a within-item variable and Boundary Type and Direction were between-items variables.

Results were in line with the error analysis. There was a large main effect of Language, $F_1(1, 68) = 36.155, p < 0.001, F_2(1, 44) = 134.986, p < 0.001$, confirming the observation that the Chinese speakers were consistently faster than their English-speaking counterparts. Boundary Type also showed a significant main effect, $F_1(1, 68) = 17.515, p < 0.001, F_2(1, 44) = 9.098, p = 0.004$, indicating that Within questions were answered more quickly than Across ones. There was also a main effect of Direction, $F_1(1, 68) =$
33.134, $p < 0.001$, $F_2(1, 44) = 13.876$, $p = 0.001$; the Forward task was faster than the Backward task.

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 two interaction effects confirmed two observations: (1) Chinese speakers were faster in Within Year questions than Across Year questions, as shown by a large main effect of Boundary Type, $F_1(1, 29) = 76.919$, $p < 0.001$, $F_2(1, 44) = 35.523$, $p < 0.001$. This pattern is as predicted, because no modulo calculation was necessary in the Within Boundary questions if the Chinese speakers were using an arithmetic strategy, while the performing of modulo calculations in the Across Boundary questions would slow down their answering. On the other hand, English speakers spent about the same amount of time answering Within and Across Year questions, as shown in Figure 3-4; and (2) English speakers were much faster in answering Forward questions, as shown by the main effect of Direction, $F_1(1, 39) = 31.226$, $p < 0.001$, $F_2(1, 44) = 14.877$, $p < 0.001$, whereas Chinese speakers did not show this contrast in a significant way.

A three-way interaction among Language, Boundary Type, and Direction was also found significant in the participants analysis, $F_1(1, 68) = 15.486$, $p < 0.001$. Further tests were performed with the two language groups separately in order to look into this three-way interaction. For the English speakers, they were consistently faster in answering Forward questions, in both Within and Across Year conditions, as there was no interaction between Boundary Type and Direction found. On the other hand, Chinese speakers were showing different patterns in answering Forward and Backward questions.
in the two Boundary types, as confirmed by an interaction between Boundary Type and Direction, $F_1(1, 29) = 11.845, p = 0.002, F_2(1, 44) = 2.652, p = 0.1$. More specifically, they were faster with Forward than Backward questions only in the Across Boundary task, $F_1(1, 29) = 3.905, p = 0.001, F_2(1, 44) = 3.329, p = 0.082$, but this direction advantage disappeared for Within Year questions.

**DOW task.** The results of the days of the week task are presented in Table 3-2 and Figure 3-5. As the accuracy analysis showed, the data from the DOW task are more complicated and provide more information on the strategies used by the two language groups in temporal calculations.

<table>
<thead>
<tr>
<th>Language</th>
<th>Boundary Type</th>
<th>Direction</th>
<th>Mean RT (ms)</th>
<th>SD (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td>Within</td>
<td>Forward</td>
<td>667</td>
<td>268</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backward</td>
<td>664</td>
<td>241</td>
</tr>
<tr>
<td></td>
<td>Across</td>
<td>Forward</td>
<td>1200</td>
<td>490</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backward</td>
<td>1571</td>
<td>666</td>
</tr>
<tr>
<td>English</td>
<td>Within</td>
<td>Forward</td>
<td>1093</td>
<td>543</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backward</td>
<td>1561</td>
<td>888</td>
</tr>
<tr>
<td></td>
<td>Across</td>
<td>Forward</td>
<td>1598</td>
<td>678</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backward</td>
<td>2614</td>
<td>1371</td>
</tr>
</tbody>
</table>

The Chinese speakers were consistently faster than their English-speaking counterparts, as shown by a main effect of language, $F_1(1, 68) = 20.617, p < 0.001, F_2(1, 28) = 63.776, p < 0.001$. The Within questions were answered faster than the Across ones, and the Forward questions were answered faster than the Backward ones, as shown by the two main effects of Boundary Type, $F_1(1, 68) = 161.850, p < 0.001, F_2(1, 28) = 27.446, p < 0.001$, and Direction, $F_1(1, 68) = 42.243, p < 0.001, F_2(1, 28) = 11.496, p = 0.002$. 
Figure 3-5: Reaction times on the DOW task for Chinese and English speakers

There was also an interaction effect between Boundary Type and Direction significant by participants, $F_1(1,68) = 26.585, p < 0.001$, but not by items. Reaction times for Forward and Backward questions showed different patterns in the two boundary types. For the Across condition, the Forward questions were answered more quickly than the Backward ones, as confirmed by a main effect of Direction, $F_1(1,68) = 46.498, p < 0.001, F_2(1,14) = 9.917, p = 0.007$. For the Within task only, Forward questions were also answered faster than the Backward ones, but the difference was of much smaller magnitude compared to that in the Across Boundary condition—it is significant by participant, $F_1(1,68) = 14.173, p < 0.001$, but not by item.

There was also an interaction effect between Language and Direction, $F_1(1,68) = 15.407, p < 0.001, F_2(1,28) = 9.244, p = 0.005$. For the English speakers, there was a large main effect of Direction, $F_1(1,39) = 86.977, p < 0.001, F_2(1,28) = 13.090, p = 0.001$—they were much faster with Forward questions. Chinese speakers were also faster
in answering Forward questions than Backward questions, but the difference was much smaller, $F_1(1,29) = 13.460, p = 0.001, F_2(1,28) = 4.933, p = 0.035$.

I also found an interesting, unexpected interaction effect with English speakers, an interaction between Boundary Type and Direction significant by participants, $F_1(1,39) = 15.221, p < 0.001$, but not by items. This discrepancy between participants and items analyses needs to be checked more closely by looking into the English speakers’ performance in the two boundary conditions separately. The Direction effect for the Across condition was consistent in the by participant and by item analyses, $F_1(1,39) = 38.862, p < 0.001, F_2(1,14) = 12.071, p = 0.004$—they were much faster with Forward questions when the questions were Across Week. This is in line with the prediction made earlier in this chapter. Because reciting the list backwards was much more difficult than reciting it forwards, if English speakers were using the list-reciting strategy, they would be much faster in the Forward questions, which would drive the Direction effect.

However, the difference between the two Directions was not significant in the items analysis in Within Week calculations, although it was significant by participants $F_1(1,39) = 20.341, p < 0.001$. This difference in significance in the item analysis and the participant analysis suggests that the Direction effect found with the Within Week calculations was driven by only some of the items. In other words, some of the Within Week questions were solved with approximately the same speed, despite the direction. Because the calculation distance for Within Week questions was controlled to be from 1 to 4, it is highly likely that the times spent for short distance (such as 1 and 2) calculations were about the same, no matter whether the calculation was forward or
backward. More discussion on this issue can be found in the next subsection—Section 3.1.4.3.

**Questions involving Sunday.** The analysis of the accuracy data showed that the Chinese speakers had problems with questions involving Sunday, the only irregularly named weekday in Chinese, in that they made more errors with questions that had Sunday as the starting point of calculation and with questions that had Sunday as the correct answer. As a result, compared to their English-speaking counterparts, the Chinese speakers had a lower accuracy rate in answering questions involving Sunday. However, the difficulty induced by Sunday that Chinese speakers encountered did not affect their reaction times negatively. As Table 3-3 and Figure 3-6 show, the irregularity of Sunday’s name did not slow down Chinese speakers’ calculations in general, though it did result in more incorrect answers. The result of a two-way ANOVA shows that there is a main effect of Language, $F_1(1,68) = 15.681, p < 0.001, F_2(1, 30) = 34.942, p < 0.001$; the Chinese speakers were consistently faster than the English speakers in questions either with $(F_1(1,68) = 18.227, p < 0.001, F_2(1,13) = 18.214, p = 0.001)$ or without $(F_1(1,68) = 9.265, p = 0.03, F_2(1,17) = 14.683, p < 0.001)$ Sunday. There is also an interaction between Language and the Sunday factor that is significant by participant and marginally significant by item, $F_1(1,68) = 10.632, p = 0.002, F_2(1, 30) = 3.611, p = 0.067$.

Table 3-3: Mean RT and SD in Sunday and Non-Sunday conditions for Chinese and English speakers

<table>
<thead>
<tr>
<th>Language</th>
<th>Sunday involved</th>
<th>Mean RT (ms)</th>
<th>SD (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td>No</td>
<td>1615</td>
<td>555</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>1513</td>
<td>735</td>
</tr>
<tr>
<td>English</td>
<td>No</td>
<td>2325</td>
<td>878</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>2459</td>
<td>1308</td>
</tr>
</tbody>
</table>
It is interesting that the Chinese speakers were actually faster when Sunday was involved, $F_1(1,29) = 13.016, p = 0.001$, and the English speakers performed in the opposite way—they were faster when there was no Sunday involved, $F_1(1,39) = 3.042, p = 0.004$. However, neither of these effects was significant by item, suggesting that the effect was only driven by some of the items.

Table 3-4: Summary of results in error and reaction time analyses

<table>
<thead>
<tr>
<th></th>
<th>Lg</th>
<th>Boundary</th>
<th>Direction</th>
<th>Lg x Boundary</th>
<th>Lg x Direction</th>
<th>Boundary x Direction</th>
<th>3-way</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOY</td>
<td>Error Analysis</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>RT Analysis</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>DOW</td>
<td>Error Analysis</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>RT Analysis</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

1. significant only in items analysis; 2,3,4 significant only in participants analysis
So far, the error analysis and reaction time analysis have revealed a lot of instructive results. For a clearer big picture, the results are summarized in Table 3-4.

Most of the effects found were consistent and as predicted, such as the main effects of Language, Boundary, and Direction. However, some discrepancies appear, as shown in Table 3-4. First, the interaction between Language and Direction was found in the reaction time analysis but not the error analysis for both MOY and DOW tasks, suggesting that Chinese and English speakers showed similar accuracy rate patterns in answering Forward and Backward questions. However, their speed in answering the questions of the two directions was different. More specifically, Forward questions were answered more accurately by both Chinese and English speakers; however, Chinese speakers spent about the same amount of time on the questions in both directions, whereas the English speakers were much faster in solving Forward than Backward questions. In other words, although Chinese speakers were equally fast in answering Forward and Backward questions, they made more mistakes in Backward calculations. And it is the Chinese speakers’ performance that leads to the discrepancy of the interaction effects between Language and Direction observed in the error and reaction time analyses.

A similar discrepancy was found with the interaction between Boundary Type and Direction. It is significant in the error analysis but not in the RT analysis (or only significant in the by participant analysis). For accuracy, participants were much better at answering Forward questions in the Across Boundary condition, but the contrast was not as obvious in the Within Boundary condition. For speed, they were in general faster with Forward questions, both in the Within and Across calculations.
And further, the discrepancies of the two interaction effects described above result in differences in the three-way interactions. Nonetheless, I was not expecting the accuracy data to fully match the reaction time data, or the MOY data to match the DOW data. Higher error rates in certain conditions may or may not accompany slower responses. Also, people may have different mental representations of the month and week cycles, and thus different strategies for solving questions in these two systems. But it is somewhat surprising not to find an interaction between Language and Boundary Type in the DOW reaction time analysis.

In Section 3.1, I predicted that because of the use of different strategies, there should be an interaction between Language and Boundary Type. More specifically, I thought that the Chinese speakers’ advantage from using the arithmetic strategy would result in quicker responses in Within Week calculations, but the advantage might be weakened or even lost in Across Week calculations because they have to perform modulo calculations, which might well take longer. In other words, because Chinese speakers were probably employing the arithmetic strategy, their reaction times should be sensitive to Boundary Type. On the other hand, I predicted that the English-speaking participants would not show any difference in reaction times according to Boundary Type, as there should be no difference in the time it takes to recite a list within or across a week boundary. And the prediction should hold true for both MOY and DOW calculations.

It was, then, surprising to not find this interaction effect in the reaction time analysis of the DOW task, especially as it was found in the error analysis. Given the reaction time data, the Chinese speakers had longer reaction times in the Across Week condition than in the Within Week condition, as predicted. However, the English
speakers also, unexpectedly, showed this reaction time difference between the two conditions—they were much slower in the Across Week condition than in the Within Week condition. If the English speakers were using the list-reciting strategy, as predicted, then there should not be much reaction time difference from reciting within or across the week boundary, just as in the MOY task, where there was no Boundary Type effect.

There are probably two reasons why the English speakers show a Boundary effect in DOW calculations. First, it might be because of the calculation distance difference between Within and Across Boundary questions. As described in Section 3.1.2, the calculation distance was 1–4 for Within Week and 2–7 for Across Week questions. The longer distances to be calculated in the Across Boundary condition might have caused longer reaction times, which led to the unexpected Boundary effect. When I excluded the shortest distance (distance 1) from the Within condition and the longest distances (distance 5 to 7) from the Across Boundary data and reran the analysis, the Boundary effect disappeared, confirming that distance is a contributing factor in the Boundary effect.

Second, another implication might be that English speakers were not uniformly using the list-reciting strategy, at least for some of the items, in performing weekday calculations. The Direction effect with the English speakers helps clarify this issue. As mentioned earlier in this section, we observed a significant Direction effect consistent in participants and items analyses in Across Week calculations, but not in the Within condition (at least not for all of the items in this condition), which suggests that English speakers were not using the direction-sensitive list-reciting strategy in answering some of the Within Week questions. If they had been, they should have been much faster in
answering Forward questions than Backward questions, because it is much easier to recite a list forward than backward. So it could be that for the questions involving Within Week, short distance calculations, English speakers could visualize the calendar and perform the Within Week calculations because the distance was not great (below 4), and there is a good chance that people are familiar with these Within Week calculations, as they are performed from time to time in daily life. But when it came to Across Week questions, moving beyond the week cycle made the visualization more difficult, so the same participants might opt to switch back to the more familiar way of reciting the weekday list to solve the problem. The consistent Direction effect in both participant and item analyses in the Across condition and the exceedingly long RTs with the Backward questions in the Across condition confirm that they switched to list reciting, which took much longer when going backwards. Section 3.1.4.3 provides further clues of this possible split in strategy use with English speakers.

3.1.4.3 Calculation distances analysis

Some discrepancies between participants and items analyses were found in the previous section, as shown in Table 3-4. A few of the effects are significant by participants but not by items, suggesting that the questions were not treated in the same way. This bring us to another novel finding of this research—the strategy use difference revealed by controlling calculation distances. As discussed in Section 3.1, the calculation distance in Kelly et al.’s (1999) experiment was a 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, as one of the strategies reported by the participants was remembering
specific pairs of items (though this was not dominant). The design of the current research, in which distances vary, enabled the exploration of whether the strategy use would be affected by different difficulty levels.

The reaction times for the Within and Across Week questions of the two language groups are presented in Figure 3-7 and Figure 3-8. The distances were controlled to be 1 through 4 in the Within Week condition and 2 through 7 in the Across Week condition. From these two charts, it is clear that the Chinese and English speakers were using different strategies. As shown on the left side of both charts, the Chinese speakers’ reaction times were not sensitive at all to the distance or direction in the Within Week questions, confirming the prediction that they were probably using exclusively numerical strategies—addition and subtraction—by simply transferring calendar questions to number questions. This is because the calendar system in their native language works well with the number series, and it does not take extra effort for them to use this strategy. As a matter of fact, it would be strange if the Chinese speakers chose not to use the arithmetic strategy, as the relationship between the time and number series is too obvious to use any other strategies. However, when it came to the Across Week calculations (Figure 3-8), the Chinese speakers were slower in answering Backward questions. I argue that the longer reaction times with Backward questions only indicate an extra step of applying modulo calculations and is not the result of using a verbal list strategy (a more detailed discussion is in Section 3.1.5). This can be confirmed by the reaction time-by-distance information, which shows that the time spent with Backward questions was not linearly related to the temporal distance to be calculated. However, the fact is that the Chinese speakers did spend more time with increased distance in the Forward Across
Week condition, suggesting that at least some of the Chinese speakers may be counting forward instead of applying the arithmetic strategy in this condition. This is understandable because considering the not-too-long distance (2–7) in the Across Week condition, it might be more convenient to count forward through the weekday list than to do modulo calculation, which may involve non-numerically named Sunday. At the same time, counting backward through the weekday list is an unfamiliar activity that leaves the Chinese speakers no option but to deal with the modulo calculations, which are not that sensitive to the distance (distances of 4–6 take about the same amount of time, as shown on the left side of Figure 3-8).

Figure 3-7: Reaction times of Chinese (left) and English (right) in Within Week calculations (The x-axis represents the calculation distances, and the y-axis is the reaction times in milliseconds.)
On the other hand, the English participants’ reaction times were heavily affected by the length of temporal distances. As can be seen on the right side of Figure 3-7 and Figure 3-8, the questions that required the participants to count backwards took much longer than the Forward questions, because counting backwards is not a daily activity that people are familiar with. Moreover, for shorter distance calculations, their reaction times increase steadily as the distances increase. However, this pattern reversed after the distances reached a certain length. More specifically, the reaction times dropped when the distance was longer than 4 in the Across tasks. The rise and drop of reaction times pertinent to distances strongly suggests that the English speakers were able to apply different approaches when encountering calendar questions with different distances. When the distances grew longer, which makes verbal list counting a less efficient strategy, they were flexible with their strategy use and able to switch 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 the arithmetic strategy was self-reported by some of the English participants in Kelly et al.’s work, but there was no empirical evidence for it until the current research with its varied distance design.

Now we can turn back to the question raised in Section 3.1.4.2—the unexpected interaction between Boundary Type and Direction with English speakers. With reaction times broken down by distance in Figure 3-7 and Figure 3-8, it is clear how the Direction effect is different with the two Boundary types. There was a very small reaction time difference between Forward and Backward questions for distance 1 and 2 in Within Week calculations, which is responsible for the no-show of a Direction effect, whereas Forward questions were answered much faster in the Across Week task, even for distances as short as 1 and 2, which leads to a significant Direction effect. The difference of Direction effect in the two Boundary Types further confirms that English speakers were able to use different strategies in solving temporal calculations with different difficulty levels.

Because the year cycle has longer distances and so can provide more information about strategies used in response to possible calculation difficulties, it is helpful to check the MOY data with the focus on the varied distances. The two language groups’ reaction time data were categorized by distances for the Within task and for the Across task, as shown in Figure 3-9 and Figure 3-10.

The Chinese speakers and the English speakers perform in very different ways in the Within Year calculations. As observed in the DOW tasks, the Chinese speakers’ reactions were not affected by the length of distances in the calculation, and this is true in
solving both the Forward and Backward questions. This result provides clear evidence that the Chinese speakers solve calendar problems with simple addition and subtraction, and there was no difficulty for them in applying the numerical strategy in the Within Year calculations either forward or backward.

![Figure 3-9: Reaction times of Chinese (left) and English (right) in Within Year calculations.](image)

On the other hand, the English speakers present a more complicated picture of responding to different distances. In answering the Forward questions, they spent more time when the distances got longer, and this pattern persists until the distance reaches 7. When the distance was longer than 7, the reaction times begin to drop. This is in line with the prediction that the English speakers were basically using the list reciting strategy, and it is not surprising that counting farther takes more time. And that reaction times decrease for longer distances indicates that the participants switch to other strategies, such as adding the complement of 12 in order to get the correct answer. The switch to numerical strategies helps speed up the responses, but also results in more errors, as they
had a lower accuracy rate with longer distance calculations (see Section 3.1.4.1). The possible change of strategy is even more obvious in the Backward condition. Because to recite the English month names backwards is not natural and is particularly difficult, the English speakers spent much more time solving these problems. And this pattern begins to be present even with distances as short as 2—the adjacent month term would be relatively easy to retrieve, but still take much longer in Backward questions with a distance of 2. And as soon as people needed to count backwards for a distance of 3 or greater, it required a fair amount of extra work, and the time spent skyrocketed. However, the reaction times did not show a pattern of steady increase as observed with the Forward questions, and this indicates that the English speakers might be using various strategies, possibly including counting forward, visualizing, and counting on their fingers, as reported in Kelly et al.’s (1999) work. There is a sharp drop of reaction times from a distance of 7 to a distance of 8, and it is interesting to see that for longer distances (8 to 10) the English speakers actually spent less time than for distances of 3 to 7. This pattern cannot be explained if the English speakers were still using the verbal list strategy. When the temporal distance is longer than 7 months, the English speakers abandon the less efficient reciting method and begin to relate months to numbers. This is extremely natural given that people in the English-speaking world do use numbers to represent months, such as when writing checks and filling out forms.

Now let us take a look at the performances in the Across Year condition, presented in Figure 3-10, in which less regularity is found. Again, the Chinese speakers’ reactions seem not to systematically relate to the distances. In answering the Forward questions, the reaction times do not show a linear relationship with the increasing
distances, and they hit the highest point when the distance is 10, then drop with distances of 11 and 12, showing that the speakers were aware of the strategy of deducting 12 for getting answers in these situations. The use of the 12 complement strategy is even clearer in the Backward performances. The abrupt drop of the reaction times dealing with the distance of 10 (e.g., in answer to the question, “If now is March, 10 months earlier was what month?”) is because the 10 complement of 12 is obviously 2, and people were fast in arriving at 2 and adding it to the month in question in order to get the answer. In a word, the distance data reveal that the Chinese speakers were generally using arithmetic methods in the Across Year calculations. However, I do not, so far, have an explanation for what makes the Backward questions with a distance of 9 so difficult.

![Graph](image)

**Figure 3-10:** Reaction times of Chinese (left) and English (right) in Across Year calculations (The x-axis represents the calculation distances and the y-axis is the reaction times in milliseconds.)

For the English speakers, their reaction times for Forward Across Year calculations increase until the distance is 10, then drop quickly for distances of 11 and 12.
As has been discussed in the previous paragraphs, this trend shows that the English speakers were counting forward through the list for shorter distances, but when it came to longer distances such as 11, they were able to switch to counting backwards the complement number to solve the problem. The Backward responses by distance are much more complicated. They can be divided into two parts—short distances (2 to 5) and longer distances (6 to 11). It is obvious that the participants were solving the shorter distance questions using the most direct way—trying hard to recite the list backwards. This method is direct, but may cost a lot of time due to its unfamiliarity. When the distances get longer, reciting backwards becomes even more difficult, driving the English speakers to seek other methods. It is possible that the specific strategies they were using involved arithmetic, which is not sensitive to the distances (6 through 11), with clear evidence from the distance of 10. Processing distance 10 questions took significantly less time, indicating that at least in this case, the English speakers were using an arithmetic strategy by adding the complement number.

In sum, the Chinese speakers were consistently using arithmetic strategies in calculating temporal distances, and the method they used was not very sensitive to either Boundary Type or Direction. Some of the Chinese speakers do sometimes switch to list reciting methods, but only in limited circumstances such as short distance Across Boundary calculations when performing addition or subtraction involving Sunday.

On the other hand, the English speakers were making use of various strategies including (1) the verbal list-reciting strategy in both directions—either in the direction queried, or, in the reverse direction using a complementary numerical value to that presented in the question, i.e., 7-x for DOW and 12-x for MOY; the verbal-list reciting
method was used overwhelmingly in answering Forward questions and long distance Backward questions; and (2) the numeric strategy including at least two ways of computing Across Boundary responses arithmetically; adding the distance in question is more natural for Forward problems and deducting the complementary numeral is more natural for Backward ones. The numeric methods were mostly used when the calculation distance was long, so that the list-reciting methods became less practical. In a word, when the distances to be calculated exceeded a certain limit that made counting a less efficient method, the English speakers would switch to numerical strategies. This switch of strategy is particularly pronounced in the Backward conditions. This suggests that the difficulty of counting through the month list backwards encouraged the English speakers to search for other possible ways to solve the problem, and the most available method is to relate months to numbers, following the familiar convention, often encountered in daily life, of numbering the months of the year.

Then how does distance affect English and Chinese speakers’ accuracy? Figure 3-11 and Figure 3-12 present the accuracy data by difference in calculation distances in DOW and MOY tasks.
In answering Within Week questions, Chinese speakers have no particular problems with any of the distances or directions, as shown by the upper left figure in Figure 3-11. On the other hand, English speakers’ accuracy steadily decreases when calculation distances got longer. However, the advantage of accuracy with Chinese speakers disappears in the Across Week condition, where the two language groups display similar patterns. Both the Chinese and English speakers maintained high accuracy rates in answering Forward questions, but English speakers were better at answering Backward questions, especially for questions with short distances (distance 2
to 4). This difference implies the possible use of different strategies by the two language groups. Chinese speakers had a hard time getting correct answers in Backward Across Week calculations even with short distances, probably because of the unfamiliar and irregular modulo 7 calculations. However, the English speakers experienced difficulty only with questions involving longer distances, because reciting a list backward for a long distance was particularly difficult.

Figure 3-12 presents an interesting comparison between English and Chinese speakers in doing DOW calculations. For Within Year calculations, Chinese speakers’ performance was not affected by the distance, in both Forward and Backward directions. English speakers, on the other hand, made more mistakes as distances increase, and the pattern is more obvious with Backward questions. Looking at Forward and Backward conditions separately, the high accuracy rate for distance 10 calculations in the Forward condition indicates that English speakers were probably switching to numeric strategies in this situation, as discussed earlier, as the complementary value of 10 can be easily targeted. English speakers experienced difficulty with Backward questions, especially with the ones involving longer distance calculations. The different patterns observed between Chinese and English speakers suggest the different use of strategies in Within Year calculations. More specifically, Chinese speakers seemed to overwhelmingly use the arithmetic strategy, as distance and direction did not affect their performance, whereas English speakers’ better performance in Forward and short distance calculations reveals that they were not using the same strategy as the Chinese speakers.
Interestingly, the situation is totally different in the Across Year calculations, as shown in the lower part of Figure 3-12. Both of the language groups performed better in the Forward than in the Backward condition. And the patterns presented by these two groups almost resemble each other, especially for the Backward condition.

For Forward calculations, English speakers show a general decreasing accuracy rate until distance 10, while Chinese speakers did not show a steady decrease; instead,
their accuracy drops abruptly with distance 10. As mentioned in Section 3.1.4.1’s discussion on the kinds of errors made, Chinese speakers made a lot of base-ten or missed-by-two errors, probably because of confusion between 10 and 12 cyclical systems. Obviously, adding 10 months leads to most of this kind of confusion.

For Backward calculations, the two language groups show similar patterns in that the accuracy rate first drops quickly from distance 2 to 5, then bumps back up, to drop again until distance 10, after which it rises for distances 11 and 12. In Figure 3-10, which shows the reaction times of Chinese and English speakers in Across Year calculations, we can see that Backward distance 10 questions cost the least amount of time for both Chinese and English speakers. But Figure 3-12 shows that those questions also had the lowest accuracy rate. It appears there might a speed-accuracy trade-off in distance 10 calculations for both English and Chinese speakers.

To sum up, the accuracy and reaction time by distance data together suggest that for Within Week and Within Year calculations, Chinese speakers were overwhelmingly using arithmetic strategies, as their speed and accuracy was not affected much by Direction or distance of the questions. English speakers were probably using very different strategies compared to Chinese speakers, as shown by their increasing response times and decreasing accuracy rates when distances became longer. When solving Across Boundary questions, the two language groups sometimes used similar strategies, as suggested by the similar accuracy patterns and similar reaction time patterns. More specifically, when the distance to be calculated was long, English speakers would choose not to use the most direct list-reciting method. Instead, they might switch to more efficient numeric methods, especially in Across Year rather than Across Week conditions.
In addition, the complicated and inconsistent patterns of the accuracy and reaction time data of English speakers also indicate various strategy uses among English speakers.

3.1.5. Discussion

Because of the different levels of transparency in numerical time word systems and arbitrary time word systems, I hypothesized that differences exist in the temporal cognition of adult speakers of languages with distinct systems. The results of this experiment confirm that the Chinese and English speakers indeed perform differently in temporal calculation tasks in that the former complete temporal distance calculations with higher accuracy rates and shorter reaction times. There remains the question, however, of whether the difference comes from the use of different strategies made available by the speakers’ native languages or from other differences, such as in general mathematics ability.

Two predictions, presented in Section 3.1, were made in regard to whether the source of performance difference was the use of specific strategies by the two language groups. First, if the performance difference is due to the Chinese speakers’ use of arithmetic strategies and the English speakers’ use of list-reciting strategies, then the Chinese speakers’ advantage should disappear in cross-week or cross-year calculations, because they need to perform modulo 7 or modulo 12 transfers, which add extra steps to the calculation process and thus should result in longer reaction times and higher error rates. This was confirmed by the results, as it is the case that the 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 that sensitive to Boundary Types; with the list-reciting strategy used most of the time, reciting names within or across a boundary does not make much difference in accuracy. 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 the transferring of numbers to calendars leads to problems as well as an increase in speed. In a word, the use of the arithmetic strategy made the Chinese speakers lose their advantage in accuracy, and it weakened their advantage in reaction time when they were answering Across Week questions.

The second prediction is that if Chinese speakers use arithmetic in temporal calculations more than English speakers, then they would lose their advantage in the calculation of distances involving Sunday, the one week day with a non-numerical name in Chinese (“weekday sun”). The results confirmed that the irregularity of Sunday’s Chinese name causes trouble in applying the arithmetic strategy, resulting in more mistakes by the 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. It is interesting that the Sunday questions did not slow down the Chinese speakers’ calculations, suggesting that the effect of Sunday was subconscious—the Chinese speakers did not bother spending extra time transferring Sunday into “weekday 7” in their calculations, although this was not a wise choice, as suggested by the increased mistakes in solving problems involving Sunday.
The different kinds of errors found with the Chinese and English speakers also suggest that the former group more readily employed the arithmetic strategy. The missed-by-two errors made by a lot of Chinese speakers are obvious evidence that they were transferring calendar problems to more familiar numerical problems, yet made mistakes because of the possible confusion of base-10 and base-12 systems. In contrast, the English speakers’ errors were more various, which suggests a more versatile strategy when encountering different calendar calculation questions. This was confirmed by the discrepancy between participant and item analyses of several tests for the English speakers. Also, for the English speakers, a frequently found pattern is that the errors increase as the distance counted gets longer, which is reasonable, as the verbal list strategy may result in more mistakes when counting greater distances.

Therefore, the differences in calendar terms between the two languages probably lead to dominant arithmetic or list-reciting strategies for the two language groups. The verbal list strategy is substantially slower than the number-transferring one, yet the latter results in more possible errors across boundaries because of the modulo 7 and modulo 12 calculations.

Up to this point, everything is in line with the predictions, yet some details about specific strategy use still need attention. First, were the Chinese speakers using the arithmetic strategy consistently? Taking a close look at the Chinese speakers’ reaction time results, it is worth noting one phenomenon—the Direction effect found in the Across Boundary conditions. If the Chinese speakers were treating calendar terms as a series of numbers and were using the arithmetic strategy in their problem solving, then theoretically, performing simple addition and subtraction should take approximately the
same amount of time. But we observe a Direction effect—longer reaction times for Backward questions than Forward questions in the Across Boundary conditions, as shown in Figure 3-13. The pattern shown in Across Boundary questions resembles the English speakers’ pattern (see Figure 3-14). Does this imply that the Chinese speakers might be reciting a list, at least in the Across Boundary tasks? Possible, yet there is another possibility.

Figure 3-13: Reaction times for Forward and Backward questions of the Chinese speakers

Figure 3-14: Reaction times for Forward and Backward questions of the English speakers
Kelly et al. (1999) collected data on self-reported strategy use. In their report, 90% of the adult Chinese speakers were generally using the arithmetic strategy, but it is not clear whether there was a split of strategy by Boundary Type with the Chinese speakers. Moreover, there were no data on strategies used for individual problems. As a result, although the self-reported strategy use data were very useful in explaining the general differences in strategy use between the two language groups, the mysterious Direction effect for the Chinese speakers was not discussed in detail in Kelly et al.’s paper.

Looking more closely into the modulo calculations, it appears that the forward modulo calculations were easier than the backward ones. For example, 7 months after July (“7 month”) can be transferred to $7 + 7 = 14$, and the answer is simply 14 modulo 12, which is 2 (February “2 month”). However, for backward modulo calculations, such as 7 months before April (“4 month”), one would not first transfer April into “16 month” ($4 + 12$) and then subtract 7 to arrive at the answer “9 month” (September). Instead, two steps of calculations are necessary: the first is to calculate the distance between the starting month, “4 month” (April), and the year boundary, “12 month” (December), which is 4; because the question asks for 7 months before April, 3 more months must be subtracted from “12 month” (December); this is the second step, which gives the correct answer “9 month” (September). This two-step backward calculation method was confirmed by the debriefing with the Chinese speakers after they completed the experiment. As a result, direction does make a difference for Across Boundary calculations—the Backward calculations take an extra step and thus are much longer than the Forward ones. So I argue that the Direction effect found in the Across Boundary calculations with the Chinese speakers could be a combination of employing a verbal list strategy and an
extended two-step arithmetic strategy, to which the latter contributes more (90% according to self-reports collected by Kelly et al. [1999]).

The differences between these two language groups in accuracy rates, response times, and reaction times for different calculation distances support the hypothesis that how a language codes calendar concepts plays a role in what strategy the language’s speakers adopt in solving time-related problems. Because of their language’s employment of number systems, the Chinese speakers automatically use the arithmetic strategy in most of their calculations, which results in shorter reaction times and higher accuracy rates. However, these higher accuracy rates were only exhibited when boundaries were not crossed, and when Sunday was not involved. This suggests that the differences in adults’ temporal reasoning are not differences between the populations in general, but rather differences in the strategies adopted for specific time calculations.

The English speakers lack direct access to a method of relating the temporal series to a numerical series, and therefore they usually recited the name list to solve the calendar questions, especially in circumstances when no boundary is crossed and/or the calculation distance is not long. And the greater difficulty in applying this verbal list strategy in questions of long distances or backwards may have driven the English speakers to other strategies, including the numerical one. The switch of strategy is especially visible in month calculations, probably because the English language has the convention of numbering months in some daily activities.

It is worth noting that language grants an advantage in Chinese speakers’ solving of the temporal problems, yet also keeps them from using other more practical strategies in some situations. For example, the consistent use of the arithmetic strategy in solving
questions involving Sunday led to more mistakes compared to the English speakers’ recitation of the weekday list. By contrast, the English speakers do not have direct access to numerical names for calendar terms, but they are at the same time not so confined to using only one strategy. They are thus more flexible and can switch among various strategies to solve problems in different situations, such as questions of different direction and questions with long calculation distances.
CHAPTER 4
PRIMING STUDY WITH ADULTS

The results presented in Chapter 3 showed that the Chinese and English speakers might be using different strategies in solving calendar problems. However, Chinese speakers might differ from English speakers in one of two ways. First, they could both use the same strategy—simple numerical arithmetic—to reason about time; it could simply be that it is easier for Chinese speakers to map from the DOW and MOY names to numbers because of the transparency of their time words. After all, English speakers are aware of the mapping of months to numbers, as evidenced by the use of numerical abbreviations for months. Alternatively, it could be that English and Chinese speakers use different strategies. English speakers might, most of the time, employ the name recitation strategy, that is, reciting DOW or MOY while tracking the number of terms in the sequence that they have gone through. The results from Experiment 4 suggest that the Chinese speakers might have been unvaryingly using the arithmetic strategy whereas the English speakers were generally using the verbal list-reciting method, but switching to different strategies in different conditions. Experiment 5 tests whether English speakers use list-based strategies, rather than arithmetic ones, in solving most of the time-related problems, and it is intended to further confirm the dominant use of the arithmetic strategy by Chinese speakers by means of a different methodology.

This experiment uses a priming methodology to determine whether first performing an arithmetic or list-based task speeds or slows time calculations for English and Chinese speakers. The priming method is widely used in psychology, cognitive
Generally, the priming effect is “a nonconscious influence of past experience on current performance or behavior” (Schacter & Buckner, 1998, p. 185). In other words, it is an implicit memory effect in which exposure to a stimulus influences a response to a later stimulus. It is ubiquitous in daily life (Tulving & Schacter, 1990) and can be very salient and long-lasting (Tulving, Schacter, & Stark, 1982). The commonly found priming effect can be positive or negative in its influence on the speed of response. Positive priming is thought to be caused by spreading activation (Reisberg, 2007) in that the first stimulus activates parts of a particular representation or association in memory just before an action or task is carried out. The representation is already partially activated when the second stimulus is encountered, so less additional processing time for activation is needed. Experiment 5 tries to find a positive priming effect.

There are several types of priming, with perceptual and conceptual priming being the main two types (Tulving & Schacter, 1990; Schott, Richardson-Klavehn, Heinze, & Düzel, 2002). Priming studies of perception (McDermott, 1997; Wiggs & Martin, 1998; Schott et al., 2002; Kouider & Dehaene, 2007; Finkbeiner & Forster, 2008) typically show that subjects respond faster and more accurately when a prime and target are congruent (i.e., prime and target involve the same operation) than when they are incongruent (i.e., prime and target involve different operations). Conceptual priming studies demonstrate that when examples that share conceptually related features are shown to people, they subsequently use those shared features or even generate additional features that are consistent with the emergent underlying concept from the primed feature constellation (Marsh, Bink, & Hicks, 1999; Levy, Stark, & Squire, 2004). Conceptual
priming is enhanced in semantic tasks in which word recognition is typically faster when the target word (e.g., doctor) is preceded by a semantically related prime word (e.g., nurse) (Friederici, Steinhauer, & Frisch, 1999). Note that although most of these priming experiments have been conducted using linguistic material, the priming effect is not confined to word recognition and production (Voss, Lucas, & Paller, 2009) or even the sentential context level (Stanovich & West, 1983), but can also “occur with nonverbal stimuli such as pictures, shapes and facts” (Tulving & Schacter, 1990, p. 303). Priming works best when the two stimuli are in the same modality (visual cues for visual tasks and verbal prompts before verbal tasks), but research shows that it also works cross-modally (Roediger & Blaxton, 1987; Holcomb & Anderson, 1993; Zurif, 1995; Diependaele, Sandra, & Grainger, 2005).

4.1 Experiment 5

Based on the studies mentioned in the previous section, I designed Experiment 5 to use the implicit priming methodology. Only the month questions are tested in this experiment because a month task can be designed to include more various distances, which can provide more information about strategy use. Also, Section 3.1.4 shows that it is in the month condition that the English speakers were probably using different kinds of strategies, especially with long distance calculations.

The hypothesis that Experiment 5 tests is that the two language groups’ speed in solving calendar problems will be affected when they encounter a question of some type immediately before the task. If Chinese and English speakers in general use different online strategies in calendar calculation, then their calculations should be facilitated by
different types of priming tasks. More specifically, if Chinese speakers do use an arithmetic strategy in calendar calculation, then their responses should be faster after solving an arithmetic-based calculation prime task, but not after a list-reciting task; meanwhile, for English speakers, the opposite should be true. However, if both English and Chinese speakers use arithmetic calculations to perform time calculations, but have different abilities to access the relevant numbers, then English and Chinese speakers should both perform better on a time calculation when it follows an arithmetic prime task.

4.1.1 Participants

Thirty-nine native speakers of Chinese and 36 native speakers of English, none of whom participated in Experiment 4, were recruited. The Chinese speakers were undergraduate and graduate students studying at South China Agriculture University, who ranged in age from 19 to 35 \( (M = 23.8, SD = 4.5) \). All of them (23 females and 16 males) were native speakers of Mandarin, although they had studied English for 4 to 18 years \( (M = 10.13, SD = 2.48) \). The English speakers (21 females and 15 males) were undergraduate students at the University of Hawai‘i, aged from 18 to 32 \( (M = 21, SD = 3.0) \). Some of these English speakers had studied languages that include French, Ilocano, Spanish, Portuguese, Russian, Hawaiian, and Chamorro. One participant had studied more than two years of Japanese, which makes use of numbers in its month system, and the data from that participant were excluded from the analysis. The data from another two speakers were excluded from the analysis because their voices were too soft to trigger the microphone for about one third of the trials. All of the participants either
received extra credit in an introductory linguistics class or were paid five dollars or the equivalent amount for their time.

**4.1.2 Materials and Design**

A set of the questions generated for the MOY block in Experiment 4 is used in this test. In addition to the variables described in the previous experiment, there is another variable in this experiment: Prime Type with two levels—Arithmetic and List. The Arithmetic Primes are simple additions and subtractions for numbers under 10. The numbers that appear in the Arithmetic Prime do not appear in the following temporal calculation question in order to avoid a simple word priming effect. The List Primes are alphabetic counting questions for the English-speaking participants and, in order to avoid possible confounding from their study of the English language, Celestial Stem counting questions for the Chinese-speaking participants.

The Celestial Stems (天干 tiāngān) are 10 characters that form a Chinese system of ordinals (see Table 4-1). This system was widely used in the past to name days in a 10-day period, and its most prominent use was in combination with the 12 Earthly Branches (地支 dìzhī) in a 60-year calendrical cycle. The stems are still widely used nowadays in Chinese counting systems, similar to the way the alphabet can be used in English. For example, they are used to name organic chemicals in terms of the carbon chain length, such as methanol (甲醇 jiāchún) and ethanol (乙醇 yǐchún). They are also used in the names of certain diseases such as Hepatitis A (甲型肝炎 jiǎxíng gānyán) and Hepatitis B (乙型肝炎 yǐxíng gānyán). Another common use is in sports league names; for example, Lega Serie A is known as 意甲 yìjiǎ in Chinese. Because of these familiar
uses, the Celestial Stem system is ideal for List Primes for the Chinese speakers in the current experiment.

Table 4-1: The ten Celestial Stems as a Chinese ordinal system

<table>
<thead>
<tr>
<th>Celestial Stem</th>
<th>甲</th>
<th>乙</th>
<th>丙</th>
<th>丁</th>
<th>戊</th>
<th>己</th>
<th>庚</th>
<th>辛</th>
<th>壬</th>
<th>癸</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>jiǎ</td>
<td>yǐ</td>
<td>bǐng</td>
<td>dīng</td>
<td>wù</td>
<td>jǐ</td>
<td>gēng</td>
<td>xīn</td>
<td>rén</td>
<td>guǐ</td>
</tr>
<tr>
<td>Corresponding letters</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
</tr>
</tbody>
</table>

The two Prime Types with two Directions generate four types of Primes (1a, 1b, 2a, 2b). The critical questions are formatted differently from those used in Experiment 4. It is highly likely that the effect I am looking for is relatively small; therefore, it is essential to measure the reaction times in a more accurate way. For this reason, the critical questions are designed to require the participants to complete a sentence (see Examples 3 and 4) rather than to give an answer in response to a complete question sentence. In this way, there is less time included in the measurement. The Prime questions are formatted in the same way, as shown in (1) and (2).

(1a) 3 + 5 is? (Arithmetic/Forward)
(1b) 7 - 2 is? (Arithmetic/Backward)
(2a) Three letters after A is? (List/Forward)
(2b) Five letters before H is? (List/Backward)
(3a) If this month is July, two months from now it’ll be?
(3b) If this month is December, nine months ago it was?
(4a) If this month is October, three months from now it’ll be?
(4b) If this month is January, five months ago it was?

Prime Type and Prime Direction are counterbalanced across questions. The critical questions were presented in four blocks of the four prime conditions: List Forward, List Backward, Arithmetic Forward, and Arithmetic Backward. In each of the blocks, the critical questions were balanced in Boundary Type and Direction. That is to say, the direction of the prime task (Forward/Backward) is either the same or the opposite as the subsequent time target (Forward/Backward). The order in which the four blocks were presented was counterbalanced among the participants. Reaction times were recorded and error rates were calculated as dependent measures.

4.1.3 Procedure

The procedure is similar to that in Experiment 4, differing only in that before each critical trial a prime trial is added, so that the participants are required to do either a List or Arithmetic calculation before doing the time calculation.

Participants initiated each prime trial by pressing the SPACE bar. Once they pressed the SPACE bar, a fixation cross appeared on the screen for 500 milliseconds, followed by the playing of a recorded prime question through the speakers. As soon as they spoke, the participants were allocated 450 milliseconds to articulate an answer to the prime question, followed by a pause of 250 milliseconds before the playing of the critical question. After hearing the critical question, they had 600 milliseconds to articulate an answer. Then they could press the SPACE bar whenever they were ready to listen to the next prime question.
Four practice items preceded the experiment. In addition, the Chinese participants were asked to recite the Celestial Stems before the experiment to make sure they were familiar with the 10 terms. There was a break halfway through the experiment, after the first two blocks had been presented. The experiment took approximately 25 to 30 minutes for the Chinese participants and 35 to 40 minutes for the English participants to complete.

4.1.4 Predictions

I expected to replicate the main effect of Language that I found in Experiment 4, whereby Chinese speakers are generally faster and more accurate in responding to questions. But what is more important is that if English and Chinese speakers are really using different strategies (listing-dominant versus arithmetic-dominant), there should be a two-way interaction between Language and Prime Type; Chinese speakers should be faster after arithmetic problems and slower after listing problems. By contrast, English speakers should be faster after solving a list-recitation prime problem than after an arithmetic prime. However, if English and Chinese speakers are actually using the same arithmetic strategies, but with different degrees of success and efficiency, then there should be no such interaction effect, and instead only a main effect of Prime Type, where both language groups are more successful following arithmetic primes.

The results should make it possible not only to measure once again whether the two populations perform differently in time calculation tasks, but also to tease apart exactly what strategies they use that cause them to perform differently.
4.1.5 Results

4.1.5.1 Error analysis

No participants or items were excluded because of a low or outlying accuracy rate. Although no previous literature suggests any differences in accuracy in the current kind of task, it might be the case that different Prime Types would affect participants’ accuracy due to consistency or inconsistency between the strategies they would need to use. The accuracy data might be informative in this regard.

As described in Section 4.1.2, there are four blocks designed to test both Prime Type and Prime Direction. It is hypothesized that Prime Type would make a difference to the subsequent response to a critical question; however, it is not clear whether different Directions of the same type of Prime would cause any difference. So I first ran two four-way repeated-measures ANOVAs by participant and by item to test if there were any Compatibility or Incompatibility effects of Prime and Critical Direction (the direction of the calendrical question). The individual proportions of correct answers were arcsine transformed to comply with the normal distribution premise. The by-participant analysis has Prime Type, Prime Direction, and Critical Direction as within-subject factors and Language as the between-subject factor. For the by-item analysis, Prime Type, Prime Direction, and Language are taken as the within-item factors while Critical Direction is the between-item factor.

There is a main effect of Language, $F_1(1, 70) = 28.292, p < 0.001, F_2(1, 120) = 30.007, p < 0.001$; once again the Chinese speakers were more accurate in solving the time questions. Also, as in the previous experiment, there is a main effect of Critical Direction, $F_1(1, 70) = 59.712, p < 0.001, F_2(1, 120) = 29.934, p < 0.001$; both of the two language groups made fewer errors in the Forward conditions. There is an interaction
effect between Critical Direction and Language that is significant by participant, $F_1(1, 70) = 6.257, p = 0.015$, and marginally significant by item, $F_2(1, 120) = 2.874, p = 0.093$, indicating that Chinese and English speakers were performing differently in regard to the directions of the question.

Looking at the two language groups separately, there is a main effect of Critical Direction for both the Chinese, $F_1(1, 38) = 16.315, p < 0.001, F_2(1, 60) = 4.953, p = 0.030$, and the English speakers, $F_1(1, 32) = 43.771, p < 0.001, F_2(1, 60) = 34.195, p < 0.001$; the Forward questions were more often answered correctly. However, there is no interaction effect between Prime Type and Prime Direction with either Chinese or English speakers. Moreover, the pair-wise comparisons showed no effect of Prime Directions within the same type of Prime. In other words, the same type of Prime works in a similar way with either Prime Direction. So I collapsed the Prime Directions, and ran some more tests that instead included the Boundary Type of the critical questions for both language groups. Because the main purpose of the current experiment is to test whether the two prime types cause different effects on the two language groups, statistical tests were performed for Chinese and English speakers separately instead of in an aggregated way.

**Accuracy analysis for the Chinese group.** For the Chinese speakers, two three-way repeated-measures ANOVAs were performed, one each with participants and items as random factors. In the by-participant analysis, the data were analyzed by Prime Type (Arithmetic/List), Critical Direction (Forward/Backward), and Boundary Type (Within/Across) with Critical Direction (referred to as “Direction” hereafter in this
section) and Boundary Type as within-subjects variables. For the item analysis, Prime is taken as a within-item variable while Boundary Type and Direction are between-items variables. The results are presented graphically in Figure 4-1 and reported below.

No main effect of Prime Type was found. As can be seen from Figure 4-1, neither of the two types of primes induced higher accuracy rates on the answers to the critical questions. There are main effects of Boundary Type, $F_1(1, 38) = 104.072, p < 0.001, F_2(1, 60) = 61.758, p < 0.001$, and Direction $F_1(1, 38) = 11.257, p = 0.002, F_2(1, 60) = 11.370, p = 0.001$. Within conditions and Forward questions have more correct answers, which replicates the results in the previous experiment. There is also an interaction between Boundary Type and Direction, $F_1(1, 38) = 29.502, p < 0.001, F_2(1, 60) = 16.112, p < 0.001$; the participants were more successful in Forward than in Backward calculations in the Across Year condition only, $F_1(1, 38) = 26.213, p < 0.001, F_2(1, 30) = 27.741, p < 0.001$, whereas no such difference was found in the Within condition.
Accuracy analysis for the English group. Figure 4-2 presents the English speakers’ performance in different Prime Type and Boundary Type conditions. The same sets of tests were performed.

![Accuracy rates for the English speakers in different Prime conditions](image)

Figure 4-2: Accuracy rates for the English speakers in different Prime conditions

No main effect of Prime Type was found as error rates were not different in the List or Arithmetic Prime conditions. There is a main effect of Boundary Type, $F_1(1, 32) = 36.760, p < 0.001$, $F_2(1, 60) = 23.332, p < 0.001$, confirming better performance in the Within condition. Forward questions were answered more accurately, as shown by a main effect of Direction, $F_1(1, 32) = 38.286, p < 0.001$, $F_2(1, 60) = 14.064, p < 0.001$. There is also an interaction effect between Boundary Type and Direction, which is only significant by participants, $F_1(1, 32) = 17.539, p < 0.001$; the Direction effect is bigger in the Across Boundary than in the Within Boundary calculations. However, not all of the items work in a consistent way. There is also a three-way interaction among Prime Type,
Boundary, and Direction $F_1(1, 32) = 6.566, p = 0.015, F_2(1, 60) = 6.325, p = 0.015$, suggesting a need to look into the data in more detail.

In the Within Year task, there is an interaction between Prime Type and Direction $F_1(1, 32) = 4.862, p = 0.035, F_2(1, 30) = 13.236, p = 0.001$; the Arithmetic Primes induce higher accuracy rates than the List Primes, $t(1,32) = 2.066, p = 0.047$, in the Within Forward condition, whereas no effect of Prime Type is found in the Within Backward condition. In the Across Year condition, no interaction between Prime Type and Direction was found.

Comparing the two figures of accuracy rates for the Chinese (Figure 4-1) and English (Figure 4-2) speakers, it is interesting to see that the Chinese-speaking participants were sensitive to Prime Types in the Backward calculations, while on the other hand, the English-speaking group’s accuracy is affected by the Prime Types in the Forward conditions.

4.1.5.2 Reaction time analysis

All trials with incorrect responses, which made up 15.6% of the data, were excluded from the data analysis. I also removed all responses that were greater than 2.5 standard deviations from the mean of all responses in each of the eight conditions (Prime Type x Boundary Type x Direction) for each language group. That excluded another 3.2% of the data. No participants or items were removed for reasons of accuracy or outlying $SD$.

As in the accuracy analysis, I first tested if the Prime Directions of the same type of Prime cause any differences. Two four-way repeated-measures ANOVAs by
participant and by item were performed. The by-participant analysis has Prime Type, Prime Direction, and Critical Direction as within-subject factors and Language as the between-subject factor. And for the item analysis, Prime Type, Prime Direction, and Language are taken as the within-item factors, while Critical Direction is the between-item factor.

No main effect of Prime Type is found. There is a large main effect of Language, $F_1(1, 70) = 87.032, p < 0.001$, $F_2(1, 120) = 149.497, p < 0.001$, which replicates the results from Experiment 4, in which the Chinese speakers were generally faster than their English-speaking counterparts. Also, there is a main effect of Critical Direction, $F_1(1, 70) = 47.438, p < 0.001$, $F_2(1, 120) = 16.591, p < 0.001$; Forward calculations are faster than Backward ones. There is also an interaction effect between Prime Type and Language, $F_1(1, 70) = 10.598, p = 0.002$, $F_2(1, 120) = 7.362, p = 0.008$, suggesting that the Chinese and English speakers’ reaction times were affected in different ways by their encounters with the different types of Prime.

It seems that items were inconsistent because three interaction effects were found to be significant only by participants, and not by item—interaction between Language and Prime Direction, $F_1(1, 70) = 4.963, p = 0.029$, and interaction between Language and Critical Direction, $F_1(1, 70) = 5.385, p = 0.023$. These two interactions show that (1) Prime directions had different influences on the responses to critical questions for the two language groups; and (2) Chinese and English speakers had different speeds in answering Forward and Backward calendar questions. Another interaction effect was found that was only significant in the participant analysis, between Prime Type and Prime Direction,
\[ F_1(1, 70) = 4.789, \ p = 0.032, \] suggesting that when the directions of Arithmetic and List Primes were different, they affected critical questions in different ways.

There are two three-way interactions significant by participants and marginally significant by items: one among Language, Prime Type, and Prime Direction, \( F_1(1, 70) = 5.604, \ p = 0.021, \) \( F_2(1, 120) = 2.702, \ p = 0.1, \) and another among Language, Prime Type, and Critical Direction, \( F_1(1, 70) = 9.237, \ p = 0.003, \) \( F_2(1, 120) = 3.876, \ p = 0.051. \)

In order to look into the two three-way interactions, I performed 2 (Prime Type) x 2 (Prime Direction) x 2 (Critical Direction) repeated-measures ANOVAs separately for the Chinese- and English-speaking groups.

I first take a look at the Language, Prime Type, and Prime Direction interaction. For Chinese speakers, no interaction between Prime Type and Prime Direction was found, suggesting that Prime Direction of the same Prime Type is not an influential factor for Chinese speakers. For the English speakers, however, it is not the same case. An interaction between Prime Type and Prime Direction was found, which was significant in the participant analysis, \( F_1(1, 32) = 5.876, \ p = 0.021, \) and marginally significant in the item analysis, \( F_2(1, 60) = 3.077, \ p = 0.085. \) This is the cause that drives the three-way effect in discussion. In order to uncover where the interaction effect was coming from within the two Prime Types, one-way repeated-measures ANOVAs were performed with English speakers. In the Arithmetic condition only, there is no main effect of Prime Direction. On the other hand, in the List Prime condition, there is a main effect of Prime Direction, which is significant only in the participant analysis, \( F_1(1, 32) = 8.983, \ p = 0.005, \) in which Forward Primes induce faster responses than Backward Primes, but the
non-significance in the item analysis suggests that not all critical questions were affected by this effect.

Turning to the three-way interaction among Language, Prime Type, and Critical Direction, the effect was not mainly driven by English speakers as no interaction between Prime Type and Critical Direction was found. However, Chinese speakers showed this interaction, $F_1(1, 38) = 15.710, p < 0.001$, $F_2(1, 60) = 4.559, p = 0.037$. More discussion of this interaction can be found in the next subsection, in which Chinese speakers’ performance is discussed in detail.

The results so far show that the Chinese speakers were probably acting in a consistent way, and their performances were not affected by the directions of the primes—their responses were not different in the Addition and Subtraction conditions, nor in the List Forward and List Backward conditions. However, the English speakers present a more complicated picture in which their performances were varied when they needed to count forward or backward through the alphabetical list before answering the questions. These differences imply possible strategy use differences. As a result, some more analyses, done separately for the two language groups, are needed.

**Reaction time analysis for the Chinese group.** Two three-way repeated-measures ANOVAs were performed, one each with participants and items as random factors. In the by participant analysis, the data were analyzed by Prime Type (Arithmetic/List), Critical Direction (Forward/Backward), and Boundary Type (Within/Across) with Critical Direction (referred to as “Direction” hereafter in this section) and Boundary Type as within-subjects variables. For item analysis, Prime is taken as a within-item variable
while Boundary Type and Direction are between-items variables. The results are reported in Table 4-2 and presented graphically in Figure 4-3.

Table 4-2: Mean RT and SD in each condition for the Chinese speakers

<table>
<thead>
<tr>
<th>Boundary Type</th>
<th>Direction</th>
<th>Prime Type</th>
<th>Mean RT (ms)</th>
<th>SD (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within</td>
<td>Forward</td>
<td>Arithmetic</td>
<td>700</td>
<td>604</td>
</tr>
<tr>
<td></td>
<td></td>
<td>List</td>
<td>748</td>
<td>596</td>
</tr>
<tr>
<td></td>
<td>Backward</td>
<td>Arithmetic</td>
<td>860</td>
<td>657</td>
</tr>
<tr>
<td></td>
<td></td>
<td>List</td>
<td>927</td>
<td>663</td>
</tr>
<tr>
<td>Across</td>
<td>Forward</td>
<td>Arithmetic</td>
<td>1498</td>
<td>655</td>
</tr>
<tr>
<td></td>
<td></td>
<td>List</td>
<td>1589</td>
<td>739</td>
</tr>
<tr>
<td></td>
<td>Backward</td>
<td>Arithmetic</td>
<td>1943</td>
<td>852</td>
</tr>
<tr>
<td></td>
<td></td>
<td>List</td>
<td>2573</td>
<td>1081</td>
</tr>
</tbody>
</table>

As can be seen from Figure 4-3, the participants answered the questions faster after they encountered the Arithmetic Primes than after they encountered the List Primes. This is confirmed by a main effect of Prime Type, $F_1(1, 38) = 21.224, p < 0.001$, $F_2(1, 60) = 12.143, p = 0.001$. This is compatible with the view that the Chinese speakers were
using an arithmetic strategy in temporal calculations, as their responses became faster after they performed addition or subtraction. Again, the Within questions were answered faster than the Across ones, as shown by a large main effect of Boundary Type, $F_1(1, 38) = 338.787, p < 0.001, F_2(1, 60) = 111.248, p < 0.001$. There is also a main effect of Direction, $F_1(1, 38) = 96.756, p < 0.001, F_2(1, 60) = 16.112, p < 0.001$; Forward tasks are always faster than Backward tasks.

There are significant interactions (1) between Prime Type and Boundary Type, $F_1(1, 38) = 16.376, p < 0.001, F_2(1, 60) = 5.742, p = 0.020$; (2) between Prime Type and Direction, $F_1(1, 38) = 15.710, p < 0.001, F_2(1, 60) = 4.559, p = 0.037$; and (3) between Boundary Type and Direction, $F_1(1, 38) = 48.966, p < 0.001, F_2(1, 60) = 7.050, p = 0.010$.

These three interactions indicate that (1) the difference between Arithmetic and List Primes is much larger in the Across Boundary condition than in the Within Boundary condition; (2) the difference between Forward and Backward Directions was greater after participants encountered the Arithmetic Primes; and (3) the difference between Forward and Backward Directions is much larger in the Across Boundary condition.

In addition, there is a significant three-way interaction between Prime Type, Boundary Type, and Direction, $F_1(1, 38) = 16.286, p < 0.001, F_2(1, 60) = 4.711, p = 0.034$. Separate analyses were performed for the two Boundary Types in order to get a clearer picture of this interaction.

For the Within Year task, I found no main effect of Prime Type, although the mean response times were shorter in the Arithmetic Prime conditions. This might be a ceiling effect, as the Chinese speakers were already very fast with the Within Boundary questions, so perhaps the priming effect could not be fully realized. What was found is a
main effect of Direction, $F_1(1, 38) = 35.194, p < 0.001, F_2(1, 30) = 6.869, p = 0.014$.

Calculation times decreased with Forward questions.

For the Across Year tasks, on the other hand, there is a main effect of Prime Type, $F_1(1, 38) = 23.512, p < 0.001, F_2(1, 30) = 10.229, p = 0.003$. Questions were answered faster after participants encountered Arithmetic Primes than after they encountered List Primes, suggesting that the Chinese speakers were using arithmetic strategies in solving the temporal calculation questions. Again, there is a Direction effect, and it is of much larger size than that in the Within condition, $F_1(1, 38) = 81.601, p < 0.001, F_2(1, 30) = 11.920, p = 0.002$, favoring the Forward Direction. There is also a Prime Type by Direction effect, $F_1(1, 38) = 18.696, p < 0.001, F_2(1, 30) = 5.483, p = 0.026$; the difference between Forward and Backward Directions is larger in the Arithmetic Prime condition. Pair-wise comparisons show that in the Backward calculations, the Arithmetic Primes cause faster responses than the List Primes, $t(1,38) = 5.350, p < 0.001$. No difference of the Prime pairs in other Boundary and Direction conditions reaches significance.

It is interesting that the priming effect seems not to be affected by the compatibility of the Prime Direction and Critical Direction. As Table 4-3 shows, although the reaction times between the Compatible and Incompatible conditions are slightly different, it is not statistically significant.
Table 4-3: Chinese speakers’ reaction times by compatibility of Prime Direction and Critical Direction

<table>
<thead>
<tr>
<th>Prime Type</th>
<th>Compatibility</th>
<th>Mean RT (ms)</th>
<th>SD(ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arithmetic</td>
<td>Compatible</td>
<td>1179</td>
<td>626</td>
</tr>
<tr>
<td></td>
<td>Incompatible</td>
<td>1224</td>
<td>691</td>
</tr>
<tr>
<td>List</td>
<td>Compatible</td>
<td>1409</td>
<td>711</td>
</tr>
<tr>
<td></td>
<td>Incompatible</td>
<td>1393</td>
<td>712</td>
</tr>
</tbody>
</table>

Figure 4-4: Chinese speakers’ reaction times when Prime Direction and Critical Direction are Compatible (left) or Incompatible (right)

The pattern for the Incompatible condition data (Figure 4-4, right side) looks similar to the pattern for the Compatible condition data (Figure 4-4, left side). Statistical tests showed that whatever effects were found in the Compatible condition were found in the Incompatible condition, as well. In other words, even when the Prime Direction and the Critical Direction are incompatible (e.g., an Addition Prime followed by a Backward temporal question; Subtraction followed by a Forward question; a Forward list followed by a Backward question; or a Backward list followed by a Forward question), there is still a strong Prime Type effect, $F_1(1, 38) = 9.922, p = 0.003$, $F_2(1, 28) = 7.274, p = 0.012$, as
Arithmetic Primes always induce faster responses regardless of the Prime Direction compatibility.

This finding suggests that the compatibility of Primes and Criticals is not a necessity for the priming effect, at least for the Chinese speakers in the temporal calculation tasks. As long as the participants had solved an arithmetic question, the participants were set to the “arithmetic mode,” and their subsequent reaction to a temporal question was facilitated, even when the initial arithmetic problem required addition while the following temporal question required subtraction, or vice versa.

**Reaction time analysis for the English group.** The same sets of tests were applied to the data of the English-speaking group. The results are reported in Table 4-4, and are perhaps best understood visually, as they are presented in Figure 4-5.

<table>
<thead>
<tr>
<th>Boundary Type</th>
<th>Direction</th>
<th>Prime Type</th>
<th>Mean RT (ms)</th>
<th>SD (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within</td>
<td>Forward</td>
<td>Arithmetic</td>
<td>2488</td>
<td>868</td>
</tr>
<tr>
<td></td>
<td></td>
<td>List</td>
<td>2355</td>
<td>713</td>
</tr>
<tr>
<td></td>
<td>Backward</td>
<td>Arithmetic</td>
<td>3517</td>
<td>1815</td>
</tr>
<tr>
<td></td>
<td></td>
<td>List</td>
<td>3052</td>
<td>1988</td>
</tr>
<tr>
<td>Across</td>
<td>Forward</td>
<td>Arithmetic</td>
<td>2934</td>
<td>851</td>
</tr>
<tr>
<td></td>
<td></td>
<td>List</td>
<td>3033</td>
<td>891</td>
</tr>
<tr>
<td></td>
<td>Backward</td>
<td>Arithmetic</td>
<td>3700</td>
<td>1476</td>
</tr>
<tr>
<td></td>
<td></td>
<td>List</td>
<td>3605</td>
<td>1253</td>
</tr>
</tbody>
</table>
In a difference from the Chinese language group, with the English language group there is no main effect of Prime Type. As Figure 4-5 suggests, the effect of Prime is complicated and varied by condition. Responses are faster in the Within Year calculations, $F_1(1, 32) = 18.259, p < 0.001, F_2(1, 60) = 10.336, p = 0.002$. This Boundary effect suggests that some arithmetic strategy use is involved because it is arithmetic rather than list-reciting that is sensitive to Boundary Types. Overall, Forward is faster than Backward, as confirmed by a main effect of Direction, $F_1(1, 32) = 21.038, p < 0.001, F_2(1, 60) = 13.195, p = 0.001$. There is also a marginal interaction effect between Prime Type and Boundary, $F_1(1, 32) = 3.186, p = 0.084, F_2(1, 60) = 3.119, p = 0.082$, which suggests that a further look into the Prime effect with the two Boundary types might be necessary.

The Within condition shows a main effect of Prime Type, $F_1(1, 32) = 5.005, p = 0.032, F_2(1, 30) = 4.006, p = 0.054$, in which the List Primes induce faster responses than
the Arithmetic ones, as predicted. Again, there is a Direction effect, \( F_1(1, 32) = 20.370, p < 0.001, F_2(1, 30) = 20.146, p < 0.001; \) Forward is faster than Backward.

As for the Across condition only, there was no main effect of Prime Type, but there was a main effect of Direction, as in the Within condition, although of smaller magnitude, \( F_1(1, 32) = 12.985, p = 0.001, F_2(1, 30) = 4.692, p = 0.038. \)

At the first look, the English speakers’ reaction times were not affected by the Compatibility of Prime Direction and Critical Direction, as shown in Table 4-5. There is no significant Compatibility or Incompatibility effect.

Table 4-5: English speakers’ reaction times regarding Compatibility of Prime Direction and Critical Direction

<table>
<thead>
<tr>
<th>Prime Type</th>
<th>Compatibility</th>
<th>Mean RT (ms)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arithmetic</td>
<td>Compatible</td>
<td>3023</td>
<td>984</td>
</tr>
<tr>
<td></td>
<td>Incompatible</td>
<td>3039</td>
<td>958</td>
</tr>
<tr>
<td>List</td>
<td>Compatible</td>
<td>2866</td>
<td>767</td>
</tr>
<tr>
<td></td>
<td>Incompatible</td>
<td>3018</td>
<td>940</td>
</tr>
</tbody>
</table>

Figure 4-6: English speakers’ reaction times when Prime Direction and Critical Direction are Compatible (left) or Incompatible (right)
However, unlike the Chinese speakers whose reaction times showed similar patterns for the two Prime Types in the Compatible and Incompatible conditions, the English speakers’ performances were very different depending on compatibility, especially in the Across Boundary task, as shown in Figure 4-6.

When the Prime Direction and the Critical Direction are compatible (shown in Figure 4-6, left side), there is a Prime effect found in the Within Boundary condition, $F_1(1, 32) = 5.248, p = 0.029, F_2(1, 14) = 3.194, p = 0.069$; List Primes induce faster responses. This shows that, in the Within Year calculations the English speakers were more readily using a list-reciting strategy, which was facilitated by solving a list problem in the priming question; however, this effect disappears when the Prime Direction and the Critical Direction are incompatible (shown in Figure 4-6, right side). The difference between the two compatibility conditions is even larger in the Across Boundary condition. For the Compatible condition, the Arithmetic Prime induced faster responses in the Backward condition; however, in the Forward condition, the List Prime induced faster responses. For the Incompatible condition, the pattern is the opposite.

These complex interactions are perhaps best understood visually, as in Figure 4-7, where the reaction times for the two compatibility conditions are presented separately for the List and Arithmetic Primes (List Prime on the left and Arithmetic Prime on the right).

List Primes show a Compatibility effect in the Across Forward condition. On the other hand, Arithmetic Primes show a Compatibility effect in the Across Backward condition.
To summarize the important findings of Experiment 5 thus far, the List and Arithmetic Primes affect Chinese and English speakers’ temporal calculations in different ways. Arithmetic Primes induce faster responses with the Chinese speakers, but only in the Across Year task; List Primes induce faster responses with the English speakers, but only in the Within Year tasks.

**Reaction time analysis by calculation distance.** The results from Experiment 4 suggest that distance might be an influential factor in determining what strategy gets used in the temporal calculation tasks, especially for the English speakers. In the current experiment, the calculation distances were set between 2 and 9 in order to explore whether the English speakers switch between different strategies according to the level of difficulty. The non-significance of some by item analyses shown in the previous subsections also suggests that items were not treated consistently, probably due to the different distances.
Thus statistical tests were applied to the data from the Chinese and English language groups separately, and the results are reported below.

Distance, categorized into two levels as Short (2–5) and Long (6–9), is taken as a within-participant factor and submitted to a two-way repeated-measures ANOVA together with Prime Type; Language is a between-participant factor. There is a large main effect of Distance, $F_1(1, 70) = 173.441, p < 0.001$, $F_2(1, 124) = 20.634, p < 0.001$; the response times increase with distance. There is also an interaction between Language and Distance that is significant by participant, $F_1(1, 70) = 17.040$, and marginally significant by item, $p < 0.001$, $F_2(1, 124) = 2.621, p = 0.1$; the speed difference between Short and Long Distance calculations was different for Chinese speakers and English speakers, suggesting that it is necessary to check how the two language groups react to the two primes in the Short and Long Distance tasks.

![Figure 4-8: Chinese speakers’ reaction times in different Prime conditions for Short (left) and Long (right) Distance calculations](image-url)
For the Short and Long Distance tasks, the Chinese speakers’ performances were of similar patterns, as Figure 4-8 shows—the Short Distance condition is on the left side and the Long Distance condition is on the right side. Responses are almost always faster in the Arithmetic than in the List Prime condition, regardless of the calculation distance.

There are Prime Type effects in both Short Distance \((F_1(1, 38) = 18.631, p < 0.001, F_2(1, 28) = 10.463, p = 0.007)\) and Long Distance \((F_1(1, 35) = 4.109, p = 0.050, F_2(1, 28) = 3.463, p = 0.057)\) tasks—questions were answered faster after the Chinese participants encountered Arithmetic Primes than after they encountered List Primes. Boundary and Direction also had significant effects for both the Short and Long Distance tasks. Within Boundary calculations are always faster, whether in the Short Distance \((F_1(1, 38) = 124.21, p < 0.001, F_2(1, 28) = 18.431, p = 0.001)\) or Long Distance \((F_1(1, 35) = 198.265, p < 0.001, F_2(1, 28) = 68.308, p < 0.001)\) conditions. Forward is always faster than Backward in both Short Distance \((F_1(1, 38) = 20.715, p < 0.001, F_2(1, 28) = 7.618, p = 0.041)\) and Long Distance \((F_1(1, 35) = 42.172, p < 0.001, F_2(1, 28) = 17.342, p < 0.001)\) calculations. Interaction between Boundary Type and Direction is significant in the Short Distance task, \(F_1(1, 38) = 12.638, p = 0.001, F_2(1, 28) = 6.881, p = 0.022\), and in the Long Distance task, \(F_1(1, 35) = 35.780, p < 0.001, F_2(1, 28) = 4.984, p = 0.045\).

In short, the two types of priming questions affect the Chinese speakers’ performance in similar ways whether the calculation distance is short or long. Arithmetic primes always induce faster responses of the following temporal calculations, indicating that the Chinese speakers were probably using the arithmetic strategy in their calendar problem solving.
For the English speakers, distance did influence their performance and the picture is more complicated than it is for the Chinese speakers. The results in each condition are presented in Figure 4-9, with the Short Distance condition on the left side and the Long Distance condition on the right side. They show similar patterns in the Within conditions, but in the Across condition calculations, the responses were very different according to the length of the calculations.

![Graph showing reaction times for English speakers in different Prime conditions for Short and Long Distance calculations](image)

**Figure 4-9: English speakers’ reaction times in different Prime conditions for Short (left) and Long (right) Distance calculations**

For the Short Distance calculations, no effect of Prime Type or Boundary Type is found, but there is a main effect of Direction, $F_1(1, 32) = 20.49, p < 0.001$, $F_2(1, 28) = 10.331, p = 0.003$, which is also found in the Long Distance condition, $F_1(1, 32) = 20.49, p < 0.001$, $F_2(1, 28) = 12.103, p = 0.002$; Forward is faster than Backward. What is more important is that there is a main effect of Boundary Type found only in the Long Distance condition, $F_1(1, 32) = 58.106, p < 0.001$, $F_2(1, 28) = 43.789, p < 0.001$, in which the Within Boundary questions were answered much faster than the Across Boundary.
questions. This suggests that when the calculation distance becomes longer, the English participants might tend to use an arithmetic strategy that is sensitive to boundary types due to the modulo calculations, while when the calculation distances are short, they would probably stick to the more direct list-reciting strategy, so that the response times were about the same when counting either within or across the year boundary. The different Boundary effects can be seen clearly in Figure 4-10. The reaction times in the Short Distance calculations (left side) are about the same for the Forward and Backward tasks, whereas the responses to the Across questions were much slower than those to the Backward ones in the Long Distance condition (right side).

Figure 4-10: English speakers’ reaction times for Short (left) and Long (right) Distance calculations by Boundary Type

Checking the influence of distance in the calculations by taking it as a categorical factor revealed the strategy use difference between the Chinese and English speakers. In order to further analyze what specific strategies might be used by the two language groups in solving calendrical problems, it is informative to take a look at the reaction times with continuous distances between 2 and 9. The reaction time data is presented below in Figure 4-11 for the Chinese speakers and Figure 4-12 for the English speakers.
For the Chinese speakers, the Prime Type effect is not fully presented in the Within Boundary calculations, as the two lines for List and Arithmetic Primes intervene with each other. This is probably because the participants were already very fast in solving the Within Year questions, and in this situation the facilitation of the Arithmetic Primes did not have enough space to have an effect. When it came to the more difficult Across Boundary tasks, which cost more time to calculate, the Prime effect is clearly
observed—the line of reaction times of the Arithmetic Primes is below that of the List Primes in an almost consistent way, especially in the most difficult Across Backward condition. This is compatible with the hypothesis that the Chinese speakers were overwhelmingly using the arithmetic strategy in temporal distance calculations, and thus their responses were facilitated by encountering an arithmetic problem.

Figure 4-12: English speakers’ reaction times by Distance in each Boundary and Direction condition (x-axis is distance; y-axis is mean reaction time in milliseconds)
For the English speakers, the Prime effect is mainly in the Within Boundary conditions, as shown in the upper parts of Figure 4-12. For the Within Year Boundary calculations, List Primes generally induced faster responses. Although the responses were slower in the Arithmetic Prime condition, the slopes tend to go down when the distance is longer, especially in the Backward condition, suggesting that at least some English speakers might be using the arithmetic strategy, thus producing shorter reaction times, because reciting the month-list backwards is difficult. Moreover, the use of the list-reciting strategy might be even less popular in the Across Boundary tasks, as the Prime effect of List disappears.

No substantial difference between the two Prime Types is found in the Across Boundary Short Distance calculations, yet the picture is complicated with increasing distance. It is clear that at least for the distance 6 calculations, Arithmetic induced faster responses in both the Backward and Forward conditions, indicating that the English speakers were sensitive to 6-month-before-or-after calculations and were probably transferring the temporal problem to numerical form for the half-year calculation. In a word, the complicated picture indicates that a variety of strategies were used among the English speakers when they encountered the different temporal calculation tasks.

4.1.6 **Discussion**

The results from Experiment 5 are instructive in several ways. First, the experiment as a whole proves that the priming methodology works well not only with tasks such as word recognition or production, but also with more complicated tasks such as temporal calculations. When the prime and critical questions were solved by employing the same type of strategy, the response is faster than when the prime and
critical questions are solved by using different strategies. This implies, more specifically, that when an arithmetic question is answered, the part of the brain in charge of mathematical functioning might be activated, facilitating the subsequent solving of a calendrical problem that requires an arithmetic strategy. In the same way, if the brain has been carrying out list reciting of some sort (e.g., alphabetical), the reactions to an immediately following recitation of another kind of list (e.g., temporal) would be faster. It seems, however, that the accuracy of the responses to the subsequent questions would not be significantly affected in any direction by the priming effect.

Second, this experiment has replicated the Language effect in both accuracy and reaction time analyses that was found in Experiment 4, confirming the overwhelming use of arithmetic strategies by the Chinese speakers, which results in faster responses in all of the conditions. And a more important finding is that it is an interaction between Language and Prime Type, rather than a main effect of Prime Type, which further confirms that the reaction time difference between the English and Chinese speakers was not the result of different abilities to access the use of arithmetic strategies, but of the different strategies they employ. The different priming effects—of Arithmetic Priming with the Chinese speakers (in the Across Year condition) and List Priming with the English speakers (in the Within Year calculations)—demonstrate that it is not the case that the English speakers are using arithmetic strategies in most cases as the Chinese speakers do in temporal distance calculation but with less success and efficiency. On the contrary, the Chinese and English speakers actually tend to use online strategies in different ways, due to the different levels of transparency of the numerical and temporal calendrical terms in their native languages. Chinese speakers were mostly using the
arithmetic strategy, and this tendency is so strong that Direction and Distance can only alter their strategy choice in a very subtle way. English speakers, on the other hand, were much more flexible in strategy use.

Third, the finding of a significant Priming effect with the English speakers in the Within Year but not in the Across Year calculations is consistent with the hypothesis that the English speakers were able to switch between different strategies according to the nature of the calendrical question in terms of Boundary, Direction, or Distance. When list-reciting became a less efficient method, such as in the longer distance calculations, the English speakers would turn to the arithmetic method. As Figure 4-12 shows, in the Across Forward condition, List and Arithmetic Primes caused no difference in the reaction times when the calculation’s distance was from 2 to 4. Arithmetic Primes did, however, induce faster responses when the distance to be calculated was more than 5. Although the difference was not statistically significant, it shows the trend of strategy shifting. The situation in the Across Backward condition is even more complicated, as the intervening lines of the two Prime conditions probably indicate varying strategies among the English speakers. It is highly likely that some were counting forward with the complement of 12, while others were mapping month names to numbers and using arithmetic strategies. Another piece of evidence that supports the suggestion that they used arithmetic in long distance calculations is the Boundary effect found in the Long Distance but not in the Short Distance calculations. To recite the month list within or across a year boundary should not cause any differences in reaction times, so the Boundary effect that was found is an indication of the use of arithmetic strategies, which would cause longer responses due to the modulo calculations. In a word, the complicated
picture produced by the factor of distance reveals versatile strategy use by the English speakers. By contrast, Chinese speakers’ reactions in the two Prime conditions were almost consistent—responses were faster after encountering Arithmetic Primes in the Across condition, as shown in Figure 4-11. As discussed in Section 4.1.5, the lack of Priming effect in the Within condition might be due to a ceiling effect. The same pattern of reactions—despite Direction and Distance—demonstrates the almost unvarying use of arithmetic strategies by the Chinese speakers in solving temporal calculations.

Therefore, the results of Experiment 5 not only prove once again that the two populations perform differently in time calculation tasks, but also make it possible to tease apart exactly what strategies they use that cause them to perform differently. There is one interesting finding that raises a question that the current experiment cannot give definitive answers to and that will need to be answered by future studies. As Section 4.1.5 shows, the Chinese speakers were not sensitive to the compatibility of Prime Direction and Critical Direction; however, the English speakers’ reactions were affected by that compatibility. It is not clear why the English speakers’ reactions are sensitive to the compatibility of Prime Direction and Critical Direction, but it might be related to the various strategies they were using, as the Chinese speakers, who overwhelmingly use arithmetic strategies, were not affected by the compatibility. Definite answers must wait for further empirical studies, which might employ methodologies such as mouse trajectory tracking (Spivey & Dale, 2006; Spivey, Dale, Knoblich, & Grosjean, 2010) to investigate online processing of temporal problem solving.
CHAPTER 5
GENERAL DISCUSSION

This study has explored how language is influenced by and influences general cognition. More specifically, it explores whether the numerically transparent naming of time concepts in a language confers advantages on the language’s speakers in their acquisition of time concepts and reasoning about time. Even more specifically, it tests whether the use of numerical systems in the names of time concepts affects how children acquire time concepts and how adults reason about time. For example, in Mandarin Chinese, the word for Monday is literally “weekday one” and January is “one month,” whereas English uses arbitrary names with astronomical (“Monday”) or mythical (“Thursday”) origins. Children generally learn basic numbers and counting before acquiring the time sequences, and therefore, the relationship between number and time-word terms may affect at what age children acquire the time concepts.

Chapter 2 reports a series of acquisition studies that test whether numerical time naming systems facilitate children’s acquisition of time sequences by comparing the behavior of monolingual Chinese-, English-, Latvian-, and Korean-speaking children. The results demonstrate that, first, Chinese monolingual children comprehend and use time words at an earlier age than English monolinguals. Second, in Latvian, which has numerical names for days of the week but arbitrary names for months of the year, children acquire days of the week earlier than months of the year, showing that the effect of naming systems can occur even within one language. Third, in Korean, in which the two systems are the reverse of the Latvian systems (arbitrary day names and numerical month names), the children’s performance on DOW and MOY tasks shows the opposite
pattern as the Latvian data, providing strong support for the naming systems’ effects and demonstrating that these effects do not result from other factors, such as word frequency or accessibility.

Chapter 3 and Chapter 4 describe a set of behavioral experiments (Experiments 4 and 5) using reaction time measurements and a priming methodology to investigate whether Chinese- and English-speaking adults’ time calculation abilities and the online strategies they employ while performing temporal reasoning are influenced by their language’s time naming system. For Experiment 4, reaction times, strategy use, and error distribution results indicate that there are differences between English and Chinese speakers—Chinese speakers may more readily employ arithmetic when calculating because their temporal naming systems relate transparently to numerals; this in turn improves the speed and accuracy of their time calculations. English speakers, on the other hand, use other strategies, such as sequential recitation, which may lead to slower responses. In Experiment 5, Chinese speakers were faster in answering time calculation questions after encountering Addition and Subtraction Primes (in the Across Year calculations). By contrast, English speakers’ reactions were facilitated after doing alphabetic list reciting, but only in the Within Year calculations. The List Primes did not affect their response times in the Across Year calculations, indicating a possible mixed use of strategies and the ability to switch between strategies in different conditions. These results confirm that Chinese and English speakers may employ different strategies in solving time distance calculations. Chinese speakers probably unvaryingly use arithmetic strategies in temporal distance calculations, while English speakers may mainly use the list-reciting strategy but can switch to other methods as needed.
To summarize, the dissertation research addresses a fundamental question about the relationship between language and cognition, namely whether linguistic coding of a concept in different languages shapes how people acquire and subsequently use that concept.

5.1 Implications

5.1.1 Educational Implications

The acquisition studies reported in this dissertation investigate the link between the early mastery of specific nomenclature systems, such as simple numeric sequences, and the subsequent acquisition of more cognitively complex systems, such as time concepts. This research sheds light on the constructional nature of human knowledge, where more complex systems develop on the scaffolding of less complex, previously learned systems.

The study is important for understanding how children in different cultures have different forms of “linguistic infrastructure” to transfer to or transform related cognitive processes. Scholars and educators thus should pay attention to distinguishing the phenomena that reflect the structure of the child’s language from the phenomena that reflect universal aspects of cognitive development. The findings could help inform ways in which such seemingly simple (but in fact quite complex) temporal reasoning processes may be augmented in some languages by pedagogical approaches that draw explicit connections to other representational systems that are not expressed in the native language of the children. Thus, the research could have benefits for non-academic stakeholders, as its findings should encourage early childhood educators to help children
establish links between numerical and temporal domains, which may help children develop more diverse and efficient strategies in temporal problem solving.

In a broad perspective, the results are of interest to the general public as well as scholars, as they point to advantages conferred on speakers based purely upon the language that they speak. Given the importance of multilingual education and cross-cultural exchange between the English-speaking world and China, this research will provide useful insights that might shape language policy as well as individual parents’ decisions on the languages they wish their children to acquire.

5.1.2 Linguistic Relativity

The two behavioral experiments with adults investigate whether the way a system is learned affects how it is used in adulthood. This part of the dissertation revisits the famous Sapir-Whorf hypothesis, which claims that human languages differ from one another in considerable ways in how they describe the world and that cross-linguistic differences in lexicon and grammar have nonlinguistic consequences. The research adds new information to the on-going discussion of this theory. In general, the findings support the hypothesis that linguistic differences can produce non-linguistic consequences, in this case by affecting people’s reasoning about time.

The exact nature of the relationship between language and cognition has always been a matter of debate. Since the 1980s, empirical studies in the cognitive linguistic field show evidence both for and against the notion of linguistic relativity initially proposed by Edward Sapir (1958 [1929]) and developed by Benjamin Whorf (1956). The idea is that the particular language that people speak affects thought. Among the many
speculations on the impact of language on thought, the following passage might be the best-known:

Human beings do not live in the objective world alone, nor alone in the world of social activity as ordinarily understood, but are very much at the mercy of the particular language which has become the medium of expression for their society. It is quite an illusion to imagine that one adjusts to reality essentially without the use of language and that language is merely an incidental means of solving specific problems of communication or reflection. The fact of the matter is that the “real world” is to a large extent unconsciously built upon the language habits of the group….we see and hear and otherwise experience very largely as we do because the language habits of our community predispose certain choices of interpretation. (Sapir, 1958 [1929], p. 160)

The strong version of linguistic relativity (also known as linguistic determinism) has long been abandoned. Recently, however, there has been a surge of systematic and compelling behavioral studies using fine-tuned methods to show that even little quirks of language influence speakers’ perception (Boroditsky, 2003; Casasanto, 2010), attention (Chan & Bergen, 2005), categorization (Lucy, 1992; Imai & Gentner, 1997; Roberson, Davies, & Davidoff, 2000; Lucy & Gaskins, 2001), spatial reasoning (Levinson, 1996; Pederson et al., 1998; Boroditsky, 2001; Levinson et al., 2002), numerical reasoning (Miller & Paredes, 1996; Gordon, 2004; Pica, Lerner, Izard, & Dehaene, 2004) and the acquisition process (Choi, McDonough, Bowerman, & Mandler, 1999). More relevant to the current dissertation research, languages also affect people’s reasoning about time (Boroditsky, 2000, 2003; Boroditsky & Ramscar, 2002, Matlock et al., 2005; Núñez & Sweetser, 2006; Casasanto & Boroditsky, 2008). Therefore, although the precise psychological implications of these studies continue to be debated in the literature (e.g., Gumperz & Levinson, 1996; Levinson et al., 2002; Li & Gleitman, 2002; Chen, 2007), the cross-linguistic evidence does show that language affects some aspects of general cognition in certain ways.
However, most of the ongoing discussion of linguistic relativity has missed some of the important ways in which linguistic structures contribute to cognitive development (Miller & Paredes, 1996). Symbolic communication is one of the most important higher cognition abilities in the vast network of complex cognitive skills that humans possess. How symbol systems such as calendar terms are organized and learned may have consequences for their users’ ability to perform certain tasks. The current study has shown that the way calendar symbols are coded in the languages children acquire has significant impacts on adults’ temporal reasoning with calendar calculations in terms of speed and accuracy rates. The marked linguistic difference also has a psychological impact on how the tasks are viewed and what cognitive strategies, or habits of thought, are recruited to solve a given type of cognitive problem.

In a word, the current research offers a new piece of evidence for the pervasive influence of language on thought, in the specific domain of cognition of time, by showing that the way calendars are coded can have a substantial effect on both the developmental trajectory and the subsequent employment of strategies in problem-solving processes.

5.2 Limitations and Possible Future Studies

People make use of richly structured cultural contexts as they develop cognitive abilities (Tomasello, 1999; Norenzayan, Schaller, & Heine, 2006). To the extent that societies diverge in their linguistic conventions, so would cognitive processes, to some degree, as shown by the research presented here.

Clearly, however, language is an indispensable part of culture, and linguistic conventions are almost always under the influence of cultural conventions. As a result, the examining of cognitive processes that mediate the linguistic shaping of thought can
hardly be considered to only measure linguistic effects sans implicit cultural effects. For instance, after reading about this research, people may even ask whether there might be a cultural basis for the linguistic differences between the types of names the Chinese language uses for weekdays and for months. The Chinese weekday names used to use the “Seven Luminaries”—a system that referred to arbitrary planetary names. With the ending of the last dynasty and the founding of the Republic of China in 1911, Monday through Saturday in China came to be numbered one through six, with the reference to the sun remaining for Sunday. This kind of linguistic change is clearly a consequence of social and cultural change.

Therefore, more various and refined methods still need to be developed for controlling possible confounding factors in the acquisition studies designed to investigate this research topic. It is important to distinguish linguistic effects from other cultural effects on thought—for example, effects due to educational practices that vary according to economic and geographical factors. This is particularly challenging given that non-linguistic cultural patterns and linguistic conventions tend to be intertwined.

I here propose some procedures that I intend to implement to improve my next acquisition studies. For example, a short interview with the children’s kindergarten/preschool instructors will be conducted to collect background information about possible cross-cultural differences. Questions will include: (1) How much time do the children spend doing homework at home? (2) How much does the home life emphasize education? (3) How much mathematics instruction do the children get, and in what way? (4) How are days and months being explained to the children? The data
collected will serve as an important source for understanding factors other than linguistic differences that could generate the results found in the next stage of this research.

However, no matter how much background information is collected, it is very difficult to control cultural and educational variables. It is desirable, therefore, to compare the acquisition of time terms and time concepts across two languages within individuals; that is, in bilinguals. I propose to use the same tests developed for the studies described in this research with Chinese-English bilingual children, which will eliminate some uncontrolled factors in the current studies. In addition, investigating bilinguals will allow me to tease apart whether it is the learning of time words or of time concepts that happens at different ages for the different languages. If bilinguals perform differently on the same task in the two languages, then this suggests that it is the terms, and not the concepts, that are acquired at different rates depending on numerical transparency of time terms. But if bilinguals perform equally well on the task in their two languages, then it is likely that monolinguals who show different acquisition rates depending on their language are actually learning the concepts at different rates.

Another benefit of conducting both the DOW and MOY studies with English-Chinese bilingual children is that it will establish a baseline for comparison with the Latvian and Korean data. In contrast with Latvian and Korean, in both English and Chinese the MOY and DOW systems have the same level of transparency, which means that any differences in the ages at which the two systems are acquired are due to factors other than numerical transparency. The results from testing bilinguals may tell us more about the advantages of acquiring a second language that has different structural properties from the first language.
Another line of future studies will use the methods described for the behavioral studies in this research, but will test different, second-language learning populations. The project will involve testing Chinese-speaking learners of English and English-speaking learners of Chinese in their non-dominant language with the behavioral tests designed for the dissertation research. An interesting question that arises here is whether reasoning about time will be affected by language uni-directionally (with a facilitation effect only from the more systematic way of coding, i.e., from Chinese to English) or bi-directionally (i.e., facilitation from Chinese to English and inhibition from English to Chinese). I hypothesize that with better command of Chinese, the second language learners will have more-Chinese-like performance on the tests—that is, shorter reaction times and lower error rates—because the learning of a more systematic way of conceptualizing time may give them a new way of perceiving time sequences and performing temporal tasks using other strategies.

The future research described here will tell us more about the effect of naming systems on acquisition of and reasoning about time concepts, contributing to the study of language development and cognitive linguistics.
APPENDICES

APPENDIX A: DAYS OF THE WEEK EXPERIMENT STIMULI (ENGLISH VERSION)

Picture Cards:

Questions:
Part 1  Levels 1-3:
1. How many days are there in a week? _______1
2. On what day does Winnie the Pooh climb a tree? _______2
3. How many days in a week do you go to school? _______1
4. On what day does Winnie the Pooh fly a balloon? _______2
5. On what day does Winnie the Pooh read a book? _______2
6. How many days in a week do you stay home? _______1
7. Today is Tuesday and Winnie the Pooh flies a balloon. What will he do tomorrow? _______3
8. On what day does Winnie the Pooh go dancing? _______2
9. How many days in a week do your parents go to work? _______1
10. On what day does Winnie the Pooh eat some honey? _______2
11. Today is Friday and Winnie the Pooh goes swimming. What will he do tomorrow? _______3
12. Today is Tuesday and Winnie the Pooh flies a balloon. What did he do yesterday? _______3
13. How many days of the week do you know the names of? _______1
14. Today is Wednesday and Winnie the Pooh reads a book. What will he do tomorrow? _______3
15. Today is Sunday and Winnie the Pooh eats some honey. What did he do yesterday? _______3

Part 2 Levels 4-5:
1. Today is Wednesday and Pooh reads a book. On Saturday Pooh climbs a tree. How many days must he wait to climb a tree? _______4
2. Today is Tuesday and Pooh flies a balloon. On Thursday Pooh goes dancing. How many days must he wait to go dancing? _______4
3. Today is Thursday and Pooh goes dancing. How many days must he wait to go dancing again? _______5
4. Today is Sunday and Pooh eats some honey. On Friday Pooh went swimming. How many days ago did he go swimming? _______4
5. Today is Monday and Pooh rides in a boat. On Wednesday Pooh reads a book. How many days must he wait to read a book? _______4
6. Today is Saturday and Pooh climbs a tree. Next Monday Pooh rides in a boat. How many days must he wait to ride in a boat? _______5
7. Today is Saturday and Pooh climbs a tree. On Thursday Pooh went dancing. How many days ago did Pooh go dancing? _______4
8. Today is Sunday and Pooh eats some honey. Next Tuesday Pooh flies a balloon. How many days must he wait to fly a balloon? _______5
10. Today is Saturday and Pooh climbs a tree. Next Tuesday Pooh flies a balloon. How many days must he wait to fly a balloon? _______5
APPENDIX B: STIMULI OF THE POSTTEST (DATE AND HOUR CALCULATIONS)

1. Winnie the Pooh is going to visit Tigger by boat. Now it is 5 o’clock and the train will leave at 10. How many hours does Winnie the Pooh have to wait? (5)

2. Winnie the Pooh is going to swim with his friends. Now it is 8 o’clock and his friends will come at 12. How many hours does Winnie the Pooh have to wait? (4)

3. Winnie the Pooh is going to have a party with his friends. Now it is 6 o’clock and his friends will come at 9. How many hours does Winnie the Pooh have to wait? (3)

4. Winnie the Pooh ate some honey at 9 o’clock. Now it is 11 o’clock so how many hours ago did Winnie the Pooh eat honey? (2)

5. Winnie the Pooh read his book at 2 o’clock. Now it is 6 o’clock so how many hours ago did Winnie the Pooh read his book? (4)

6. Today is the 4th of December. Six days from now Mickey Mouse has a soccer contest. What date will that be? (10)

7. Today is the 5th of December. Three days from now Mickey Mouse has a painting contest. What date will that be? (8)

8. Today is the 7th of December. Five days from now Mickey Mouse is going to give Minnie a gift. What date will that be? (12)

9. Today is the 9th of December. Four days ago Mickey Mouse had a tour with Minnie. What date was that? (5)

10. Today is the 8th of December. Two days ago Mickey Mouse took a walk. What date was that? (6)
APPENDIX C: MONTH OF THE YEAR TEST STIMULI (ENGLISH VERSION)

Picture Cards:

January | February | March | April
---------|----------|-------|------
[Images of Mickey Mouse with snowman, train, soccer ball, paintbrush]

May | June | July | August
-----|------|------|--------
[Images of Mickey Mouse driving, Easter bunny, umbrella, rainbow]

September | October | November | December
-----------|---------|-----------|----------
[Images of Mickey Mouse with backpack, walking, taking photo, gift]

Questions:
Part 1  Levels 1-3:
1. How many months are there in a year? _______1
2. In what month does Mickey Mouse like to go walking? _______2
3. How many months out of the year do you go to school? _______1
4. In what month does Mickey Mouse like to play football? _______2
5. In what month does Mickey Mouse like to drive a car? _______2
6. How many months out of the year do you stay home from school? _______1
7. Now it is April and Mickey Mouse likes to paint. What will he do next month? _______3
8. In what month does Mickey Mouse like to visit friends? _______2
9. How many months out of the year do your parents go to work? _______1
10. In what month does Mickey Mouse like to make movies? _______2
11. Now it is August and Mickey Mouse likes to walk in the rain. What will he do next month? _______3
12. Now it is March and Mickey Mouse likes to play football. What did he do last month? _______3
13. How many months of the year do you know the names of? _______1
14. Now it is May and Mickey Mouse likes to drive a car. What will he do next month? _______3
15. Now it is December and Mickey Mouse likes to deliver gifts. What did he do last month? _______3

Part 2 Levels 4-5:
1. Now it is May and Mickey likes to drive a car. In September Mickey likes to build things. How many months must he wait to build things? _______4
2. Now it is April and Mickey likes to paint. In July Mickey likes to play with a big fish. How many months must he wait to play with a big fish? _______4
3. Now it is June and Mickey likes to visit friends. How many months must he wait to visit friends again? _______5
4. Now it is November and Mickey likes to make movies. In August Mickey walked in the rain. How many months ago did he walk in the rain? _______4
5. Now it is January and Mickey likes to make a snowman. In May Mickey will drive a car. How many months must he wait to drive a car? _______4
6. Now it is September and Mickey likes to build things. Next February Mickey will ride in a train. How many months must he wait to ride in a train? _______5
7. Now it is October and Mickey likes to go walking. In July Mickey played with a big fish. How many months ago did Mickey play with a big fish? _______4
8. Now it is December and Mickey likes to deliver gifts. Next April Mickey will paint something. How many months must he wait to paint something? _______5
9. Now it is August and Mickey likes to walk in the rain. Next January Mickey will make a snowman. How many months must he wait to make a snowman? _______5
10. Now it is October and Mickey likes to go walking. Next March Mickey will play football. How many months must he wait to play football? _______5
REFERENCES


