Using space to conceptualize time and other domains crosslinguisticly

By

Yunzhijun Yu

Department of Psychology, University of Colorado at Boulder

Defended March 19, 2014

Thesis Advisor:
Dr. Eliana Colunga, Department of Psychology

Defense Committee:
Dr. Richard Olson, Department of Psychology
Dr. Martin Goldman, Department of Physics
Dr. Andrea Feldman, Program for Writing & Rhetoric

TABLE OF CONTENT

Abstract

The question of whether language shapes the way people think of the world has been under debate for decades. Mandarin Chinese speakers use vertical metaphors to talk about time more often than English speakers do (Chen, 2007). Boroditsky (2001, 2008) demonstrated that this linguistic difference in metaphors causes Mandarin speakers to have vertical mental representations of temporal relations more frequently than their English counterparts. Nevertheless, some studies (Chen, 2007. January & Kako, 2007) failed to replicate Boroditsky's findings. These studies raised questions as to the replicability of Boroditsky'findings. Furthermore, it was questioned whether Boroditsky's experimental designs captured only the linguistic cues rather than other factors such as Mandarin speakers' preference on rank. While linguistic cues might play a role in determining speakers' mental representation of time, other factors, such as cultural preference on ranking items, must also be taken into account and carefully examined. This study adapts the experimental paradigm of Boroditsky (2008) and examines how speakers' mental representations of time, ranking, and nominal categories onto a two dimensional space. The purpose of this study is to determine whether native Mandarin and English speakers think of time differently and, if so, whether that difference is limited to the factor of language alone.

Keywords: linguistic relativity, English, Mandarin Chinese, spatial temporal metaphors, language and time

Our ability to understand abstract ideas is one of the most mysterious aspects of human minds. Psychologists have been putting much effort into researching how abstract concepts, such as the concept of color, are understood in human cognition. Some linguistic and psychological researchers have tried to use language to explain the way we conceptualize the world around us. Language is one of the most fascinating gifts that has been given to the human race, as it is a highly structured mechanism that is cognitively learned through our experiences of the world. This is similar to the learning of abstract concepts, which are obtained and conceived through our experiences as well.

The approach to understanding the relationship between language and mind has been tremendously discussed and studied since the early 1900s (Chomsky, 2006). We have the ability to perceive, incorporate, and reproduce the world in the form of language, even on the abstract areas such as space and time. Chomsky believes that the languages we speak functionally interact and construct our perceptions of the physical into abstract cognitions and representations that our minds are able to conceive.

One of the most influential studies, which suggests the language we speak influences our cognition, is Boroditsky's Mandarin time conception study (2001). The results of this study indicated that English speakers almost always prefer horizontal representations of time, whereas Mandarin speakers actually prefer vertical representations of time 7 to 8 times more often. Boroditsky argued that this effect was due to the language differences between English and Mandarin, as the Mandarin language uses more vertical metaphors to talk about time.

This study became very successful and has been cited in numerous scholars' research. Nevertheless, it was later brought into question after two other replications that failed to find the same effect (Chen, 2007; January & Kako, 2007). These studies challenged the validity of Boroditsky's conclusion by addressing some limitations and questioning variables that appeared in her study. Boroditsky responded to these articles in another study (2008). In this

study, she directly asked native English speakers and Mandarin speakers to point out temporal events in space and recorded their responses with “horizontal” or “vertical”. This new study found the same effect that appeared in Boroditsky’ original paper (2001), which suggests that Mandarin speakers are more likely to conceptualize time in vertical representation than English speakers are. Nevertheless, it is still debatable whether the observed difference was truly due to the language factor alone. For instance, if the preference for vertical representations that is observed in Mandarin speakers persists to other domains like rank or nominal category (for example: pen, fruit, etc.), then this particular preference may just be a cultural difference instead of direct consequence of language metaphors.

The present research project came up with a new experimental design that was developed based on Boroditsky’ experiment (2008). It examined more than linguistic factors, including rank and category, to detect if Mandarin speakers use the vertical dimension more often when they think about time, and if this preference extends to other domains beyond time. The hypothesis is that native Mandarin speakers will map time onto vertical relations more often than native English speakers do, because the Mandarin language uses more vertical metaphors to talk about time. However, for other domains such as rank and nominal category, it is expected that both Mandarin and English speakers will prefer horizontal and vertical representations equally, as the metaphorical difference between Mandarin and English does not extend to these domains. While trying to replicate the original study of Boroditsky (2008), this research provides further insights into the validity of the current debate and will serve as a reference for future research.

THE SAPIR-WHORF HYPOTHEIS AND LINGUISTIC RELATIVISM

The idea that the language we speak enables us to sort out and categorize the world has been discussed and debated in various articles and literatures. Amy Tan, a first generation American Chinese who was raised in California, described her experiences that were

associated with this idea in her famous article "The Language of Discretion" (Tan, 1990). In the article, Tan used different situations and examples from her bilingual experiences to illustrate the idea that individuals are convolutedly related to particular languages and cultures. For example: *"In English, we see "cats" and "dogs"; what if the language had also specified* ***glats****, meaning 'animals that leave fur on the sofa', and* ***glotz****, meaning 'animals that leave fur and drool on the sofa'?"* Through this funny but vivid example, Tan brought up a question: how would describing the animals differently in these two languages change their speakers' conceptions of the animals with slight vocabulary variations? In other words, are we more likely to think of dogs as animals that leave fur and drool, when we speak this hypothetical language as compared to when we speak English?

This idea traces back to the Sapir-Whorf hypothesis that was developed in the 20th century. During that time, an American anthropologist-linguist, Edward Sapir, and his student, Benjamin Whorf, hypothesized that the way individuals think is strongly affected by their native language. Furthermore, as each language in the world differs on their grammatical and semantic meanings, there are certain thoughts of an individual in one language that cannot be interpreted by an individual who speaks a different language (Kay & Kempton, 1984, Sapir, 1951[1929], Whorf, 1956[1940]).

The Sapir-Whorf hypothesis later divided into two slightly different hypotheses. The first hypothesis is called linguistic determinism. Linguistic determinism suggests that individuals' thoughts are largely determined by the categories made available by their native language (Pinker, 1994). Linguistic determinism is the extreme form of the Sapir-Whorf hypothesis as it implies that the language an individual speaks directly determines and constrains that individual's thinking and conceptualizations of the world. According to this hypothesis, if a child is born in the wilds and raised by a group of wolves, his cognitive abilities would be extremely limited. If this child looks up at the sky, he would not be able to

have thoughts like “how beautiful this sky is”, because he does not speak a language that makes the concepts of “sky” or “beauty” available for his cognition. The hypothesis of linguistic determinism was criticized by Pinker, who argued that this hypothesis is barely testable and cannot be used to model multi-linguals’ cognitive processes.

More researchers and scholars embrace the hypothesis of linguistic relativism, which is the other hypothesis developed from the Sapir-Whorf hypothesis. Linguistic relativism is considered to be the less extreme version of linguistic determinism. Instead of suggesting that our language determines how the world is formed in our cognition, linguistic relativism believes that language plays the role of setting up a predisposition. Specifically, the hypothesis of linguistic relativism suggests that the differences among languages cause differences in the thoughts of speakers. Among the studies that focus on this hypothesis, one of the major approaches that serve as the central direction for investigations is the domain-centered approach (Lucy, 1997).

According to Lucy’s definition, a domain-centered approach usually begins with a certain domain of experienced reality, and asks how different languages would encode or construct this particular reality. Typically, speakers of various languages are asked to refer to the same materials or situations, and the experimenters will observe the different patterns that speakers will exhibit. This approach is often used in various experiments due to its strength of precision and control. It allows researchers to facilitate rapid and certain comparisons among large sets of languages. However, this approach faces a serious issue: in seeking influences on thought, studies adopting this approach often have difficulty establishing the significance of purported effects, as the approach sometimes might emphasize what is *possible* for participants to express, and overlook what is habitually structured and expressed. Nevertheless, due to the natural advantage of the domain-centered approach, which is the ease of experimental control, this approach is widely used in current psycholinguistic

research.

TIME METAPHORS IN ENGLISH AND MANDARIN CHINESE

As mentioned earlier, the Sapir-Whorf Hypothesis has been under debate for decades. For complicated issues such as attitudes and values, it is difficult to account for the role of culture on language and assess directly how the structure of language might affect speakers' cognition. However, it is possible that the language we speak impacts our thoughts on some basic cognitive level, such as how we conceptualize time.

One way to research the impact of linguistic difference between languages regarding time is to use the domain-centered approach that was mentioned earlier. In a well-controlled experiment, Boroditsky (2001) examined the representation of time in Mandarin speakers and English speakers. In this study, Boroditsky discussed the spatial metaphors of temporal relations that are used in Mandarin and English. Specifically, English speakers usually use front/back terms to talk about time (e.g., "The meeting is AHEAD", "I will leave the memories BEHIND", etc.), and this spatial relation is horizontal. However, in Mandarin Chinese, different meanings usually are combined into one letter. For instance, the letter "上" means "up" in spatial context, but it also means "last" when it is used in a temporal context (e.g., "上面" – "Up there" vs. "上个月" – "Last month". The same character "上" that describes a vertical spatial relation is also used to describe time or events that happened in the past). To test if this difference has any effect on the speakers of these two languages, Boroditsky measured the response times of the speakers after exposing them to horizontal and vertical priming task. In a typical trial of the priming task, the participants saw two pictures in a row, each depicting two objects aligned horizontally or vertically, and they saw a sentence describing the spatial relationship of the two objects. The participants' task was to determine if the sentence was a correct statement. The sentence used a spatial metaphor (such as before and after) and a time word (such as earlier and later) (see Figure 1a & 1b).

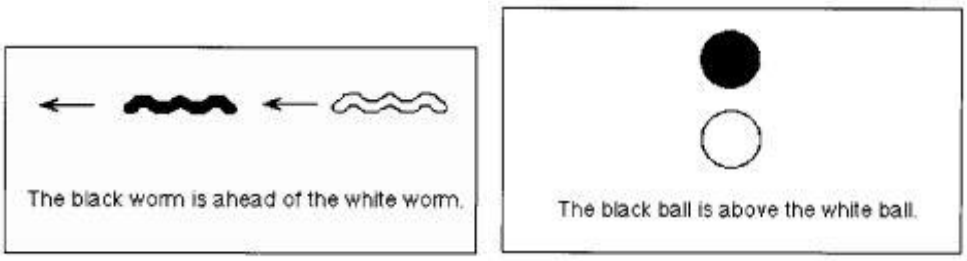


Figure 1a & 1b. Examples of the horizontal spatial prime used in Boroditsky (2001)'s priming tasks

Based on the results, Boroditsky reported that English speakers showed cross-domain priming from horizontal spatial relations to temporal relations, whereas Mandarin-English bilinguals were more likely to show vertical priming. Such a tendency persisted even when the Mandarin speakers were processing the instructions and the primes in English. Moreover, Boroditsky trained a portion of the English participants to think of time vertically by exposing them to several examples of sentences making use of the vertical metaphor for time. The results showed that the trained English speakers showed a similar response pattern to Mandarin speakers. Therefore, Boroditsky concluded that the use of spatial metaphors could change the way speakers think about time. Furthermore, in a regression test, Boroditsky determined that the extent to which Mandarin English bilinguals think about time vertically is related to how old they were when they first began to learn English. The bias to think about time vertically was greater for Mandarin speakers who started learning English later in life. This study directly provides evidence that supports the hypothesis of linguistic relativism.

However, Chen (2007), January and Kako (2007) failed when they attempted to replicate Boroditsky (2001)'s study. The failed attempt made the researchers from these two studies question whether Mandarin speakers really do conceptualize time differently than English speakers do. Chen (2007) pointed out that the use of horizontal spatial metaphors in Mandarin Chinese to express time was actually more frequent than the use of the vertical spatial metaphors. Meanwhile, January and Kako (2007) pointed out that if a short training

session for English speakers could change how time is represented in their minds, then why would years of exposure to horizontal time metaphors not change Mandarin speakers representations? Boroditsky (2001)'s results were therefore challenged due to the questionable replicability.

In order to take a better look at the spatial representations English and Mandarin speakers construct for time, Boroditsky designed a new study (2008). This study used a "pointing" design, which allows the participants to point to the spatial locations of some temporal events. Basically, the researcher pointed to a spot in the air and told the participants: "If this space right here is TODAY, where will you put YESTERDAY?" This design allowed the participants to point anywhere around them in 3-dimensional space rather than 2-dimensional space (Boroditsky, 2001), and instead of forcing participants to conceptualize time in an artificial environment, the design gave its participants more freedom in terms of expressing how exactly time is conceptualized in their minds. This study's finding again confirmed that Mandarin speakers think about time vertically more often than English speakers do, and specifically, Mandarin speakers are eight times more likely to construct vertical representations of time than are English speakers.

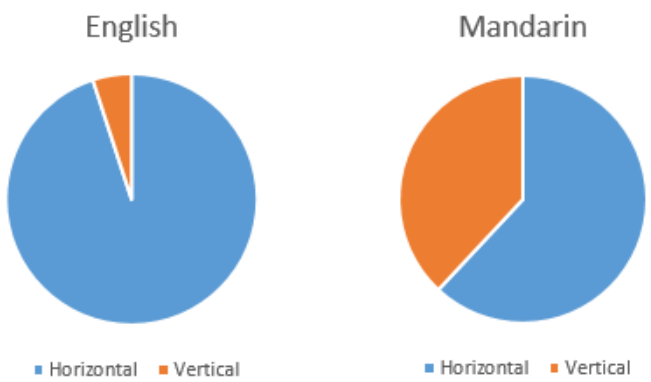


Figure 2. Proportions of vertical and horizontal spatial metaphors used in English and Mandarin. The frequency of vertical metaphors used in Mandarin is about 38%, comparing to 5% in English (Chen, 2007).

One critical assumption that Boroditsky made in her studies, was that Mandarin speakers talk about time vertically more often than English speakers do. In order to refute this assumption, Chen (2007) searched the Yahoo and the Google News Taiwan to estimate the frequency of horizontal and vertical spatial metaphors' usage in Mandarin Chinese expressions. Figure 2 summarizes the results that were found in Chen's study (2007). Although Mandarin speakers most often talk about time using horizontal metaphors, the prevalence of vertical metaphors is significantly greater in Mandarin than in English. This result provided support for the original hypothesis, that Mandarin speakers conceptualize time vertically more often than English speakers do because their native language contains more vertical temporal metaphors.

Although the frequency of vertical metaphors in Mandarin Chinese is higher in English, it is important to discuss the impact that traditional Asian writing direction has on the perceived direction of time. Literature showed that the factor of writing direction indeed has an impact on individuals' perception of time. Native speakers of Arabic and Hebrew read text arranged from right to left are more likely to arrange time from right to left (Fuhrman, & Boroditsky, 2010). Traditionally, Chinese and other East Asian text was usually written in vertical columns arranged from right to left, top to bottom (see Figure 3.).

Figure 3. This ancient Chinese official memorial, "*Chu Shi Biao*", was written around 227 AC. in the Three Kingdom period. It was written from top to bottom, and right to left, according to the official customs during that specific period of time.

This custom was kept until the People's Republic of China officially switched it to international standard in 1956. Since then, all the text in newspapers, books, and online articles is nearly always arranged horizontally from left to right. In Taiwan, the vertical writing system was also switched to horizontal writing system since 2004 (Fuhrman et al., 2011). The impact of the vertical writing system may be lessened due to Mandarin speakers' decreasing exposure to it.

One important point that Boroditsky also discussed in her studies is that the lives and cultural experiences of English and Mandarin speakers differ in various ways beyond differences in language. Therefore, even showing that Mandarin and English speakers' performance differs in these tasks cannot simply be explained by this difference in language. For instance, it may be that Mandarin speakers use vertical metaphors to represent time simply because they were ranking the time units instead of conceptualizing time. Also, Mandarin speakers might use vertical space to represent time more often than English speakers do because they prefer vertical organizations overall. If this is true, then Mandarin speakers should use vertical representations in other domains besides time. Importantly, due to the fact that two independent studies failed to replicate Boroditsky's results, it is still unclear if there is any truth to the claim that native Mandarin speakers do, in fact, think of time differently than native English speakers do. In summary, it is unclear what the

mechanism for the effects are in Boroditsky's studies, and whether other mechanisms (such as culture, preference, etc.) might cause the effects rather than language.

Although one way of approaching this question is to artificially change the distributional patterns in people's language environment and see if this causes a corresponding change in thinking, like the priming tasks in Boroditsky (2001), the other way to approach the answer is to design an experiment that tests other domains besides time alone. In this research project, a new research design that developed from the paradigm in the study of Boroditsky (2008) will be used, and it will serve to determine: 1) whether Mandarin speakers do use vertical representations to think of time more often than English speakers do; 2) whether Mandarin speakers map other domains onto vertical space more often than English speakers do.

Methods

Participants

This study recruited 23 Mandarin-English bilinguals who speak Mandarin Chinese as their native language, and 24 native English speakers at University of Colorado-Boulder. All Mandarin speakers speak English as their second language. One of the native Mandarin speakers was thrown out due to a failure to complete the control condition correctly. All English speakers had low or no exposure to Mandarin Chinese at the time of the experiment. All the participants received $5 compensation for their time.

Materials

Each participant received a set of datasheets that contained different time frames. On each paper sheet, the participants saw a white square on the right with an "X" in the middle (see Figure 4), and the instructions on the left. The paper sheets were attached to a binder that stood on the table, so that the participants would be able to see the paper sheets perpendicular

to the table.

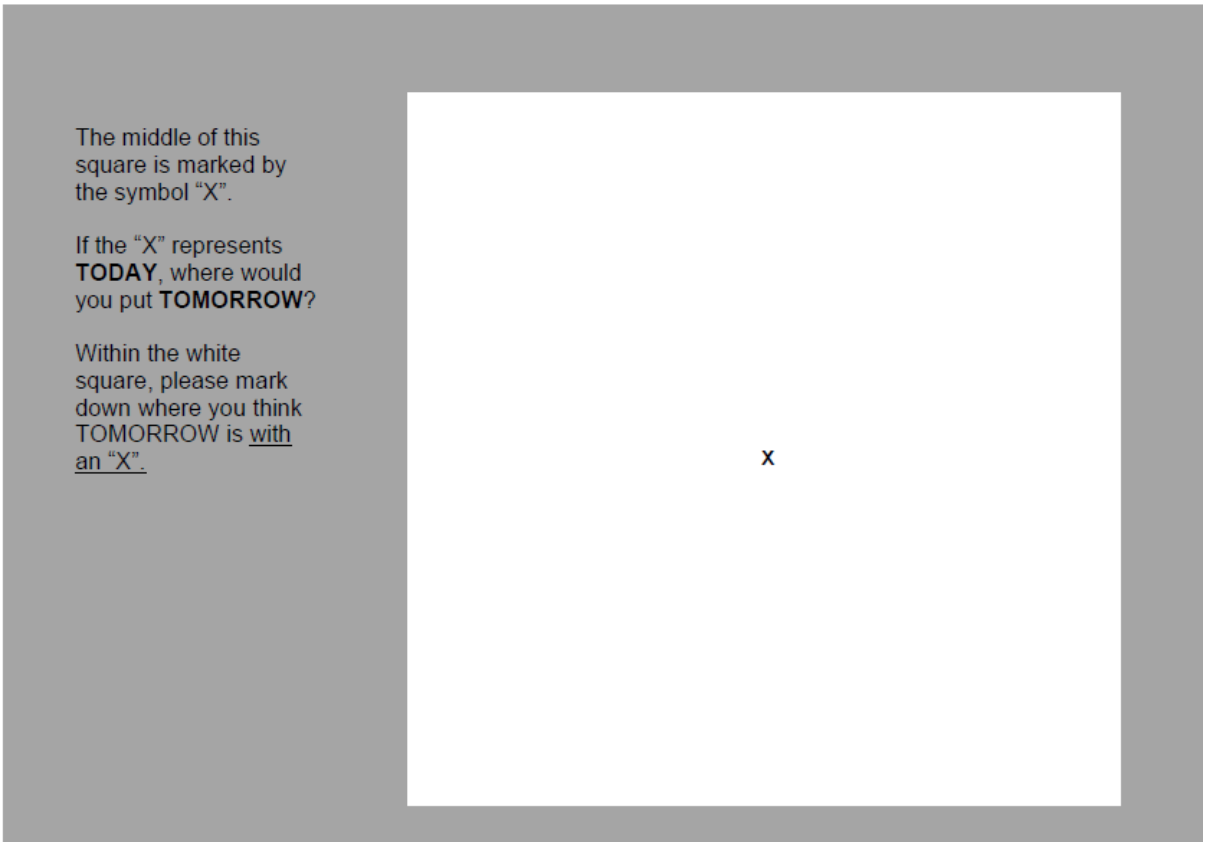


Figure 4. This is a sample datasheet that was presented to participants in the study. A set of 20 datasheets with different words (see Table 1) was presented to participants, and they were asked to mark down where they think the events should be with an "X". All the datasheets were attached to a presenter that stood on the table, so that the datasheet was presented vertically to the participants.

Procedures

The experiment was set up in accordance with the paradigm of Boroditsky (2008). On each paper sheet, the middle "X" would represent a different time (such as "today" or "lunch"), and the participants were asked to mark down the location of another time (such as "tomorrow" or "breakfast"). An example of the instruction is "If the 'X' represents TODAY, where would you put TOMORROW? Within the white square, please mark down where you think TOMORROW is with an 'X'." Each participant was asked to complete 20 questions involving 4 types of domains (time, rank, nominal category, and space) in the experiment (see

Table 1). All the instructions and questions were given in English. The order of the 4 sets of domains was given in a fixed order (Time – Rank – Nominal Category – Space/Control), but the trials within each set were counterbalanced. The order of these 4 sets was fixed to control for priming affects that may have contaminated the results (e.g., showing participants the spatial set first might have primed their responses to the temporal set).

The trial of time was given first and was designed to replicate Boroditsky's study. Although the experiment was designed in a 2D space rather than 3D, it was still expected that Mandarin speakers would mark down the locations vertically more often than English speakers would, which was also shown in the earlier studies of Boroditsky (2001, 2008). The second set, rank, was designed to see there is a difference between Mandarin speakers and English speakers on how they conceptualize rankings. If the same bias is also observed in the trial of rank, it suggests that Mandarin speakers conceptualize various domains vertically, possibly regardless of the prevalence of vertical metaphors in Mandarin language. The third set, nominal category, was designed to see if Mandarin or English speakers have an overall preference on vertical/horizontal group representations when no spatial metaphors exist in their native language to rely on to guide their responses. If they do not have a preference, then we should expect to see a tendency of random responses to this set. The very last set, spatial, served as a control condition to see if the participants were following the instructions correctly. Questions in this set included marking left/right and up/down. One participant who failed to follow the instructions was dropped.

In addition to the main task, participants were given a survey at the end of the study (see Appendix A & B). If the participants were native English speakers, they would receive a survey that asked them about their Mandarin Chinese language exposure. If the participants were native Mandarin speakers, they were given a survey that asked about their English fluency. The main purpose of these two surveys was to ensure that participants were not pre-

influenced by extraneous factors (e.g., A native English speaker turned out to be majoring in Mandarin Chinese/had lived in mainland China for years, or a native Mandarin speaker who stopped speaking Mandarin Chinese since started living in the US/did not speak English at all and could not understand the study's instructions).

Results

Grouping & Coding (see Table 2)

The responses of each participant were coded based on their relative locations in a coordinate: the reference point that was given to participants (for example: "today" in the yesterday-today-tomorrow question) was considered as the origin (0, 0) of the coordinate. The locations of Xs that were marked down by participants were recorded as (X, Y). (See figure 5).

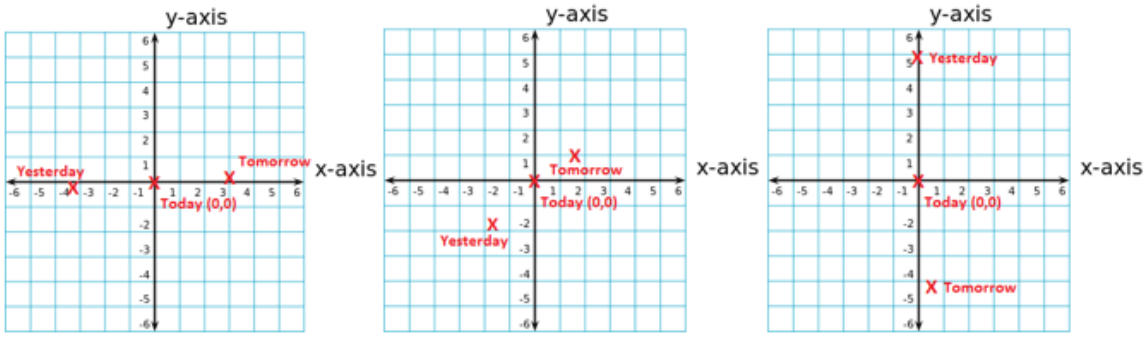


Figure 5. These coordinates demonstrate the computational rules behind the grouping approach. The coordinates on the left, middle, and right illustrate some typical responses from participants. Relatively, the response on the left would be scored as 1H (horizontal), 0V (vertical). The response in the middle would be scored as 0H, 1V. The response on the right would be scored as 1H, 1V. Note that the responses do not have to be perfectly horizontal/vertical. A maximum variation of 20 degrees is permitted.

In order to have a consistent measure on the patterns of arrangements, a formula was

constructed and used to group the data points. This formula uses the Cartesian coordinates to transfer each pair of data points, (x1, y1) and (x2, y2), to an angle that describes the relative location of the given pair of data points. From there, the relationship between the two data points was categorized as horizontal, vertical, or diagonal (see table 3 for a detailed logic flow of this formula). If the degree between two given data points is smaller than 20 degrees (minimum = 0), then this relationship would be described as "horizontal"; if the degree is larger than 70 to 90 degrees (maximum = 90), then it would be described as "vertical". Any degree that fell into the range of 20 to 70 degrees would be described as "diagonal". The reason behind the choice of 20 degrees was due to the consideration of operation errors, which could be made by participants during the marking task or by coders during the measuring process.

Statistical Analysis (see Table 3 & 4)

After the relations of the responses are determined, each relation (horizontal, vertical, and diagonal) is given an H (horizontal) value and a V (vertical) value. For any given horizontal representation, it would be scored 1 on H (horizontalness) and 0 on V (verticalness). Similarly, a vertical representation would be scored as 0 H and 1 V. A diagonal representation implies both horizontalness and verticalness, and it would be scored as 1 H, 1 V. This particular coding approach's purpose is to capture the horizontalness and verticalness of the responses, as well as to increase the analysis similarity to Boroditsky's studies (2001, 2008). In the final analysis, each response was given a value of the difference between proportions of Hs and proportions of Vs for each individual condition. Thus, a positive value indicates that the proportion of horizontal arrangements is higher than the proportion of vertical arrangements. A negative value indicates the opposite.

All the responses collected from English and Mandarin speakers are displayed in four

different scatter plots (see Table 4). The control condition was not included in the factorial analysis, because the purpose of this condition was only to ensure that participants were able to understand and follow the instructions properly. A mixed factorial analysis of variance (2x3 ANOVA) indicated that there was no significant effect of language (English and Mandarin), F (1, 44) = 1.508, $p > .10$. Also, there was no significant main effect of trial type, F (2, 88) = 2.198, $p > .10$. If we ignore other variables and look at the three different trial types (time, rank, and nominal category), participants' responses were the same across all trials. Furthermore, there was no significant interaction between the trial types and the native languages (English or Mandarin Chinese) of the participants, F (2, 88) = 1.893, $p > .10$.

However, a regression test demonstrated that native English and Mandarin speakers conceptualize time differently. As predicted, Mandarin speakers (M = -.23, SD = .4529) were more likely to lay out time vertically than English speakers (M = -.50, SD = .6217) with a two-tailed marginal significance, $R^2 = .061$, F (1, 44) = 2.847, $p < .10$. This trend was not found in the condition of rank. Numerically, the recruited Mandarin speakers (M = -.45, SD = .5778) and English speakers (M = -.68, SD = .5426) did not exhibit the same trend on their spatial representations of rank, $R^2 = .041$, F (1, 44) = 1.872, $p = .178$. Additionally, no significant difference was found between Mandarin speakers (M = -.43, SD = .5411) and English speakers (M = -.31, SD = .5675)'s mental representation of nominal category, $R^2 = .012$, F (1, 44) = .530, $p > .10$. These findings indicated that there is a trend that Mandarin speakers are more likely to conceptualize time vertically than English speakers, and this trend does not extend to the examined domains, including rank and nominal category.

SUPPLEMENTARY TESTS

In additional to the main task, a regression test was done to test if there is a positive correlation between Mandarin speakers' age of acquisition (M = 10.05, SD = 2.61) and the

frequency of using vertical temporal representation. The results of the regression test indicated that the ages of when Mandarin speakers started learning English is not a good predictor of how frequent they conceptualize time vertically, $R^2 = 0.058$, F (1, 20) = .929, p =.347 (see Figure 6).

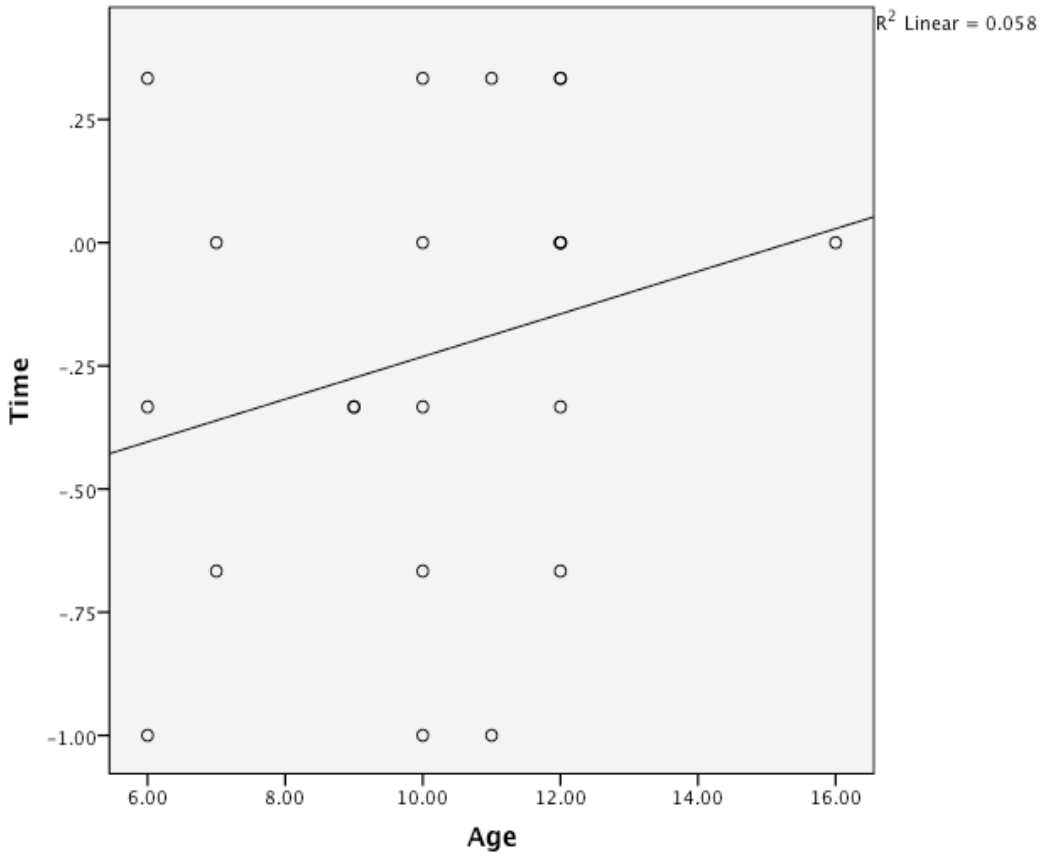


Figure 6. This is a regression line that describes the relationship between Mandarin speakers' self-reported age of English acquisition and the spatial arrangements used to represent time.

During the task, Mandarin participants were also asked to rate their English reading, writing, listening, and speaking ability in English on a scale of 1-9. These 4 items were designed to measure each Mandarin participant's English proficiency. This scale was found to be highly reliable (α = .998). However, it was found that Mandarin participants' English proficiency (M = 6.34, SD = 1.31) was not a good predictor for how frequent they

conceptualize time either, R^2 = .025, F (1, 20) = .504, p =. 486 (see Figure 7).

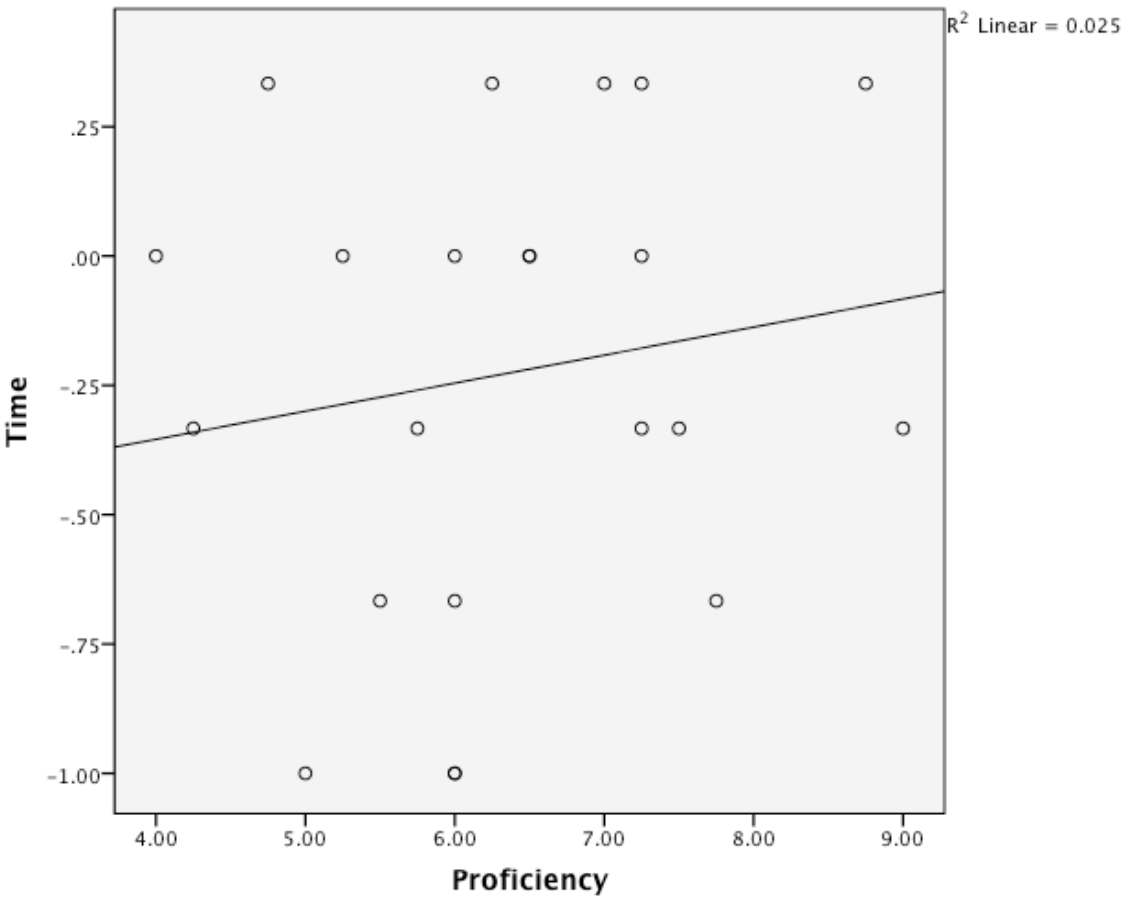


Figure 7. This is a regression line that describes the relationship between Mandarin speakers' self-reported English proficiency and the spatial arrangements used to describe time.

Discussion

The results from this investigation are consistent the hypothesis and the original claim of Boroditsky that Mandarin speakers are more likely to think about time vertically than English speakers do. When asked to mark down the location of temporal events, there was a trend that Mandarin participants were more likely to mark them down with vertical representations, comparing to English speakers, who were more likely to use horizontal representations. Additionally, based on the results obtained, this particular trend observed

among Mandarin speakers does not persist to other domains such as rank or nominal category in a statistical sense. Therefore, in terms of general mapping preference, Mandarin speakers do not have a different conceptualization habit on general concepts than English speakers do. In other words, the difference that was observed was highly likely due to the difference in linguistic cues that exist in English and Mandarin Chinese. This finding supports Boroditsky's results and further provides evidence to support the hypothesis of linguistic relativism. However, the effect (marginal significance) was not as large as the effect found by Boroditsky (2008).

One of the particular strengthens that was established in this study, is that the investigation was done under a full understanding of Mandarin Chinese. Chen (2007) pointed out that researchers can reach erroneous conclusions when they examine a crosslinguistic issue but do not have competent knowledge about the languages they examine. This barrier was overcome in this particular study, as the researcher is a native Mandarin speaker who also speaks English as the second language. Additionally, Mandarin participants were always able to ask for clarifications if they could not understand the English instructions during the task. Having a native Mandarin speaker also minimizes the possible errors that could be made due to a misunderstanding of language.

However, the result obtained from the regression test was not consistent with the result in Boroditsky (2001). In my study, the age of English acquisition of Mandarin participants was not correlated with the likelihood of using vertical metaphors to think about time. In other words, Mandarin speakers' vertical bias neither became stronger nor weaker due to how early they acquired English. However, this could also be due to how the questionnaire was worded. Although all of the Mandarin participants indicated the specific age of when they started learning English, there is a difference between language learning and language acquisition. The participants could be learning English at a different level of

intensity, and some of them may have acquired English younger than others. This specific characteristic was not reflected in the questionnaire, it is possible that this particular questionnaire did not fully capture the factor.

The other notable factor was the language environment. Due to the limitation of location, the test was done at CU-Boulder campus, which is an English speaking environment. The frequency of speaking Mandarin is limited, as most Mandarin participants indicated that they were speaking Mandarin around only 50% of the time. In addition, the task was done in English. Therefore, the effect of Mandarin Chinese might be limited due to both of the constraints. It would be interesting to see whether the effect would become larger if the same task could be tested in Mandarin Chinese, and on monolingual Mandarin speakers in their local language environment (for instance, Taiwan or China). Additionally, by conducting future studies in the local environment, the issue of limited sample size could be easily solved.

Another constraint of this study was discussed by Lucy in his article (1997): studies that adopt the domain-centered approach often have difficulty establishing the significance of expected effects, because the approach can sometimes overlook how time is habitually structured expressed in Mandarin Chinese. In the study, participants' mental representation of time was forced to be expressed on a given 2-dimensional space. However, the actual mental representation of time may be, but not limited to 2-dimensional space. Thus, the artificial experimental approach could actually overlook, or even alter, how time is normally conceptualized in participants' minds.

The results for Mandarin participants' ratings of their English proficiency could also contain some problems. Although the calculated Cronbach's α indicated high internal consistency and reliability within the scale, the self-reported scores were relatively subjective, and may not reflect the actual, objective English proficiency of the Mandarin participants. A more reliable English proficiency scale, such as the scores participants obtained from TOEFL

or IELTS (the international standardized tests for English language proficiency for non-native English language speakers), could be used in future studies.

Demand effect is another considerable factor that might have occured in this study. When asked about how they think of the task, some participants reported they believed this task could predict some aspects of their personalities or creativity. Although the real purpose of this study was well hidden during the experiment, one English speaker specifically asked the researcher if this experiment was researching about the relationship between Asian writing direction and the mental representations of objects and events. It seemed that a few participants were guessing the purpose behind the experiment, and the guessing behavior might have altered their responses.

For future studies, a computer-based program could be developed and replace the actual paper-based datasheets. This way allows the operational error (ex. The observational errors that could be made by person during the measurement) to be minimized, and is less costly, comparing to the actual time invested on measurement. In addition, it is hoped that the demand effect will be controlled through computer-based tasks instead of human interactions. Also, it would be interesting to test bilinguals who speak English and Mandarin Chinese as their first languages. The results may vary based on if English or Mandarin Chinese is more dominant in these bilinguals.

Regarding the effect found in this study, the coding method that was deployed could be improved by treating the responses as a continuous variable rather than a nominal variable. Specifically, instead of grouping the responses into categories (horizontal and vertical), the degree of displacement or other numerous values of individual responses could be measured and analyzed. Although the direction of how the trend would change still remains unknown, this method is worth trying and will be the next step of data analysis.

In addition, native English speakers should be further screened by their exposure to

East Asian languages (such as Japanese) instead of Mandarin Chinese alone. This is due to the similarity that exists across East Asian languages. For instance, the traditional writing of Japanese is also arranged from top to bottom, and right to left. The exposure to Japanese writing could also have some certain effects on English speakers mind and make them arrange time more vertically than other monolingual English speakers who have no exposure to Japanese. A more careful screening method will be deployed in the future.

Acknowledgement

This thesis was accomplished with the supports of my thesis committees. Without their encouragements, academic advices, and personal inspirations, I would not be able to complete this honors thesis on my own.

Dr. Eliana Colunga is the direct adviser of this thesis. I have benefited a lot from the high level of intelligence and enthusiasm of hers. I also have received enormous help and guidance from Dr. Andrea Feldman, who has been my writing instructor since four years ago. Dr. Martin Goldman is my inspirational model and an excellent professor who centered on plasmas research. Also, Dr. Richard Olson, has taught me the critical thinking as well as the responsibilities I should take as a researcher.

I also would like give my heartfelt thanks to Dr. John Lynch, Dr. Donald Lichtenstein, Lauren Becker and Philip Pendergast. I feel truly thankful for their encouragements and academic supports. I would like to thank Qiaoxuan Zhang, who had always helped me with various issues I encountered during this research project.

This thesis was financially supported by an individual research grant from the Undergraduate Research Opportunity Programs at CU-Boulder. Lastly, I would like to give thanks to the Mandarin speakers who participated in my study particularly.

References

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

Boroditsky, L. (2008). Does English and Mandarin speakers think differently about time? *Proceedings of the 30th annual conference of the cognitive science society*, 64-70.

Chen, JY. (2007). Do Chinese and English speakers think about time differently? Failure of replicating Boroditsky (2001). *Cognition*. 104, 427-436.

Chomsky, N. (2006). *Language and mind*. Cambridge University Press. viii.

Fuhrman et al. (2011). How linguistic and cultural forces shape conceptions of time: English and Mandarin time in 3D. *Cognitive Science*. 35, 1305-1328

Fuhrman, O., & Boroditsky, L. (2010). Cross-cultural differences in mental representations of time: evidence from an implicit nonlinguistic task. *Cognitive Science*. Doi: 10:1111/j.1551-6709.2010.011-5.x.

January, D., & Kako, E. (2007). Re-evaluating evidence for linguistic relativity: Reply to Boroditsky (2001). *Cognition*. 104, 417-426.

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

Lucy, J, A. (1997). Linguistic relativity. *Annual Review of Anthropology*. 26, 291-312.

Sapir, E. (1951 [1929]), The status of linguistics as a science. *In selected writings*. David Mandelbaum, ed. Berkeley: University of California Press. Orig. pub. In Language 5:207-214.

Tan, A (1990). The language of discretion. *The opposite of fate: A book of musings*, 2003.

Whorf, B. L. (1956 [1940]). Science and linguistics. *In language, thought, and reality*. John B. Carroll, ed. Cambridge, Mass,: M.I.T. Orig. pub. in Technology Review 42:229-231, 247-248.

Table 1

Words that were used in the main task

Domains	Words Used
Time	1) Yesterday, Today, Tomorrow 2) Breakfast, Lunch, Dinner 3) August, September, October
Rank	1) First, Second, Third 2) Small, Medium, Large 3) A, B, C
Nominal Category	1) Crayon, Pencil, Brush 2) Apple, Banana, Grapes
Space (control)	1) Top, Middle, Bottom 2) Left, Middle, Right *This condition used the same reference (Middle) twice

Table 1. A list of words used in the task. There were 3 trials under the conditions of time and rank, 2 trials under the conditions of nominal category and space. The words underlined in the middle were the words that served as reference (i.e., "if the X represents lunch, where would you put breakfast/dinner.").

Table 2.

A Flow Chart that describes the basic logic behind the formula

if $\Delta x=0$,

    if $|\Delta y|<0.5$,

        output = "D",

otherwise,

        output = "V"

otherwise,

    if $|\arctan(|\Delta y/\Delta x|)|<$ or = 20 degrees,

output = "H”,

otherwise,

if |arctan(|Δy/Δx|)|> or = 70 degrees,

output = "V”,

otherwise,

output = "D".

Expression in Excel:

IF(BO4-BR4=0, IF(ABS(BP4-BS4)<0.5,"D","V”), IF(ABS(ATAN(ABS(BS4-BP4)/ABS(BR4-BO4)))<(PI()/9), "H”, IF(ABS(ATAN(ABS(BS4-BP4)/ABS(BR4-BO4)))>(7*PI()/18), "V”, "D")))

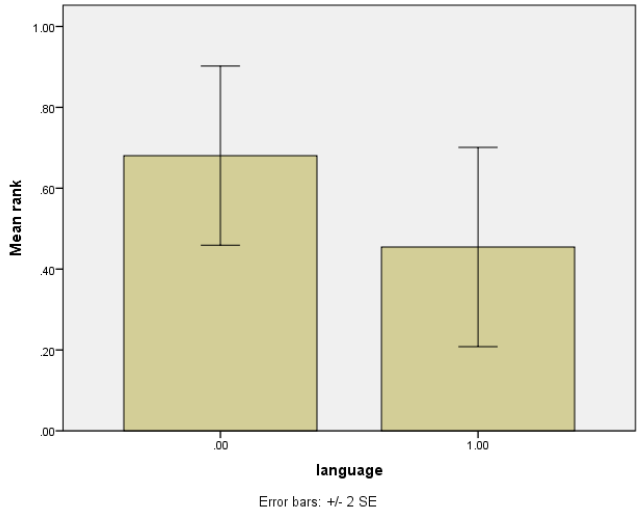
Table 3.

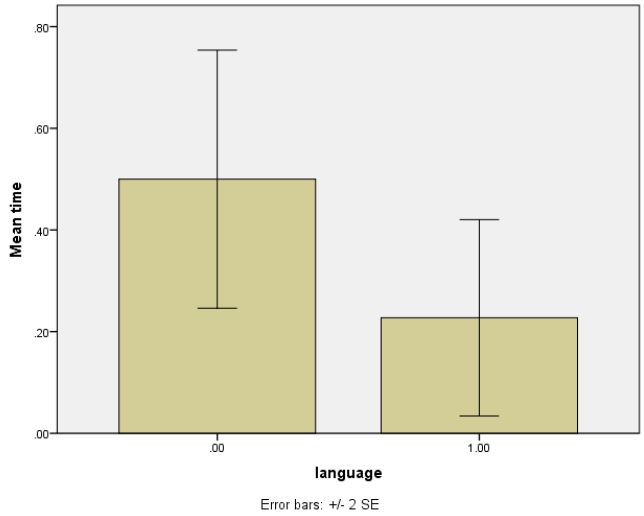
Mean and Standard Error of the Mean for the difference between the Proportion of Verticals and the Proportion of Horizontals (Maximum Value = 1, Minimum Value = -1) as an Approach to analyze English speakers and Mandarin speakers' Mental Representations of Time.

	Mandarin	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
Time	.00	24	-.5000	.62167	.12690
	1.00	22	-.2273	.45293	.09657
Rank	.00	24	-.6806	.54266	.11077
	1.00	22	-.4545	.57777	.12318
Category	.00	24	-.3125	.56746	.11583
	1.00	22	-.4318	.54106	.11535

The charts below are demonstrations of the differences between the means of English and Mandarin speakers on their response of time, rank, nominal category, and control. The y-

axis indicates the difference between the proportion of Hs and Vs. The x-axis indicates the types of language (English = .00, Mandarin = 1.00)





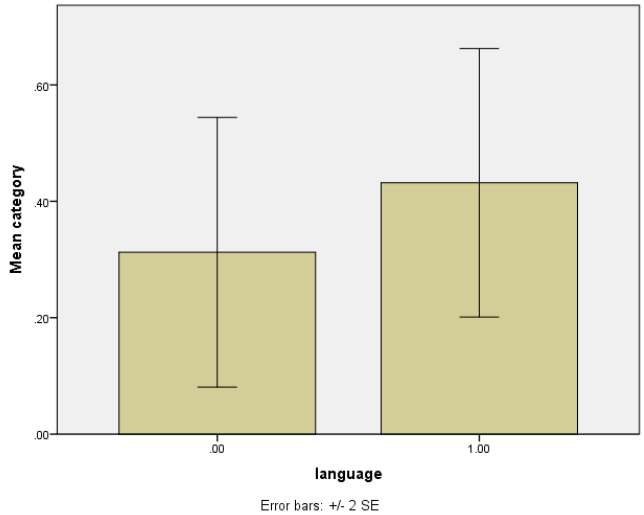
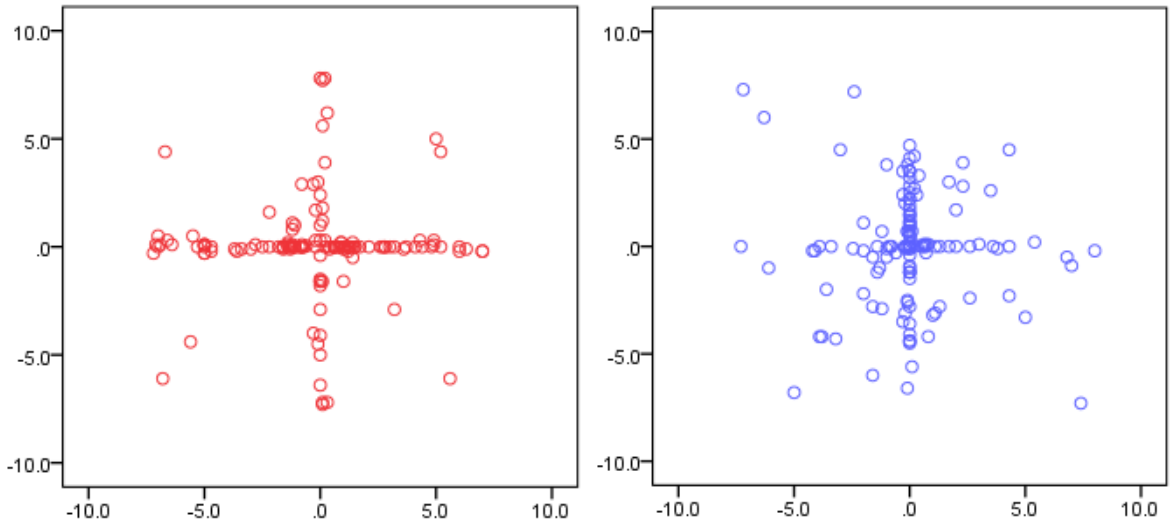


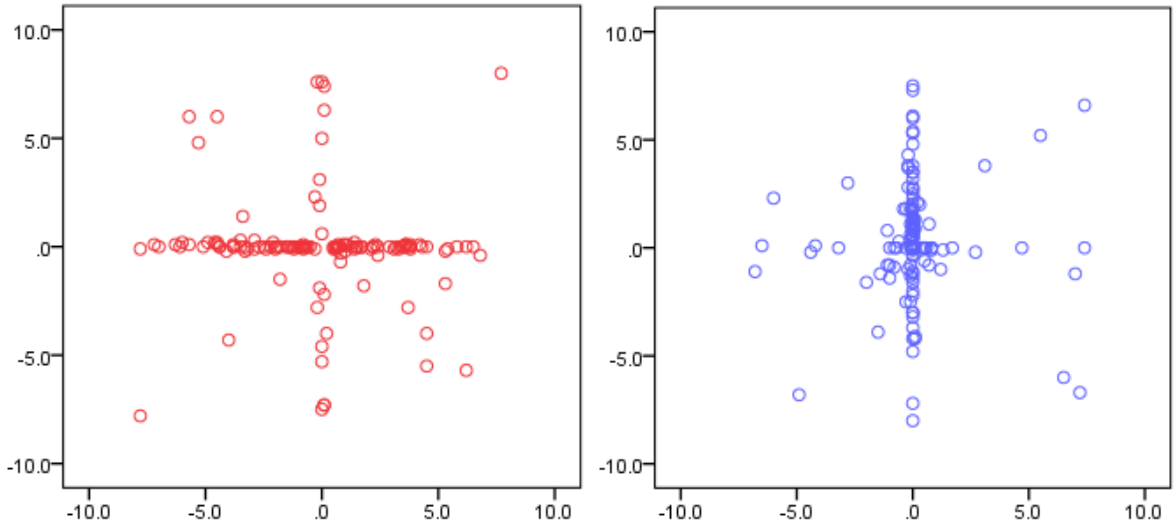
Table 4.

Distributions of responses collected from English speakers (Red, Left) vs. Mandarin Speakers (Blue, Right) in four conditions (Time, Rank, Nominal Category, Control)

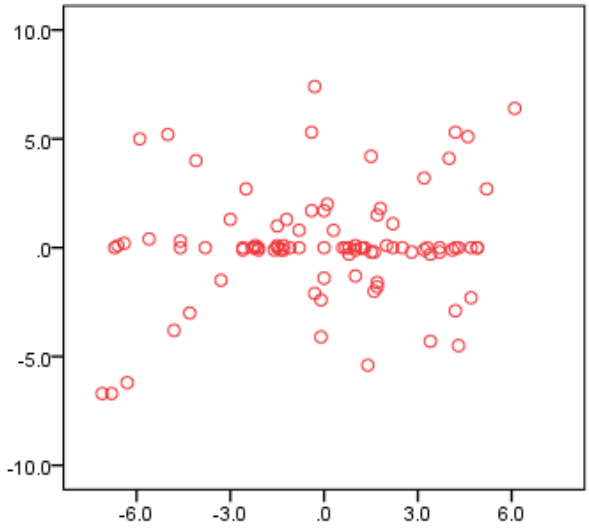
Time

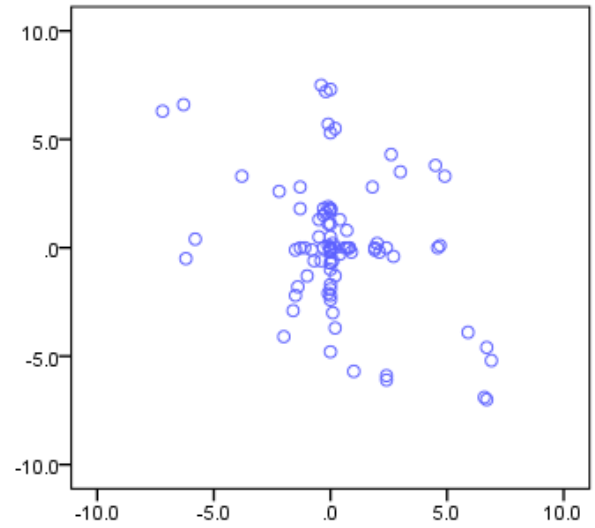


Rank

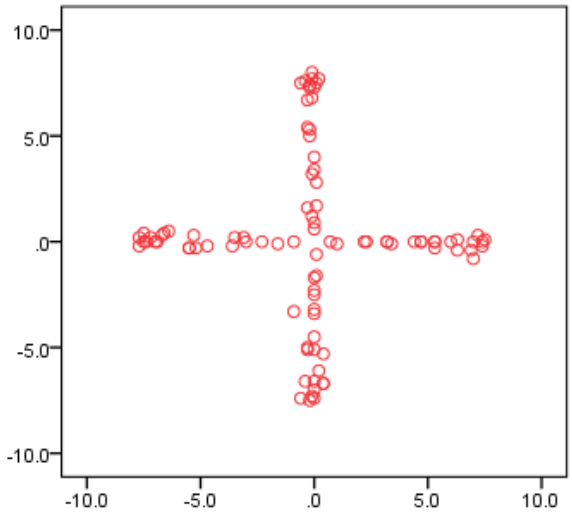


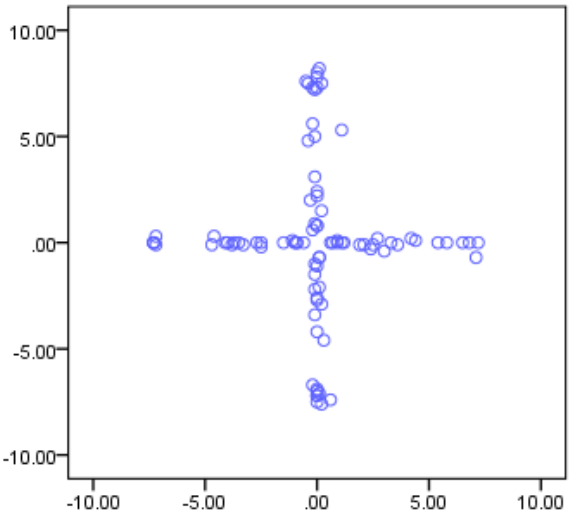
Nominal Category


10.0
5.0
.0
-5.0
-10.0
-6.0
-3.0
.0
3.0
6.0


10.0
5.0
.0
-5.0
-10.0
-10.0
-5.0
.0
5.0
10.0

Control


10.0
5.0
.0
-5.0
-10.0
-10.0
-5.0
.0
5.0
10.0


10.00
5.00
.00
-5.00
-10.00
-10.00
-5.00
.00
5.00
10.00

Appendix A: Survey for native English speakers

SOCIAL INTERACTION AND MANDARIN LANGUAGE EXPOSURE

Do you interact regularly with speakers of Mandarin Chinese?	Yes		No	

If **yes**, please provide the following information:	
Speaker or speakers (e.g., parent, sibling, nanny, sitter, neighbor, teacher, relative, friend)	
Frequency of exposure (e.g., # hrs per day, # days per week)	
Proficiency of the speaker or speakers (e.g., native speaker, excellent, good, fair, poor)	

Have you attended a school that provides instruction in Mandarin Chinese?	Yes		No	

If **yes**, please provide the following information:	
Speaker or speakers (e.g., parent, sibling, nanny, sitter, neighbor, teacher, relative, friend)	
Frequency of exposure (e.g., # hrs per day, # days per week)	
Proficiency of the speaker or speakers (e.g., native speaker, excellent, good, fair, poor)	

Have you been exposed to Mandarin Chinese through any of the following? (Check all that apply)			
books		If **yes**, please rate the level of exposure on a scale of 1-9, with 1 being extremely little, and 9 being extremely high:	
audiotapes/videotapes			
television			
computer programs			
cultural programs/camps			
travel/trips			
other (please describe):			

Please indicate your current major(s) ***<u>If you are currently in Open Option, please indicate the major(s) you want to claim</u>**	

Appendix B: Survey for native Mandarin speakers

COMPETENCE IN ENGLISH

At the time of this report, what is the frequency of your speaking in Mandarin Chinese as compared to the frequency of speaking in English: (Check one)	
	Doesn't talk in either language
	Never uses English
	Uses English less than 25% of the time
	Uses English about 50% of the time
	Uses English 75 to 100% of the time

By what age did you start learning English for the first time? (Check one)											
Age 2-6		Age 7-12		Age 13-15		Age 16-18		Age 19-22		Over Age 23	

On a scale of 1-9, with 1 being not good at all, and 9 being extremely good, how good is your English ability in reading, writing, listening, and speaking? (Please rate)							
Reading		Writing		Listening		Speaking	

Please briefly describe your experience with learning English. Be sure to indicate the age of the learning activities (eg., "I started learning English in middle school (12-year-old), and I had extra English classes on Sundays in the 8th grade (13-year-old)…").