Bilingualism: Language and Cognition 15 (2), 2012, 413–421 © Cambridge University Press 2011 doi:10.1017/S1366728911000320

Linguistically directed attention to the temporal aspect of action events in monolingual English speakers and Chinese–English bilingual speakers with varying English proficiency*

JENN-YEU CHEN
National Cheng Kung University, Taiwan
JUI-JU SU
National Cheng Kung University, Taiwan
CHAO-YANG LEE
Ohio University, USA
PADRAIG G. O'SEAGHDHA
Lehigh University, USA

(Received: September 7, 2010; final revision received: May 10, 2011; accepted: May 13, 2011; first published online October 3, 2011)

Chinese and English speakers seem to hold different conceptions of time which may be related to the different codings of time in the two languages. Employing a sentence–picture matching task, we have investigated this linguistic relativity in Chinese–English bilinguals varying in English proficiency and found that those with high proficiency performed differently from those with low proficiency. Additional monolingual English data, reported here, showed further that high-proficiency bilinguals performed similarly to the English monolinguals, suggesting that Chinese speakers' sensitivity to the time of an action event might be modifiable according to the extent of their experience with a tensed language.

Keywords: Chinese–English bilingualism, tense, aspect, action, linguistic relativity, extended present

Consider the following Chinese sentence:

(1) Zhong1-wu3 chi1 she2-mo?
noon eat what

Without an additional context, this sentence is ambiguous with respect to the time of the referenced eating event. It can refer to a completed/past action (i.e., "What did you eat for lunch?"), or to an imminent/future action (i.e., "What are you going to each for lunch?"). If English is the language of conversation, there is no ambiguity of this sort, as tense or modal adverbial is always specified in the sentence to convey explicitly the time of action.

English is a language which marks the temporal situation of an event explicitly in a sentence using tense and aspect markers (Comrie, 1985; Smith, 2006). Chinese is a language in which the time of an event is often not marked explicitly and must be inferred from the context (Li & Thompson, 1981; Lin, 2006; Smith & Erbaugh, 2005; Wu, 2009). Aside from the lack of explicit verbal markings for tense, the few aspect markers (such as *guo4*, *zhe5*, *le5* and *zai4*) must be interpreted in conjunction with the situation type of the verb (Chang, 1998; Lin, 2002, 2003) or with other factors such as viewpoint, verbal semantics, temporal adverbials, the definite/indefinite distinction, quantifier raising, informational status, pragmatics, people's knowledge of the world, etc. (Lin, 2006), in order for the correct temporal meaning to be inferred. Given that language is the primary means for expressing thought, linguistic differences such as the ones illustrated here certainly influence the mapping from thought to linguistic expression, but may also lead to subtle differences in thought processes themselves between speakers of different languages (Gumperz & Levinson, 1996; Levinson, 2003; Lucy, 1996; Slobin, 2003).

A growing number of studies have been conducted in recent years to explore the relation between cognitive diversity and linguistic diversity. These investigations have covered cognition of a wide range of attributes, such as, among others, color (Roberson, Davidoff, Davies & Shapiro, 2005), space (Levinson, 2003), time

* The work reported here was supported by the NSC-96-2752-H-006-001-PAE grant of the National Science Council of Taiwan, the Republic of China, as well as by an internal grant (D16-B0045) of National Cheng Kung University. We thank Yuh-Fang Lee, Yu Zhang, and Jennifer Harms for assistance in administering the experiment, and Danny R. Moates for assistance in recruiting the English-speaking participants at Ohio University. Special thanks go to the anonymous reviewers for their critical but helpful comments on the early versions of the paper. Jui-Ju Su is now at the Basque Center on Cognition, Brain, and Language, Donostia-San Sebastián, Basque Country (Spain).

Address for correspondence:
Jenn-Yeu Chen, Institute of Cognitive Science, National Cheng Kung University, 1 University Road, Tainan, Taiwan 701, The Republic of China
psyjyc@mail.ncku.edu.tw

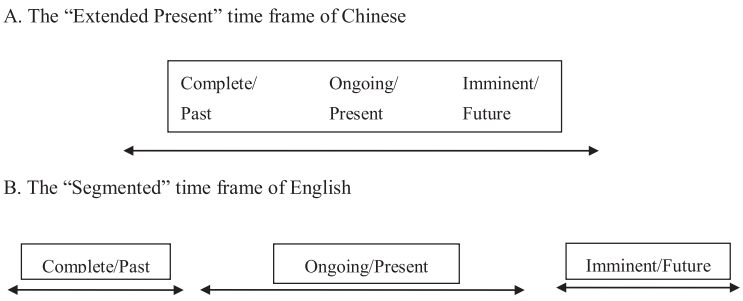

Figure 1. Chinese and English speakers' conceptions of time.

(Alloway & Corley, 2004; Boroditsky, 2001), number (Gordon, 2004; Pica, Lema, Izard & Dehaene, 2004), and object (Davidoff, Fonteneau & Goldstein, 2008). The investigations are important not only because they address the highly-appealing (though controversial) Whorfian hypothesis (Hunt & Agnoli, 1991; Whorf, 1956), but also because many researchers believe that only a full understanding and appreciation of diversity can lead to a reasonable and precise theory of universality. In this tradition, we examine here how conceptions of time may vary in speakers of languages that encode temporal information in different ways. In this paper we are specifically interested in the further question of how bilingualism might modulate the mappings between language and thought. Examination of bilingual populations not only illuminates how bilinguals adjust to particular cross-linguistic challenges, but may also clarify the nature of language–thought transactions.

In a previous study (Chen, Su & O'Seaghdha, 2011), we presented Chinese and English speakers with a set of pictures, each depicting a person performing an action (e.g., throwing a frisbee). Three temporal stages were depicted, e.g., the person is about to throw a frisbee (future), is throwing a frisbee (present), or has thrown a frisbee (past). For a given action event, participants saw only one of the three temporal aspects. Their task was simply to describe each action event with one sentence in their language. Results showed that the English descriptions matched the prescribed time categories more closely than the Chinese descriptions (88% vs. 64%). When the three temporal aspects were examined separately, we found that the Chinese participants diverged most on the past and future events, describing some of them as present ones, using the aspect marker *zai* "now". This tendency towards the present disappeared when the participants were told explicitly ahead of time that each action could usefully be described in terms of one of the three temporal phases. We interpreted the participants' improvised verbal descriptions of action events as reflecting their habitual ways of viewing these events, which correlated interestingly with the different ways of coding tense and aspect in the two languages.

We hypothesize that Chinese speakers develop and maintain a habitual extended-present time window that encompasses the near future and the recent past. Figure 1 illustrates how the Chinese and the English speakers' conceptions of time might differ. In the Extended Present of tense-less Chinese, entire actions are comprised in the present, bringing their initiation and completion closer to the center of actions but paradoxically extending their subjective durations. In the Segmented Time of tensed English, actions tend to be viewed as comprising early preparation, middle execution, and late completion phases, distancing imminent and completed actions from their centers and by the same token shrinking the center itself (Chen et al., 2011). Consistent with the hypothesis, we observed that Chinese speakers projected a narrower time window (distance between average imminent and average completed events) than English speakers when they were asked to mark the pictured actions on a timeline. At the same time, Chinese speakers projected a wider time window than English speakers when they marked the beginning and end points of ongoing actions.

In a subsequent study (Chen & Su, 2010), we extended this work by comparing the performance of high and low-proficiency Chinese–English bilinguals on a sentence–picture matching task (participants were to choose which of two pictures better matches a sentence that has just been read). Each target sentence (in Chinese) described a particular phase of an action event. The two pictures differed only with respect to the temporal phase. Filler sentences and pictures that described objects or a person's profession were included. The results showed that the high- and low-proficiency Chinese–English bilinguals performed similarly on the nontargets (fillers), but the high-proficiency participants matched the temporal phases of the target pictures more closely. In addition, the difference between the high- and low-proficiency participants came mainly from their responses to the past and future items. These findings suggest that extensive exposure to a language that systematically marks the time of events grammatically may have served to modify the conceptual system of the proficient Chinese–English bilinguals such that the initially blurred boundaries between the present and the past phases and between the

https://doi.org/10.1017/S1366728911000320 Published online by Cambridge University Press

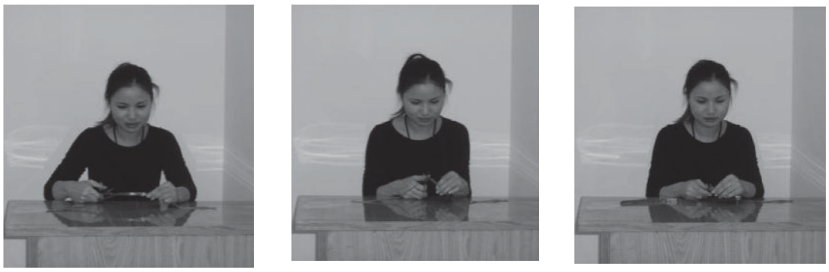

Figure 2. Pictures of three temporal phases of the action of cutting a rope (from left to right: imminent/future, ongoing/present, completed/past).

present and the future became distinctive enough to allow for a more differentiated processing of action events (Chen & Su, 2010).

In this short report, we provide additional data to supplement the findings of Chen and Su (2010). In Chen and Su (2010), participants were Chinese–English bilinguals. An additional condition comprising monolingual English speakers is reported here. Monolingual English speakers' performance on the same task provides an essential benchmark for evaluating the extent of influence of English on the Chinese–English bilinguals' conception of time. The two bilingual samples were also expanded to include 32 more participants. In the following, we re-present the method of Chen and Su (2010). In the Results section, we include the expanded bilingual data as well as monolingual English data for comparison. If a tensed language like English can modify Chinese speakers' conception of time, we should observe that the Chinese–English bilinguals with high English proficiency perform similarly to the monolingual English speakers, but differently than the low-proficiency Chinese–English bilinguals.

Method

Participants

Thirty-two native Chinese speakers with high English proficiency (scoring at or beyond the 88th percentile on the English subject of the College Entrance Exam, CE-H; 10 were new) and 49 with low English proficiency (scoring at or below the 12th percentile, CE-L; 22 were new) participated in this experiment. Twenty-two of the CE-H group and 27 of the CE-L group were reported in Chen and Su (2010). The high-proficiency participants were freshman and sophomore students recruited from National Cheng Kung University, while the low-proficiency individuals were seniors of a local high school who had taken the College Entrance Exam. The two groups of participants were at most two years apart in age. Most of them began to learn English as a foreign language since the age of eleven or twelve as part of the middle school curriculum. Some had started earlier by taking private lessons, but we did not have the details of when, for how long, and how often. The Chinese participants were paid for participation. In addition, thirty native English-speaking monolinguals (MonoE) from Ohio University (not reported in Chen & Su, 2010) also participated in this experiment as part of a course requirement in Introductory Psychology.

The mean age was around 21 years (age range from about 19 to 40 years) for the CE-H group, 18 years (range from 17 to 18 years) for the CE-L group, and 20 years (range from 18 to 30 years) for the MonoE group. There were 17 (53%), 21 (43%), and 22 (73%) women in the CE-H, CE-L, and MonoE groups, respectively.

Materials

Eighteen action events were chosen as target events (blowing a balloon, erasing a white board, crossing a log, cutting a rope, throwing a frisbee, drinking tea, folding a piece of color paper, sliding down the slide, eating a banana, lighting a candle, peeling an orange, walking on a balance bar, cutting a lettuce, doing a puzzle, walking up the stairs, ripping a piece of paper, kicking a ball, and pouring dark liquid into a glass). One woman performed all the actions. A snapshot was taken of each of the three temporal phases of each action (e.g., about to cut a rope, in the act of cutting a rope, having just cut a rope). Thus, three pictures were taken for each action event, resulting in a total of 54 pictures (see Figure 2 for an example). For each picture, a Chinese sentence was first constructed that described the action in the given temporal aspect. For the English participants, all the Chinese sentences were, then, translated into English and back-translated into Chinese to ensure the comparability of contents. Some of the English translations were reworded, following the advice of a native speaker of English, to make them more natural (e.g., "She has just finished drinking a cup of tea" instead of "She has just drunk a cup of tea").

The 54 targets were divided into three sets of 18. Each set covered all the 18 action events, with each event appearing in only one of the three temporal aspects and each temporal aspect appearing equally often. (See Appendix for a list of all the target sentences.)

https://doi.org/10.1017/S1366728911000320 Published online by Cambridge University Press

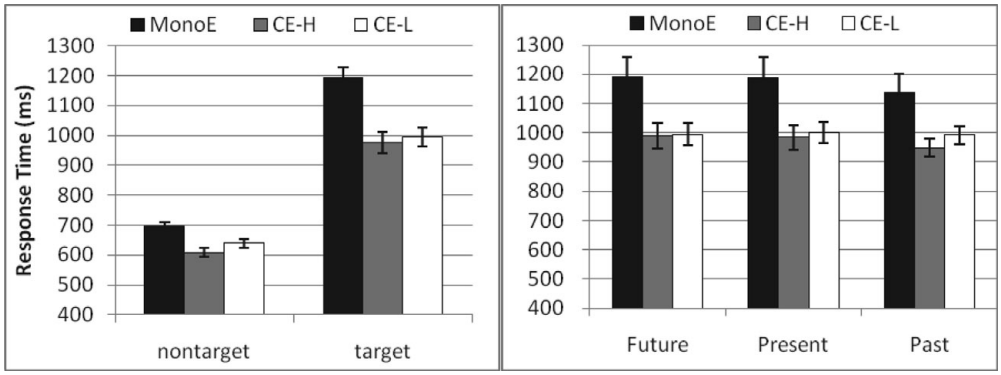

Figure 3. Left panel: Mean response times and standard errors for the target and the non-target pictures by the English monolinguals (MonoE), the Chinese–English bilinguals with high English proficiency (CE-H), and the CE bilinguals with low English proficiency (CE-L). Right panel: Mean response times and standard errors for the target pictures only broken down by group and temporal phase.

Each participant received only one set (18) of the target sentences and pictures. Another 62 sentences and pictures describing objects or people of different occupations or genders were also prepared and served as fillers (e.g., "this is a teacher", "this is a ballpoint pen", etc.). The English translations of these sentences were made in the same way as the target sentences. The MonoE group did the English version and both the CE-H and the CE-L groups did the Chinese version.

Procedure

Each participant saw 80 sentences, including one set of 18 target sentences and 62 filler sentences. Each sentence was shown on the screen one character/word at a time, 200 ms per character for the Chinese sentences and 500 ms per word for the English sentences. Immediately after the last word disappeared, two pictures appeared on the screen. One picture was a target and matched the temporal aspect of the sentence just shown. The other either depicted the same action but in a different temporal phase, if the trial involved a target, or it depicted a different object or occupation, if the trial involved a non-target filler. The participants had to choose the picture that matched the sentence they just read by pressing a designated left or right key. The response keys were counterbalanced across participants.

Participants' response times were measured with millisecond precision from the appearance of a picture to the initiation of a key-press. Normative accuracy (match or mismatch as defined above) of responses was also recorded.

Results

Response time

Figure 3 (left panel) presents the mean correct response times to the target and non-target pictures for the English monolinguals (MonoE), the Chinese–English bilinguals with high English proficiency (CE-H), and the Chinese–English bilinguals with low English proficiency (CE-L). Overall, responses to the target pictures were slower than responses to the non-target pictures: $F1(1,108) = 580.78$, $MSe = 15053$, $p < .0001$; $F2(1,78) = 155.91$, $MSe = 44273$, $p < .0001$. Responses by the MonoE participants were also slower than those by the CE-H and the CE-L participants: $F1(2,108) = 5.33$, $MSe = 78593$, $p = .0062$; $F2(2,156) = 91.04$, $MSe = 3200$, $p < .0001$, and this was especially the case for the target pictures: $F1(2,108) = 6.86$, $MSe = 15053$, $p = .0016$; $F2(2,156) = 27.2$, $MSe = 3200$, $p < .0001$.

The right panel of Figure 3 presents the mean response times for the target pictures only, broken down by time phase of the events. As the figure shows, the MonoE participants responded more slowly than the CE participants, regardless of the event phase. There was a significant main effect of group: $F1(2,108) = 5.77$, $MSe = 196587$, $p = .0042$; $F2(2,34) = 77.93$, $MSe = 7007$, $p < .0001$; but no main effect of phase, nor interaction of group with phase: $F1$s and $F2$s < 1.

Response matching rate

Figure 4 (left panel) presents the mean response matching rates to the target and the non-target pictures for the three groups of participants. Overall, there were more matches in the responses to the non-target pictures than to the target pictures: $F1(1,108) = 237.57$, $MSe = 0.0049$, $p < .0001$; $F2(1,78) = 112.44$, $MSe = 0.0070$, $p < .0001$. The three groups also responded differently: $F1(2,108) = 9.6$, $MSe = 0.0075$, $p = .0001$; $F2(2,156) = 16.8$, $MSe = 0.0018$, $p < .0001$. The CE-H participants performed differently (with higher matching rates) than the CE-L participants: $F1(1,108) = 7.22$, $MSe = 0.0075$, $p = .0084$; $F2(1,78) = 12.81$, $MSe = 0.0024$, $p = .0006$; but

https://doi.org/10.1017/S1366728911000320 Published online by Cambridge University Press

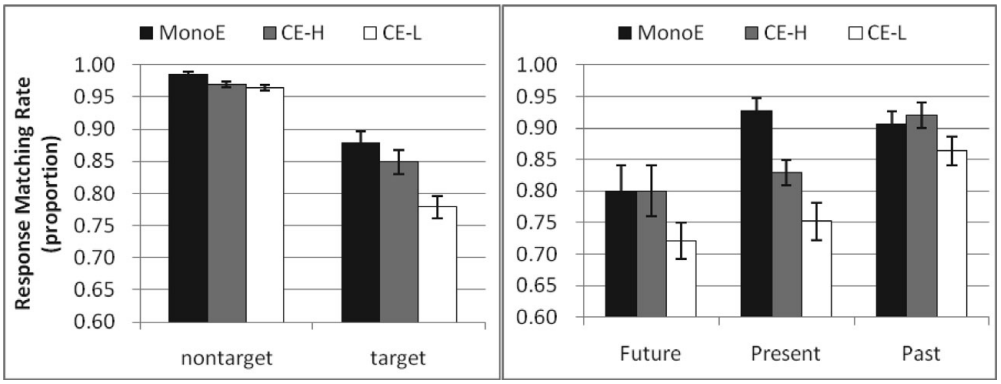

Figure 4. Left panel: Mean response matching rates and standard errors for the target and the non-target pictures by the English monolinguals (MonoE), the Chinese–English bilinguals with high English proficiency (CE-H), and the CE bilinguals with low English proficiency (CE-L). Right panel: Mean response matching rates and standard errors for the target pictures only broken down by group and temporal phase.

performed less differently than the MonoE participants: $F1(1,108) = 2.12$, $MSe = 0.0075$, $p = .1482$; $F2(1,78) = 9.11$, $MSe = 0.0033$, $p = .0034$. The patterns of the group difference were different for the target pictures and the non-target pictures: $F1(2,108) = 7.11$, $MSe = .0049$, $p = .0013$; $F2(2,156) = 13.24$, $MSe = .0018$, $p < .0001$. For the nontarget pictures, the CE-H participants performed similarly as the CE-L participants: $F1(1,108) = 0.69$, $MSe = 0.0007$, $p = .4082$; $F2(1,61) = 1.03$, $MSe = 0.0015$, $p = .3146$; but performed differently than the MonoE participants: $F1(1,108) = 6.21$, $MSe = 0.0007$, $p = .0142$; $F2(1,61) = 8.37$, $MSe = 0.0021$, $p = .0053$. Critically, for the target pictures, The CE-H participants performed differently (with higher matching rates) than the CE-L participants: $F1(1,108) = 8.06$, $MSe = 0.0118$, $p = .0054$; $F2(1,17) = 15.51$, $MSe = 0.0057$, $p = .0011$; but performed similarly as the MonoE participants: $F1(1,108) = 1.09$, $MSe = 0.0118$, $p = .2985$; $F2(1,17) = 1.91$, $MSe = 0.0078$, $p = .1845$.

The right panel of Figure 4 presents the mean response matching rates for the target pictures only, broken down by the time phase of events. As the figure shows, the three groups of participants performed differently: $F1(2,108) = 8.74$, $MSe = .0353$, $p = .0003$; $F2(2,34) = 9.96$, $MSe = .01333$, $p = .0004$. They responded to the three event times differently, too, $F1(2, 216) = 18.10$, $MSe = .0251$, $p < .0001$; $F2(2,34) = 3.31$, $MSe = .0627$, $p = .0485$, suggesting that it was easier to identify ongoing and especially completed actions as intended than incipient ones. Although it looks like high-proficiency Chinese speakers match the present phase events less than the monolingual English speakers, the relevant interaction was nonsignificant, indicating overall similar patterns of group differences across the event phases: $F1(2,216) = 1.97$, $MSe = .0251$, $p = .1002$; $F2(4,68) = 1.74$, $MSe = .0141$, $p = .1519$. The group differences are characterized primarily by the lower matching scores of the CE-L participants relative to the CE-H participants: $F1(1,108) = 8.06$, $MSe = 0.0353$, $p = .0054$; $F2(1,17) = 16.56$, $MSe = 0.0459$, $p = .0008$ and by the similar performance of the CE-H participants and the MonoE participants: $F1(1,108) = 1.09$, $MSe = 0.0353$, $p = .2985$; $F2(1,17) = 1.76$, $MSe = 0.0711$, $p = .2025$.

Discussion and conclusion

The experience of time is fundamental to human cognition and action. Therefore, all languages have ways of expressing it, although they use different linguistic (lexical and grammatical) devices to do so (Klein & Li, 2009; Li & Shirai, 2000). We investigated whether conceptions of time differ in speakers of languages that encode time differently (Chinese and English), and whether bilingualism may further modulate the language-specific conceptions of time. Our investigations show that when asked to choose one of two pictures that matched the sentence previously read, Chinese–English bilinguals with high English proficiency performed similarly to English monolinguals, and both groups produced higher matches than the Chinese–English bilinguals with low English proficiency. This pattern of similarity and difference emerged only for the target pictures, but not for the non-target filler pictures. The critical information in the target pictures was the temporal phase of the pictured actions.

Is the pattern of results the consequence of linguistic processing or nonlinguistic processing? We think nonlinguistic processing is more likely. What the participants did on the task was presumably to first construct a semantic representation of the sentence, and later determine which of two pictures matched this semantic representation. To produce a matching response, the participants needed to direct their attention to the (relevant) temporal phases of the pictured actions and to

https://doi.org/10.1017/S1366728911000320 Published online by Cambridge University Press

identify the relevant distinction between the two. In other words, although the task involved linguistic processing of sentences, responses were based on processing of pictures, when the sentences just read were no longer present. Thus, the pattern of results more likely reflects the consequence of nonlinguistic processing.

One could, of course, argue that the participants might have attempted to construct a sentence for each picture and to match them with the one they just read. In that case, the pattern of results could be said to reflect the consequence of linguistic processing, and not nonlinguistic processing. That the participants produced sentences for the pictures appears unlikely but is logically possible. A strategy they might have adopted is to name just the relevant aspect of the pictured actions, e. g., to produce nouns for the objects (nontargets) and verbs, with proper temporal markings, for the actions (targets). The nouns or the verbs were then compared with those they remembered from the sentences before responses were determined.

Another way in which our findings might reflect linguistic processing rather than conceptual change is that the fixed and shorter presentation time of the sentences might not be sufficient for some of the participants (i.e., low-English-proficiency bilinguals) to construct the appropriate semantic/conceptual representations of the sentences, leading to an increase of mismatches in later responses. However, our high-English-proficiency bilinguals, receiving the same sentences with the same presentation time as the low-English-proficiency bilinguals, performed similarly to the English monolinguals. Therefore, the different sentence presentation times were unlikely the cause of the different matching responses between the English monolinguals and the low-proficiency Chinese–English bilinguals. Note that with this interpretation, the source of the processing difference is located in the semantic/conceptual system, rather than the linguistic system (i.e., the form).

Perhaps a more satisfying interpretation of the current results is to consider the roles of both linguistic and nonlinguistic processing. For example, the sentences might have acted as linguistic cues to direct the participants' attention to the temporal phases of the pictured actions. The linguistic difference between the English and the Chinese ways of coding the temporal phases of actions led the English and the Chinese participants to focus on different temporal characteristics of action events. For the Chinese participants, the focus was on the continuity of the different temporal phases, whereas for the English participants, the focus was on the discreteness of the different temporal phases. Focusing on the continuity of time led to less differentiation of temporal phases and a result of increased mismatches in the Chinese speakers' responses. Extensive experience with English provided the high-English-proficiency bilinguals an alternative parsing of the discrete characteristics of time leading to increased matching responses that patterned approximately like those of the English speakers.

Note that with this last interpretation, it is assumed that Chinese–English bilinguals develop and maintain two conceptions of time, the continuous one and the discrete one, and they may be able to move flexibly between them. It then implies that monolingual Chinese and English speakers develop and maintain different conceptions of time. The difference, we would suggest, is likely related to the different linguistic coding of time (temporal aspect of actions) in the two languages. The mechanism via which this difference develops and establishes itself may be attentional. That is, the explicit marking of temporal phases through tense and aspect may serve to direct speakers' attention to that component of action events correspondingly. Long-term experience with a language that carries such markings could then lead to a perceptual system that becomes attuned to a tripartite temporal analysis of action events.

Alternatively, the mechanism that drives the different conceptions of time may be cultural. Extensive research has shown that when viewing a scene, westerners (e.g., Americans) tend to pay attention to the focal object, while easterners (e.g., Chinese) tend to pay attention to the background context. Easterners are inclined to adopt a holistic perspective and view an event as continuous, while westerners are accustomed to an analytic perspective and tend to view an event as discrete (Nisbett, 2003; Ji, Zhang & Nisbett, 2004). The tendency to focus on the figure instead of on the ground and the analytic perspective could lead to automatic event segmentation and render the different temporal components of an event more distinct and identifiable for English speakers. The tendency to attend to the context and the holistic perspective could encourage Chinese speakers to view an action event as a non-segmentable whole, and to include the near future and the recent past as part of the extended present of the event.

The linguistic explanation and the cultural explanation would be difficult to distinguish in cross-cultural studies. However, our bilingual participants with high English proficiency were all local residents of Tainan, Taiwan. They learned English in school as one of the many subjects in their curriculum. They were thus not immersed in a Western culture. Accordingly, their English-like performance on the task in the present study more likely reflected their linguistic experience than a cultural adjustment.

To conclude, conceptions of time may be different in speakers of different languages, which encode the temporal phases of action events differently. Bilingualism across such languages can introduce and make available alternative framings of the temporal unfolding of ordinary events in bilingual speakers.

Appendix. A list of the 18 events described in Chinese and English sentences with respect to the three temporal phases (future, present, past)

Event	Chinese sentence			English sentence	
	準備要吹			is about to blow	
1	她 正在吹	氣球	She	is blowing	a balloon
	剛剛吹完			has just blown	
	準備要擦掉			is about to erase	
2	她 正在擦掉	白板上的東西	She	is erasing	something on a white board
	剛剛擦掉			has just erased	
	準備要跨過			is about to cross	
3	她 正在跨過	橫木	She	is crossing over	a log
	剛剛跨過			has just crossed over	
	準備要剪			is about to cut	
4	她 正在剪	-條繩子	She	is cutting	a rope
	剛剛剪斷			has cut	
	準備要丟			is about to throw	
5	她 正在丟	飛盤	She	is now throwing	a frisbee
	剛剛把飛盤丟出去			has just thrown	
	準備要喝			is about to drink	
6	她 正在喝	飲料	She	is drinking	a cup of tea
	剛剛喝完			has just finished drinking	
	準備要摺			is about to make	
7	她 正在摺	紙	She	is making	an origami
	剛剛摺完			has just finished	
	準備要溜			is about to slide down	
8	她 正在溜	滑梯	She	is sliding down	the slide
	剛剛溜下			has just slid down	
	準備要吃			is about to eat	
9	她 正在吃	香蕉	She	is eating	a banana
	剛剛吃完			has just eaten	
	準備要點			is about to light	
10	她 正在點	蠟燭	She	is lighting	a candle
	剛剛點燃			has lit	

https://doi.org/10.1017/S1366728911000320 Published online by Cambridge University Press

Appendix. Continued

Event	Chinese sentence			English sentence	
	準備要剝			is about to peel	
11	她 正在剝	橘子	She	is peeling	an orange
	剛剛剝好			has peeled	
	準備要走			is about to step on	
12	她 正在走	平衡木	She	is walking on	a log
	剛剛走完			has walked towards the end of	
	準備要切			is about to cut	
13	她 正在切	菜	She	is cutting	a lettuce
	剛剛切完			has cut	
	準備要拼			is about to begin doing	
14	她 正在拼	拼圖	She	is doing	a puzzle
	剛剛完成了			has just finished	
	準備要走上			is about to step on	
15	她 正在走	樓梯	She	is walking up	the stairs
	剛剛走完			has walked to the top of	
	準備要撕			is about to rip	
16	她 正在撕	-張紙	She	is ripping	a piece of paper (in two)
	剛剛把紙撕破			has ripped	
	準備要踢			is about to kick	
17	她 正在踢	球	She	is kicking	a ball
	剛剛把球踢出去			has just kicked	
	準備要倒			is about to pour	
18	她 正在倒	可樂	She	is pouring	coke into a glass
	剛剛倒完			has poured	

References

Alloway, T. P., & Corley, M. (2004). Speak before you think: The role of language in verb concepts. *Journal of Cognition and Culture, 4 (2),* 319–345.

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

Chang, J. P. (1998). Situation types and their temporal implicatures in Chinese. Ph.D. dissertation, University of Kansas.

Chen, J.-Y., & Su, J.-J. (2010). Chinese–English bilinguals' sensitivity to the temporal phase of an action event is related to the extent of their experience with English. In V. Cook & B. Bassetti (eds.), *Language and bilingual cognition*, pp. 341–356. Hove, UK: Psychology Press.

Chen, J.-Y., Su, J.-J., & O'Seaghdha, P. G. (2011). Enduring moments: The extended present in Chinese speakers' orientation to time. Ms., National Cheng Kung University, Taiwan. https://sites.google.com/site/lccncku/publications/papers/ChenSuOSeaghdha2011Manuscript.pdf, retrieved September 10, 2011.

Comrie, B. (1985). *Tense*. Cambridge: Cambridge University Press.

Davidoff, J., Fonteneau, E., & Goldstein, J. (2008). Cultural differences in perception: Observations from a remote

https://doi.org/10.1017/S1366728911000320 Published online by Cambridge University Press

culture. *Journal of Cognition & Culture, 8,* 189–209.

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science, 306,* 496–499.

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

Hunt, E., & Agnoli, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychology Review, 98 (3),* 377–389.

Ji, L.-J., Zhang, Z., & Nisbett, R. E. (2004). Is it culture or is it language? Examination of language effects in cross-cultural research on categorization. *Journal of Personality and Social Psychology, 87,* 57–65.

Klein, W., & Li, P. (eds.) (2009). *Expression of time*. Berlin: Mouton de Gruyter.

Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.

Li, C. N., & Thompson, S. (1981). *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.

Li, P., & Shirai, Y. (2000). *The acquisition of lexical and grammatical aspect*. Berlin: Mouton de Gruyter.

Lin, J. W. (2002). Selectional restrictions of tenses and temporal reference of Chinese bare sentences. *Lingua, 113,* 271–302.

Lin, J. W. (2003). Temporal reference in Mandarin Chinese. *Journal of East Asian Linguistics, 12,* 259–311.

Lin, J. W. (2006). Time in a language without tense: The case of Chinese. *Journal of Semantics, 23,* 1–53.

Lucy, J. A. (1996). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.

Nisbett, R. E. (2003). *The geography of thought: How Asians and Westerners think differently . . . and why*. New York: The Free Press.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306,* 499–503.

Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology, 50,* 378–411.

Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (eds.), *Language in mind: Advances in the study of language and thought*, pp. 157–192. Cambridge, MA: MIT Press.

Smith, C. S. (2006). Time with and without tense. In J. Guéron & J. Lecarme (eds.), *Time and modality* (Studies in Natural Language and Linguistic Theory 75), pp. 227–249. Dordrecht: Springer Science+Business Media.

Smith, C. S., & Erbaugh, M. S. (2005). Temporal interpretation in Mandarin Chinese. *Linguistics, 43,* 713–756.

Whorf, B. L. (1956). *Language, thought and reality: Selected writings of Benjamin Lee Whorf*, edited by J. B. Carroll. Cambridge, MA: MIT Press.

Wu, J.-S. (2009). Tense as a discourse feature: Rethinking temporal location in Mandarin Chinese. *Journal of East Asian Linguistics, 18,* 145–165.