



## Conceptual and lexical effects on gestures: the case of vertical spatial metaphors for time in Chinese

Yan Gu, Lisette Mol, Marieke Hoetjes & Marc Swerts

To cite this article: Yan Gu, Lisette Mol, Marieke Hoetjes & Marc Swerts (2017): Conceptual and lexical effects on gestures: the case of vertical spatial metaphors for time in Chinese, Language, Cognition and Neuroscience, DOI: [10.1080/23273798.2017.1283425](https://doi.org/10.1080/23273798.2017.1283425)

To link to this article: <http://dx.doi.org/10.1080/23273798.2017.1283425>



© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 08 Feb 2017.



Submit your article to this journal [↗](#)



Article views: 962



View related articles [↗](#)



View Crossmark data [↗](#)

## Conceptual and lexical effects on gestures: the case of vertical spatial metaphors for time in Chinese

Yan Gu<sup>a</sup>, Lisette Mol<sup>a</sup>, Marieke Hoetjes<sup>b</sup> and Marc Swerts<sup>a</sup>

<sup>a</sup>Tilburg center for Cognition and Communication, Tilburg University, Tilburg, The Netherlands; <sup>b</sup>Centre for Language Studies, Radboud Universiteit Nijmegen, Nijmegen, The Netherlands

### ABSTRACT

The linguistic metaphors of time appear to influence how people gesture about time. This study finds that Chinese English bilinguals produce more vertical gestures when talking about Chinese time references with vertical spatial metaphors than (1) when talking about time conceptions in the English translations, and (2) when talking about Chinese time references with no spatial metaphors. Additionally, Chinese English bilinguals prefer vertical gestures to lateral gestures when perceiving Chinese time references with vertical spatial metaphors and the corresponding English translations, whereas there is no such preference when perceiving time references without spatial metaphors. Furthermore, this vertical tendency is not due to the fact that vertical gestures are generally less ambiguous than lateral gestures for addressees. In conclusion, the vertical gesturing about time by Chinese English bilinguals is shaped by both the stable language-specific conceptualisations, and the online changes in linguistic choices.

### ARTICLE HISTORY

Received 10 March 2016  
Accepted 9 January 2017

### KEYWORDS

Gesture production and perception; space; time; metaphor; Chinese–English bilinguals

### Introduction

When people speak, they tend to accompany their utterances with *gestures*, in particular movements of speakers' hands, arms, and other body parts. These gestures are not accidental, but are functionally related to the ongoing speech and to the speakers' expressive intention (Kendon, 2004). Across cultures and languages, speakers' gestures can be vastly different. This has already been shown convincingly for specific classes of gestures such as *emblems*, as these rely on culture-specific conventions to associate specific gestural forms with certain meanings (Kendon, 2004; Kita, 2009). For instance, to express the number "two", Germans may perform a gesture by extending the thumb and index fingers with other fingers closed (like an "L"), whereas Chinese typically extend the index and middle fingers (the L-like German "two" would be interpreted as a gesture of "eight" by a Chinese). Additionally, studies have shown how cultures can differ regarding speech-accompanying gestures that are more spontaneously created on the fly, and are not conventionally associated with specific functions (e.g. Kita, 2009). The current paper addresses the latter kind of gestures, where we are specifically interested in so-called *temporal gestures*, that is, gestures that represent time conceptions, in

which temporal reference is made along the body's sagittal (front-to-back), lateral (left-to-right), or vertical (top-to-down) axis (Casasanto & Jasmin, 2012; Cooperrider & Núñez, 2009). For example, when talking about specific time events such as *last week* or *next week*, English speakers may point to the back and front of the body, or in a sequence from left to right (Casasanto & Jasmin, 2012; Cooperrider & Núñez, 2009), even though there is no explicit rule that prescribes that they should use their gestures this way.

Temporal gestures differ in various languages and cultures. For instance, speakers of Aymara typically position the future behind their back, which is consistent with their language (*qhipa mara*, literally *back year*, meaning *next year*, Núñez & Sweetser, 2006). Residents of Pormpuraaw gesture the past at the direction of their *East* because they always arrange temporal order from the east to the west (cardinal directions) (Boroditsky & Gaby, 2010). A case study (Chui, 2011) suggests that Chinese speakers can employ the vertical axis to gesture about time. However, it is as yet unclear why Chinese speakers produce vertical gestures to indicate time like this. The purpose of this study is first to examine whether Chinese speakers systematically produce vertical temporal gestures. If so, we investigate how these vertical gestures are produced and perceived

**CONTACT** Yan Gu  yan.gu@tilburguniversity.edu

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group  
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

under different circumstances (e.g. Are vertical gestures more often produced for a certain type of time reference, and are they still produced when Chinese people talk in English? When perceiving gestures about time, is there a bias for vertical gestures by Chinese people?).

### Time, space, and gestures

People use spatial representations to think about time (Bender & Beller, 2014; Boroditsky, 2000; Casasanto & Boroditsky, 2008; Núñez & Cooperrider, 2013), such as sundials, graphs, hourglass, clocks, timelines, and calendars; in ancient China one could also tell the time by the burning of incense (an incense stick was marked with regular intervals and the distance between each interval corresponded to a specific length of time). Studies have revealed that bodily, cultural, and environmental experiences can influence people's conceptualisations of time (de la Fuente, Santiago, Román, Dumitrache, & Casasanto, 2014). For instance, patients with left spatial neglect also have difficulty in thinking of the past (Saj, Fuhrman, Vuilleumier, & Boroditsky, 2014). Hebrew people have a writing direction from the right to the left, and also tend to think that time goes from the right to the left (Fuhrman & Boroditsky, 2010). Yupno speakers rely heavily on topographic contrasts (environment-based absolute terms) and construct the past as downhill and the future as uphill (Núñez, Cooperrider, Doan, & Wassmann, 2012). Additionally, spatial-temporal thinking can be rapidly affected by the context. For example, people's mental timelines can be reversed after brief exposure to mirror-reversed orthography (Casasanto & Bottini, 2014).

How people conceive time can also be derived from their lexical expressions, especially through the use of spatial metaphors, although the pattern of spatial metaphors that people use to talk about time can be different across languages. For example, it is quite common for speakers of English to say "The future lies not too *far ahead*". They can use their body as a reference point for the "now" and then conceptualise the past at their back, and the future in front (Calbris, 2008; Clark, 1973), or they can use the lateral axis, to order time from left (past) to right (future) (Santiago, Lupáñez, Pérez, & Funes, 2007). Therefore, in English, as well as in many other languages (e.g. Spanish, Dutch and Chinese), two metaphorical timelines are often employed: the lateral (left to right)<sup>1</sup> and the sagittal (front to back). Chinese speakers also use the vertical axis to express time, by employing vertical spatial metaphors of "上" (shàng: above) and "下" (xià: below) to indicate the time conceptions of early and late (e.g. Boroditsky, 2001). For example, "上周" (shàng zhōu) can literally be translated

as "above week", which means "last week", while "下周" (xià xià zhōu) as "below below week", referring to "the week after next week".

Interestingly, the metaphorical use of language for representing time can also be linked to how people spontaneously gesture about time. That is, the spatio-temporal concept can also be expressed in speakers' co-speech metaphoric gestures (e.g. Casasanto & Jasmin, 2012; Cienki, 1998; Cooperrider & Núñez, 2009; Núñez et al., 2012; Núñez & Sweetser, 2006; but for an alternative view, see e.g. de la Fuente et al., 2014; Le Guen & Balam, 2012). As mentioned above, Chinese speakers can employ a vertical axis to gesture about time. However, it is not yet clear exactly why Chinese speakers produce vertical gestures.

### Theories accounting for Chinese vertical gesturing about time

Due to the differences in the use of spatial metaphors in time conceptions, Boroditsky (2001) argued that Chinese speakers may have a different conceptualisation (a vertical one) of time than English speakers. Her argument is based on Slobin's (1987) "thinking-for-speaking" hypothesis that habitual speech patterns can shape thinking online, during linguistic processing. When the preferred speech patterns are repeatedly used, language-specific conceptual schemas may be habitually formed. Specifically for Chinese, Boroditsky (2001) believes that the habitual use of vertical spatial metaphors to talk about time shapes Chinese speakers' language-specific conceptual schema.<sup>2</sup> Interestingly, after learning Chinese vertical spatial metaphors, English speakers are also more inclined to think of time vertically (Boroditsky, 2001; Hendricks & Boroditsky, 2015). Additionally, Boroditsky found that Chinese speakers can conceptualise time vertically, even when they think in English. If it is the case that Chinese speakers have a long-lasting (habitual) vertical thinking of time, one would indeed assume that they can also gesture about time vertically, irrespective of whether they speak English or Chinese.

These ideas are in line with what has been claimed in theories about embodied cognition, which propose that conceptual representations are largely grounded in sensorimotor experiences (Glenberg & Kaschak, 2002), and that representations are activated and often instantiated in the forms of gestures (Hostetter & Alibali, 2008). In other words, the production of gestures is influenced by how people think of using the body to interact with the physical environment (or in Kita's (2000) terminology, *spatio-motoric thinking*). Specifically, in this case it is the way how one thinks of time in space that affects the gestural representation.

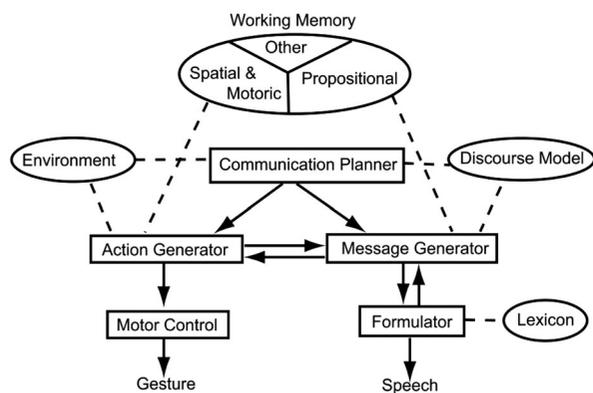
However, a slightly alternative reasoning is that Chinese may gesture vertically simply as a result of the fact that they use specific lexical words that express time, which in turn drive the way they gesture. In other words, such a view stresses the fact that speakers tend to align their gestures with the lexical representations in order to make these congruent (Kita et al., 2007; Kita & Özyürek, 2003; Özyürek, Kita, Allen, Furman, & Brown, 2005); hence Chinese speakers will produce vertical temporal gestures when speaking about time conceptions through the use of vertical spatial metaphors.

According to Kita and Özyürek's (2003) Interface Hypothesis, spontaneous gestures are not only shaped by the imagistic (spatio-motoric) representations of events, but are also adjusted to be compatible with linguistic encoding possibilities. In other words, the generation of a gesture is modulated by two forces: the spatio-motoric experience from the Working Memory, and the linguistic choices (linguistic formulation possibility, for example, different semantic and syntactic choices that a speaker can choose from when communicating, Kita et al., 2007) from the Message Generator, both of which interact with each other (see Figure 1). The linguistic influence on gesture production is represented in the model by the arrow running from the Message Generator to the Action Generator. Then the Action Generator determines the ultimate content of a gesture by taking into account the two forces, such that gestures are adjusted to fit the verbalisation.

For instance, Kita and Özyürek (2003) found that a scene of a "Rolling Event" could be expressed as "rolling down" in English, with manner and path conflated into a single clause. Accordingly, this information tended to be conflated in gesture too, with one gesture expressing both manner and path. However, in Turkish and Japanese (verb-framed languages), manner and path need to be expressed separately in two

clauses (i.e. "move down, in a rolling fashion"). Consequently, compared to English speakers, Turkish and Japanese speakers were more likely to produce manner only and path only gestures. Interestingly, the cross-linguistic differences in gestures disappeared when English and Turkish speakers were asked to describe motion events without the speech but only by gestures (silent gesture). This indicates that the verbal task had an effect on the gesturing of motion and path (Özçalışkan, 2016; Özçalışkan, Lucero, & Goldin-Meadow, 2016). Similar evidence was also revealed by an eye-tracking study, in which participants were instructed to watch animations depicting motion events (e.g. skating) that participants later had to *describe*. Native speakers of Greek (a verb-framed language similar to Turkish) were significantly more likely to look at path (where the moving character was heading) over manner (instrument regions, e.g. skating – the area of the feet that included the skates) than native speakers of English. The findings suggest that there are cross-linguistic differences in how people distribute visual attention to components of a scene when preparing for language production. However, there were no such differences when they were simply told to watch the video clip (Papafragou, Hulbert, & Trueswell, 2008). In sum, the studies above suggest that a speaker's use of gestures is not only a result of spatio-motoric processing, but is affected online by specific linguistic choices as well (such as specific syntactic or lexical surface forms).

In addition to research that focused on differences between languages, there are several studies that looked at gesture production by bilinguals, although these have yielded some mixed results. Speakers of verb-framed languages (e.g. Spanish, Turkish) often produce gestures for the path of the motion with verbs, and tend not to accumulate gestures for path and manner, whereas speakers of satellite-framed languages (e.g. English, Dutch) tend to produce path gestures with a satellite component, and tend to accumulate gestures for path and manner in a single clause. First, some case studies about bilinguals' gestures for motion events described that Spanish/Turkish learners of English maintained an L1-like gesture pattern in the L2, for example, Spanish and Turkish speakers still performed path gestures with verbs when speaking English, although the findings were based on a very small sample (Kellerman & Van Hoof, 2003; Negueruela, Lantolf, Jordan, & Gelabert, 2004; Stam, 2006). Second, another study found that there was a parallel trend in L2 speech and gesture production. Turkish learners of English were more likely to use conflated manner-path gestures when they verbally used conflated constructions in the L2 English (Özyürek, 2002). Finally, there is



**Figure 1.** Kita and Özyürek's (2003) speech and gesture production model. Reprinted with permission.

also evidence of bi-directional influences of L1 and L2 (co-activation of both languages) in both speech and gesture production. For example, Brown and Gullberg (2008) found that Japanese learners of English (L2 English) did not differ significantly in how they encoded “manner” in speech or gestures in their L1 or L2 productions, but their gesture pattern differed significantly from that of monolingual Japanese speakers and of monolingual English speakers. In short, previous studies on motion events suggest that gestures may result both from specific mental representations and from linguistic choices, but it remains to be explored how these two factors relate or interact when bilinguals describing abstract conceptions such as time events.

### The current study

In the current study we investigate gestures about time references. First we will find out whether Chinese–English bilinguals perform vertical temporal gestures. If so, the study aims to shed light on why they produce vertical gestures to indicate time. Specifically, we aim to explore whether the production of vertical gestures is due to the fact that Chinese people have a stable vertical time conceptualisation and “think vertically” when visualizing time (Boroditsky, 2001), or is it because their gestures are merely a result of the fact that they also use specific words that express vertical spatial metaphors of time (“above week” for “last week”/“below week” for “next week”), or are the vertical gestures a consequence of both these factors? We answer these questions by studying whether lexical choices of vertical temporal-spatial expressions have any an online influence on vertical gesturing.

We have set up a series of experiments in which lexical choices are manipulated in two important ways, to see how these affect the production and perception of gestures. Firstly, we introduce a within-language factor, in that we vary the linguistic expressions for time conceptions within Chinese, by comparing gestures in utterances in which time is expressed by vertical spatial metaphors (e.g. “上周”/ “shàng zhōu”, “above week”, meaning “last week”) with utterances that do not contain such a spatial metaphor (neutral words such as “昨天” *zuó tiān*, meaning “yesterday”). This will allow us to see to what extent the type and frequency of specific gestures are affected by the mere presence of specific words that express this vertical time conception. If such a lexical trigger would be the determining factor for the gestural representation of time, one would expect relatively few vertical gestures in utterances that do not contain an explicit vertical spatial metaphor.

Secondly, we make a comparison between two languages, by exploring the gestural expression of time in Chinese and English by Chinese–English bilinguals, given that time conceptions with Chinese vertical spatial metaphors have different, non-vertical, lexical correlates in English (e.g. “above week” in Chinese, “last week” in English). If the vertical mental representation of time in Chinese speakers is the most important determining factor for the choice of gestures, and assuming everyone can only have one conceptual scheme (e.g. Kellerman & Van Hoof, 2003; Negueruela et al., 2004; Stam, 2006), then one would predict that Chinese–English bilinguals will also gesture vertically even when speaking English. That means their vertical gesturing will be unaffected by the language itself. Alternatively, if gestures are more strongly caused by linguistic choices, one would expect speakers to gesture more vertically when speaking Chinese than English, especially for words with a lexical trigger in Chinese.

These questions are addressed both from the perspective of a speaker who spontaneously produces gestures while speaking (Experiment 1), and the perspective of an addressee who processes another person’s co-speech gestures (Experiment 2 and 3).

By researching the gestural representations of time in Chinese–English *bilinguals*, we can provide a better understanding of the cognitive processing of the production and perception of co-speech gestures, and we can shed light on the respective roles of lexical choices and mental representations in bilingual language processing.

### Experiment 1: production experiment

In the production experiment, we address two questions: (1) In Chinese, will verbally producing time conceptions with vertical spatial metaphors (“above week”, “below month”, etc.) lead to more vertical gestures than in the case of verbally producing time conceptions without a spatial metaphor (e.g. neither in Chinese nor in English are “yesterday” and “tomorrow” spatial expressions with reference to time)? (2) Given the fact that English and Chinese speakers may think of time differently (Boroditsky, 2001), will Chinese–English bilinguals produce more vertical gestures in Chinese than in English?

### Method

#### Participants

Forty-six late Chinese–English bilinguals (L1 = Chinese; 35 F and 11 M; mean age = 24 yrs, ranged from 19 to 38 yrs, *SD* = 3.8 yrs) were paid for participation. They were students attending English-taught international

programmes at Tilburg University (the Netherlands), who originally came from China. Bilinguals were defined as *sequential bilinguals*, who first acquired Mandarin as the L1 and then English as the L2 (Average age of acquisition = 11yrs). Their English proficiency was between intermediate and advanced, as assessed by a Quick Placement Test (UCLES, 2001) and a 5-point-scale (1 for beginner and 5 for very advanced) self-report ( $M = 3.53$ ,  $SD = .64$ ).

### Stimuli

Eleven Chinese wordlists were constructed for a word definition task. Four wordlists were relevant for the current study, which in total consisted of eleven expressions that conveyed time conceptions. The number of expressions in each wordlist ranged from two to four (Table 1). Wordlists (1) and (2) were time references containing words with vertical spatial metaphors (“上”/shàng, above, and “下”/xià, below) to indicate the conceptions of “early” and “late”. By contrast, words in wordlists (3) and (4) did not contain explicit lexical references to vertical (“shàng”, above or “xià”, below), or sagittal (“前”/qián, front or “后”/hòu, back) space dimensions.

### Procedure

Each participant was tested individually, in interactions with an unknown addressee who would not become a participant (speaker) afterwards. The experiment was ostensibly set up as a language and memory test, in which the speaker’s short-term memory and the addressee’s long-term memory would be tested. They sat face-to-face in a quiet room, where a monitor was placed in front of the speaker. The wordlists were presented in the centre of the monitor and could only be seen by the speaker (Figure 2). Each wordlist consisted of several expressions that were thematically related (e.g. “last week” and “next week”). The expressions within a wordlist were shown on separate slides to avoid that they would be presented laterally or vertically and thus suggesting a specific plane, but each word within an expression was presented laterally (Figure 2). Furthermore, the speaker was informed that s/he would have

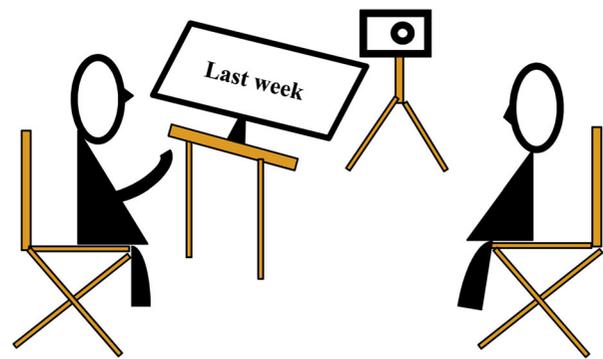


Figure 2. Schematic illustration of the experimental set-up.

to remember the wordlists shortly after having seen them twice, find the relationship between the expressions within a wordlist and explain the definitions of the expressions to the addressee as clearly as possible. Note that it was not a word-guessing game, as speakers were allowed to tell the expressions in the wordlists. The addressee was told to remember the speaker’s descriptions as much as possible for a memory test afterwards, and s/he was allowed to ask the speaker clarification questions. After the task, the addressee was taken to another room for the memory test, which s/he did not actually need to complete.

All participants took part twice in the experiment, once in Chinese and once in English. The instructions were given in the language of the experiment. The testing order of the languages was counterbalanced and the interval between the 2 tests was approximately 10 days (to reduce possible learning effects). The addressees were native speakers of Chinese for the Chinese task, and Dutch–English/English–Dutch bilinguals for the English task. The entire experiment was videotaped with participants’ consent. There was no mentioning of gestures at any point during the experiment, and speakers were not explicitly informed to gesture during the production task. After the second session of the experiment, participants were given a questionnaire to fill in some background information such as the age of acquisition, length of residence in the Netherlands, and writing experiences. Debriefing responses indicated that participants had not been aware that the purpose of the study was to investigate speakers’ gestures.

Table 1. Wordlists of targeted time referents.

	Chinese	English
(1)	上周, 下周 (vertical) shàng zhōu, xià zhōu	Last week, next week (neutral English translations of vertical)
(2)	上辈子, 下辈子 (vertical) shàng bèi zi, xià bèi zi	Previous life, next life (neutral English translations of vertical)
(3)	昨天, 今天, 明天 (neutral) zuó tiān, jīn tiān, míng tiān	Yesterday, today, tomorrow (neutral)
(4)	早晨, 晌午, 傍晚, 深夜 (neutral) zǎo chén, shǎng wú, bàng wǎn, shēn yè	Morning, noon, evening, late at night (neutral)

### Coding and measurements

The temporal gestures accompanying the speech that described the target wordlists were annotated in ELAN (Lausberg & Sloetjes, 2009). A first coder (the first author) performed an initial coding, viewing the entire video with the accompanying audio. The planes of gestures were categorised as vertical, lateral, or sagittal,

and the directionality of each plane was also indicated (Casasanto & Jasmin, 2012). When a wordlist was produced with gestures of no codable direction or no/non-temporal gestures, it was coded as “other”.

A participant could explain a wordlist by using temporal gestures from several planes, which were coded in each plane accordingly. The present study focused on vertical temporal gestures, so wordlists with sagittal or lateral temporal gestures and with “other” gestures were all treated as “non-vertical” in the later analyses.<sup>3</sup> Each wordlist got one binary score: containing a vertical gesture or not. Four participants did not produce any gestures in Chinese or in English for all target wordlists and one participant did not finish the experiment. Data from these five people were excluded from the analysis.

The reliability of the annotation was established by having 15% of the data coded by a second coder, who was naïve to the research question at the time of coding. The two coders agreed on the gesture plane judgement on 91.1% of the tokens ( $N = 56$ ), Cohen’s Kappa = 0.85 (referring to “Excellent” agreement). In cases of disagreement, the two coders discussed and reached agreement on the labels, which were used for the final analysis.

In the subsequent analyses, the wordlist-type (vertical and neutral) and the language factor (Chinese and English) were the independent variables. The type of a temporal gesture (vertical, non-vertical) accompanying the description of a wordlist was the binary dependent variable.

## Results and discussion

Participants produced a total of 328 ( $41 \times 4 \times 2$ ) target wordlists for both languages, 269 (82.0%) of which were accompanied by clear temporal gestures [128 (78.0%) for Chinese and 140 (85.4%) for English]. The remaining wordlists (18%) were produced either with gestures having no codable direction or no/non-temporal gestures. The average number of temporal gestures per wordlist in English ( $M = 3.51$ ,  $SD = .71$ ) was significantly higher than in Chinese ( $M = 3.14$ ,  $SD = 1.06$ ),  $t(40) = 1.95$ ,  $p = .029$ ,  $r = .30$ . This is to be expected, because late bilingual speakers have been shown to gesture more often in their second language than in their first language (Gullberg, 1998).

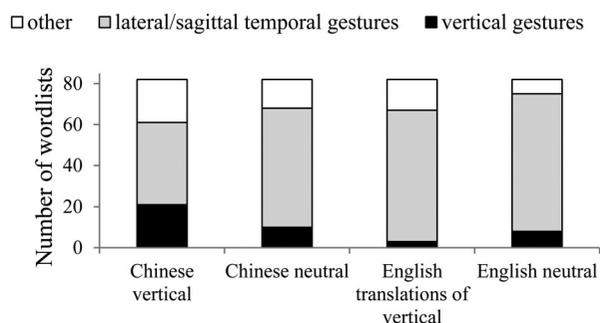
Based on the binary coding, each wordlist was accompanied with either a “vertical” or “non-vertical” gesture. The proportion of vertical temporal gestures for each type of time references was computed as the total number of wordlists that were accompanied by vertical gestures divided by the total number of wordlists. As shown in Figure 3, vertical gestures were produced for

both types of wordlists in both Chinese and English. The fact that vertical gestures were produced for time references with *neutral words* in Chinese and English suggests that Chinese–English bilinguals *can* employ a vertical conceptualisation of time, even when switching to a second language (cf. Boroditsky, 2001).

The proportion of vertical gestures was compared as a function of Wordlist-type (vertical, neutral) and Language (Chinese, English) using a random effects binary logistic regression.<sup>4</sup> This model considers multiple responses from the same participants and takes individual differences (random effects) into account.

First, Chinese wordlists with vertical spatial metaphors were accompanied by a significantly higher proportion of vertical gestures than the neutral Chinese wordlists, ( $\beta = -1.11$ , Wald  $\chi^2 = 6.04$ ,  $df = 1$ ,  $p = .0014$ ) (Figure 3). This indicates that producing time references with vertical spatial metaphors had an online effect on the production of vertical gestures. By contrast, the difference between the two types of wordlists was not statistically significant in English ( $\beta = 1.16$ , Wald  $\chi^2 = 2.51$ ,  $df = 1$ ,  $p = .113$ ). This shows that the production of vertical temporal gestures was sensitive to the linguistic choices.

Additionally, as for the comparisons between languages, firstly, the proportion of vertical gestures for wordlists with vertical spatial metaphors was significantly higher in Chinese than in the English translations ( $\beta = -2.68$ , Wald  $\chi^2 = 15.12$ ,  $df = 1$ ,  $p = .0001$ ). One may explain this as the result of simply having two different language-specific conceptualisations of time, which are pre-determined by Chinese and English. If this is true, we would expect the same pattern in the case of wordlists with neutral words. However, for the neutral wordlists, there was no significant difference in the proportion of vertical gestures between the two languages ( $\beta = -.41$ , Wald  $\chi^2 = .61$ ,  $df = 1$ ,  $p = .433$ ). This is supported by finding a significant interaction



**Figure 3.** The proportion of wordlists that were accompanied by vertical gestures for the four types of time references (82 for each type). Chinese vertical/English translations of vertical = wordlists 1 and 2; Chinese neutral/English neutral = wordlists 3 and 4 (Table 1).

**Table 2.** Summary of the output of the production results by a random effects binary logistic regression.

Comparisons	Coefficients	$\chi^2(1)$	<i>p</i> -value	Marginal effects	
Chinese vertical vs. Chinese neutral	$\beta_{\text{neutral}}$	-1.11	6.04	0.0140	-18.25%
Chinese vertical vs. English translations of vertical	$\beta_{\text{english}}$	-2.68	15.12	0.0001	-35.81%
English translations of vertical vs. English neutral	$\beta_{\text{neutral}} + \beta_{\text{interaction}}$	1.16	2.51	0.1133	
English neutral vs. Chinese neutral	$\beta_{\text{english}} + \beta_{\text{interaction}}$	-0.41	0.61	0.4335	
Language $\times$ wordlist-type interaction	$\beta_{\text{interaction}}$	2.27	6.87	0.0087	

Notes:  $\beta_{\text{interaction}}$  is the estimated coefficient of the interaction term of language and wordlist-type. Marginal effects corresponding to  $\beta_{\text{neutral}}$ : For an average person, given s/he speaks Chinese, if the wordlist-type changes from vertical to neutral, then s/he is 18.25% less likely to perform vertical gestures. Marginal effects corresponding to  $\beta_{\text{english}}$ : For an average person, given s/he speaks Chinese vertical wordlists (or English translation of vertical), if his/her language changes from Chinese to English, then s/he is 35.81% less likely to perform vertical gestures.

between language and wordlist-type ( $\beta = 2.27$ , Wald  $\chi^2 = 6.87$ ,  $df = 1$ ,  $p = .0087$ ) (Table 2).

Why did the temporal gestures of the two types of time references (vertical and neutral) display such differences across languages? It seems that apart from the spatio-motoric thinking (the stable vertical spatial-temporal mappings), there is another force – the force from the linguistic encoding possibilities, resulting in the increase in vertical gestures for time references with vertical spatial metaphors in Chinese as compared to time references with neutral spatial metaphors in Chinese. In other words, the increase in vertical gesturing for time conceptions with vertical spatial metaphors in Chinese can be triggered by the specific online lexical expressions.

To illustrate the effect of within and between-language linguistic choices, Figure 4 shows a participant



**Figure 4.** Gestures of *last week* and *next week* in Chinese, and *yesterday* and *tomorrow* in Chinese, and *last week* and *next week* in English by the same participant.

producing vertical gestures for time references of “last week” and “next week” in Chinese, while producing lateral gestures for “yesterday” and “tomorrow” in Chinese and for “last week” and “next week” in English.

In theory, it could be that, in our experiment, factors such as L2 proficiency levels, age of acquisition, culture exposure and writing experiences may have influenced the production of vertical gestures. However, when we put all these factors into the model, the effect of linguistic choices (within-language comparisons of two types of Chinese wordlists and between-language comparisons of vertical wordlists in Chinese and English translations) remained still highly significant, even after controlling for the proficiency levels of L2 English, age of L2 acquisition, length of living in the Netherlands and vertical experience (writing and reading) (all *ns.*). Furthermore, we included task order as a factor, which revealed no evidence for a possible effect that participants who did the task first in Chinese were more likely to produce vertical gestures in English, compared to those participants who did the task first in English ( $\beta = .009$ , Wald  $\chi^2 = .0$ ,  $df = 1$ ,  $p = .99$ ).

In summary, our results show that Chinese–English bilinguals’ vertical gesturing about time is not only shaped by the stable vertical conceptualisation (Figure 3), but also by the online linguistic choices. We discuss this more in general discussion.

## Experiment 2

An increasing number of studies show that the production and the perception of speech and gesture are interconnected (Pickering & Garrod, 2013). For instance, perceiving gestures automatically activates the brain areas involved in producing these corresponding actions (Hostetter & Alibali, 2008; Mashal, Andric, & Small, 2011). Additionally, studies have shown that there is an on-line integration of semantic information from speech and gesture. That is, co-occurring speech and gestures are integrated simultaneously into a preceding sentence context (e.g. Özyürek, Willems, Kita, & Hagoort, 2007). This raises the question whether the perception of vertical gestures about time by the Chinese–

English bilinguals is also influenced by the linguistic encoding possibilities.

In experiment 2, we address two questions: (1) In Chinese, will observers prefer vertical gestures for phrases with explicit vertical spatial metaphors (e.g. “above week”) over vertical gestures for phrases that do not have such an explicit spatial indicator (so without words that either a vertical or sagittal spatial metaphor)? (2) Will there be perceptual differences in that respect between Chinese and English?

If one’s perception of gesture is also sensitive to the online linguistic choices, then, firstly, Chinese–English bilinguals are likely to prefer vertical gestures to lateral ones for time references with Chinese vertical spatial metaphors. Also, Chinese–English bilinguals are likely to prefer vertical gestures for time references with vertical spatial metaphors compared to neutral words. Moreover, we expect there to be less of a preference for a vertical gesture plane in English than in Chinese, especially for time references with vertical spatial metaphors (since in English the Chinese vertical wordlists were translated into neutral words). To test these hypotheses, Chinese–English bilinguals were asked to do a rating task.

## Method

### Participants

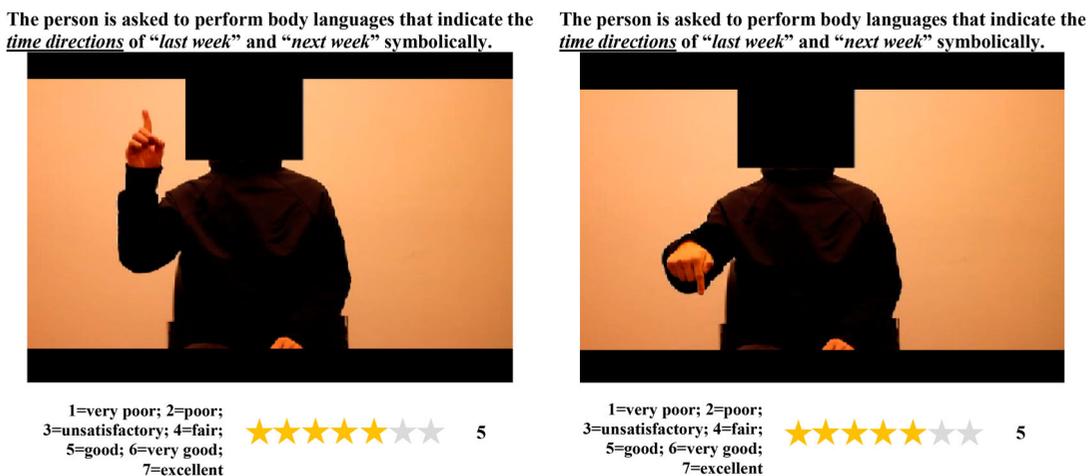
Hundred and nine Mandarin-English sequential bilinguals (L1 = Chinese; 52 F and 57 M; mean age = 18 yrs) from Nanjing University, China were paid to participate in the experiment.

### Stimuli

Thirty items, consisting of 8 target items of time references and 22 fillers were performed by an actor. Each

item consisted of a sentence which was followed by a silent video clip of the seated actor (visible from shoulders to upper legs) who made specific gestures. To avoid possible distractions due to some culture-specific facial expressions, the face was hidden by a digitally inserted black square on the face. This also enabled the same stimuli to be used both in the English and Chinese version. For instance, a sentence was shown as “The person is asked to perform body languages that indicate the **time directions** of *last week* and *next week* symbolically”. The clip underneath the sentence showed an actor who first pointed to his left side and then to the right side (lateral gesture plane) or, in a counterbalanced version, pointed upward and downward (vertical gesture plane) to indicate the time conceptions of “last week” and “next week” (Figure 5). The direction of the movements (from left to right in the lateral axis, from top to bottom in the vertical axis, both from the speaker’s perspective) was determined by observations in the production study, where these movements were much more likely than their counterparts (right to left, bottom to top). The exact same video clips were used in English and Chinese and the format of sentence instruction was consistent for all the target time references; the only thing that varied was the instructions in either English or Chinese.

To prevent participants from being too conscious of the many repetitive judgements of temporal gestures, only eight sentences had time references among all the item instructions. As shown in Table 3, half of time references (1–4) contained vertical spatial metaphors, and the other half were literally neutral (no spatial metaphors) in Chinese. Half of the time references [(1) and (2); (5) and (6)] in the sentences were performed in the vertical gesture plane and the other half were performed in



**Figure 5.** An example, of a gesture clip. Stills from the two gestures that were used as a stimulus of *last week* and *next week* (the vertical plane).

**Table 3.** Time references with vertical spatial metaphors and time references with neutral words, as used in the perception experiment.

	Chinese	English
(1)	上周, 下周 (vertical) shàng zhōu, xià zhōu	Last week, next week (neutral English translations of vertical)
(2)	上午, 下午 (vertical) shàng wǔ, xià wǔ	a.m., p.m. (neutral English translations of vertical)
(3)	上月, 下月 (vertical) shàng yuè, xià yuè	Last month, next month (neutral English translations of vertical)
(4)	上学期, 下学期 (vertical) shàng xué qī, xià xué qī	Last semester/term, next semester/term (neutral English translations of vertical)
(5)	昨天, 明天 (neutral) zuó tiān, míng tiān	Yesterday, tomorrow (neutral)
(6)	早晨, 傍晚 (neutral) zǎo chén, bàng wǎn	Morning, evening (neutral)
(7)	去年, 明年 (neutral) qù nián, míng nián	Last year, next year (neutral)
(8)	早春, 晚春 (neutral) zǎo chūn, wǎn chūn	Early spring, late spring (neutral)

the lateral plane. These references were counterbalanced in gesture planes by creating a second version.

Additionally, 11 fillers (unrelated to time conceptions) had incongruent gestures with the sentence instructions (e.g. an incongruent gesture for the concept of “over there” was presented as pointing to the actor’s own body), and the other 11 fillers were congruent. The text was created in Chinese and the English text was a translation of the Chinese. Note that wordlists (1), (2), (3) and (4) in the sentences were again neutral English translations of vertical words.

### Procedure

Participants took part twice in the experiment, once in Chinese and once in English, with an interval of one week. The sequence of the languages tested was counterbalanced. In the first test, participants were randomly assigned to one of the two versions (counterbalanced in gesture planes) to fill out a 1-7-Likert scale rating task in a large computer classroom. They were shown the sentence instructions and the silent video clips. Participants were asked to judge the extent to which the gestures in the clip expressed the instruction correctly, with 1 meaning “very poor” and 7 meaning “excellent” (see Figure 5). Data from 30 participants were excluded because they either did not show up for the second part of the experiment (19), or failed to comply with the instructions (11).<sup>5</sup> All data were collected via a survey programme called Qualtrics.

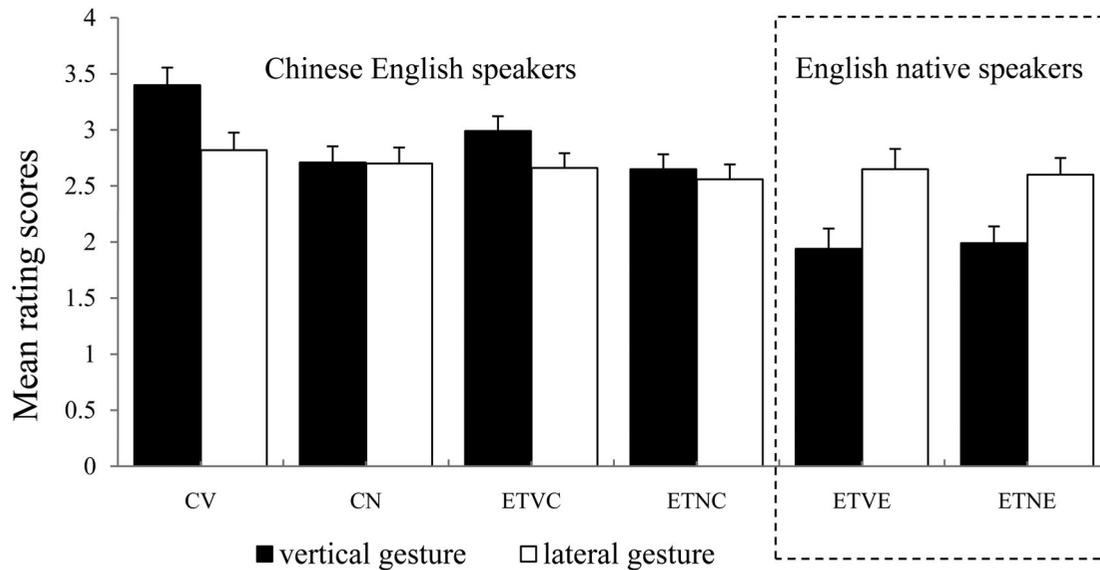
### Results and discussion

There was no significant difference between the two versions in counterbalancing gesture planes,  $F(1, 77) = .031$ ,  $p = .86$ ,  $\eta_p^2 = .00$ , so the data from the two versions was merged for further analyses. A  $2 \times 2 \times 2$  repeated measures ANOVA with Wordlist-type (vertical and neutral)  $\times$  Gesture plane (lateral and vertical)  $\times$  Language (Chinese and English) as within subject factors, and rating scores as dependent variable revealed that there were main effects of wordlist-type,  $F(1, 78) = 37.11$ ,  $p < .001$ ,  $\eta_p^2 = .32$  and gesture plane,  $F(1, 78) = 23.65$ ,

$p < .001$ ,  $\eta_p^2 = .22$ , but there was no main effect of language,  $F(1, 78) = 1.83$ ,  $p = .18$ ,  $\eta_p^2 = .02$ . Furthermore, there was a significant interaction between wordlist-type and gesture plane,  $F(1, 78) = 17.08$ ,  $p < .001$ ,  $\eta_p^2 = .18$ , and a significant interaction between wordlist-type and language,  $F(1, 78) = 4.26$ ,  $p = .042$ ,  $\eta_p^2 = .052$  (see Figure 6 for mean rating scores).

Follow-up Bonferroni adjusted *t*-tests (with alpha level corrected) were conducted. Firstly, for Chinese wordlists containing vertical spatial metaphors (Chinese vertical), participants preferred vertical gestures ( $M = 3.40$ ,  $SD = 1.38$ ) to lateral gestures ( $M = 2.82$ ,  $SD = 1.27$ ),  $t(78) = 5.53$ ,  $p < .001$ ,  $r = .53$ , 95% CI = (.37, .79). At first sight, this might suggest a Chinese speakers’ general preference of vertical gestures over lateral ones. However, there were no such preferences in time references with neutral words (Chinese neutral),  $t(78) = .13$ ,  $p = .90$ ,  $r = .01$ , 95% CI = (−.18, .20). Also, participants rated vertical gestures for “Chinese vertical” wordlists ( $M = 3.40$ ,  $SD = 1.38$ ) higher than those for “Chinese neutral” wordlists ( $M = 2.71$ ,  $SD = 1.17$ ),  $t(78) = 8.20$ ,  $p < .001$ ,  $r = .68$ , 95% CI = (.52, .86), but they did not rate the lateral gestures significantly different for the two type of wordlists,  $t(78) = 1.25$ ,  $p = .21$ ,  $r = .14$ , 95% CI = (−.07, .31) (Chinese–English speakers in Figure 6). Assuming that the neutral and metaphoric time references share the same conceptualisation of time within a language, the major discrepancy between the two is the difference in linguistic choices. This indicates that the vertical spatial metaphor of time referents has an online influence on perceiving the vertical temporal gestures.

With respect to English stimuli, the pattern of results was quite similar to that in Chinese. For the English translations of Chinese vertical spatial metaphors (ETVC), participants also preferred vertical gestures ( $M = 2.99$ ,  $SD = 1.32$ ) to lateral gestures ( $M = 2.66$ ,  $SD = 1.13$ ),  $t(78) = 2.90$ ,  $p = .005$ ,  $r = .31$ , 95% CI = (.10, .56), but they did not have such a preference for wordlists without spatial metaphors (English translations of neutral) (even though all these translations no longer contained explicit lexical markers of space). Also, vertical gestures for “English translations of vertical (ETVC)” ( $M = 2.99$ ,  $SD =$



**Figure 6.** The mean rating scores and standard errors of the two types of time references (words with vertical spatial metaphors and with neutral words) in Chinese and in English by Chinese–English observers, and in English by native English observers. Notes: CV = Chinese vertical, CN = Chinese neutral, ETVC = English translations of vertical for Chinese, ETNC = English translations of neutral for Chinese, ETVE = English translations of vertical for English, ETNE = English translations of neutral for English.

1.32) were rated significantly higher than vertical gestures for the “English translations of neutral (ETNC)” wordlists ( $M = 2.65$ ,  $SD = 1.06$ ),  $t(78) = 8.20$ ,  $p < .001$ ,  $r = .68$ , 95% CI = (.16, .53).

Further comparisons between Chinese and English revealed that participants rated vertical gestures with a higher score when perceiving Chinese wordlists containing vertical spatial metaphors ( $M = 3.40$ ,  $SD = 1.38$ ) than when perceiving the English translations of the wordlists ( $M = 2.99$ ,  $SD = 1.32$ ),  $t(78) = 2.05$ ,  $p = .022$ ,  $r = .23$ , 95% CI = (.11, .81), but the bias towards vertical gestures in Chinese no longer existed when it came to wordlists with neutral words,  $t(78) = .43$ ,  $p = .67$ ,  $r = .05$ , 95% CI = (–.21, .35) (black bars for Chinese–English bilinguals in Figure 6). As for the rating score for the lateral gestures, the difference between the two languages was not significant for either type of wordlist (white bars). In short, only when perceiving the “Chinese vertical” and the corresponding “English vertical translations”, did participants prefer vertical gestures in Chinese over those in English.

The similar pattern of results from the L1 and L2 may seem to suggest that Chinese speakers still thought in their first language when perceiving the English words. If Chinese speakers used L1 as a mediating factor in the L2 English for the translations of vertical wordlists, we would expect that participants with a lower L2 proficiency were more likely to translate the words from Chinese than participants with a higher L2 proficiency, and thus would be influenced more by the Chinese vertical linguistic choices and would therefore be more in

favour of the vertical axis. Participants’ 5-point-self-assessment of English proficiency showed that their proficiency levels were between intermediate and advanced,  $M = 2.7$ ,  $SD = 0.84$ . Further analysis on the correlation between the scores of English vertical translations and L2 proficiency showed that the rating of verticality was not related to the L2 English proficiency,  $r = -.005$ ,  $p = .97$ . Thus the similar pattern of L1 and L2 results suggests that Chinese vertical spatial metaphors were activated in the English context, which is consistent with Wu and Thierry’s (2010) proposal that languages are co-activated in a bilingual speaker, even if only one language is contextually relevant.

However, it could be that a subjective self-assessment of L2 proficiency is uninformative, and an objective language proficiency test is recommended in future studies. Another possibility could be that participants, when doing the English test, have translated the sentences internally into their native language, Chinese. It is still possible that the translation process for these frequent and salient abstract concepts is so pervasive and automatic that it is unaffected by the bilingual’s self-reported proficiency level.

To sum up, the results from this perception study show that firstly, in Chinese, the bilingual speakers preferred vertical gestures for time references with vertical spatial words to vertical gestures for neutral words. Secondly, when perceiving time references with vertical spatial metaphors, they preferred vertical gestures to lateral gestures. Additionally, participants preferred vertical gestures for time references containing vertical

spatial metaphor in Chinese to vertical gestures for the English translations.

However, as the lateral temporal gestures took the speakers' perspective, the left-right mapping onto to a timeline was mirrored (right-left) for the addressees. It could have been difficult for the addressee to deal with the mirror effect whereas describing expressions by vertical gesture is unambiguous for the addressee, as top and bottom are similar for speakers and addressees. This may also be a reason why vertical gestures are preferred over lateral gestures. To rule out the possibility of a vertical preference that is brought about by some general non-linguistic factors, we did another perception experiment, with native speakers of English.

### Experiment 3

#### Method

##### Participants

73 American English speakers (40 F and 33 M; mean age = 39 yrs) were paid to take part in this experiment via Crowd-Flower, a crowdsourcing service similar to Amazon Mechanical Turk. The validity of this method for behavioural studies has been previously tested and studies assessing data quality have been positive about using crowdsourcing as an alternative to more traditional approaches of participant recruitment (e.g. Buhrmester, Kwang, & Gosling, 2011; Crump, McDonnell, & Gureckis, 2013). Data from 19 respondents were excluded from the analyses because they were not native English speakers (5), or did not finish the task in a proper way (15) (e.g. finished the task in less than two minutes, clicked the same choice for the whole task or misunderstood the task).

##### Procedure

Participants were instructed to do the same rating task as that in Experiment 2. The Qualtrics links were provided via CrowdFlower.

### Results and discussion

A 2×2 repeated measures ANOVA with Wordlist-type (vertical and neutral) × Gesture plane (lateral and vertical) as within subject factors, and rating scores as dependent variable revealed that there was no main effect of wordlist-type,  $F(1, 53) = .003$ ,  $p = .96$ ,  $\eta_p^2 = .00$ , but there was a main effect of gesture plane,  $F(1, 78) = 23.65$ ,  $p < .001$ ,  $\eta_p^2 = .32$ . Furthermore, there was no interaction between wordlist-type and gesture plane,  $F(1, 53) = .34$ ,  $p = .56$ ,  $\eta_p^2 = .006$ . The results show that English native speakers preferred lateral gestures to vertical gestures for both types of time references (English

translations of Chinese vertical time references for English, ETVE:  $M = 2.65$ ,  $SD = 1.52$  vs.  $M = 1.94$ ,  $SD = 1.07$ ; English translations of Chinese neutral time references for English, ETNE:  $M = 2.58$ ,  $SD = 1.36$  vs.  $M = 1.99$ ,  $SD = 1.07$ , Figure 6).

The pattern is very different from that of Chinese–English speakers. First, English speakers did not rate the vertical gesturing differently for two types of time references whereas Chinese–English speakers rated the vertical gesturing for English translations of vertical time reference with a higher score than that for the English translations of neutral time references. This indicates that Chinese–English speakers' vertical tendency in time references with vertical spatial metaphors was not due to some general non-linguistic factors, but to the activation of L1 Chinese lexicons in the L2.

Second, for the English translations of Chinese neutral time reference, English speakers preferred the lateral axis to vertical axis whereas Chinese–English speakers rated the vertical axis the same as the lateral axis. Presumably, this discrepancy is due to their differences in time conceptualisations, in that Chinese speakers can think of time vertically (Boroditsky, 2001).

### General discussion

In experiment 1 we observed that Chinese–English bilinguals produced vertical temporal gestures, both in Chinese and in English. The between-language and within-language comparisons showed that Chinese–English bilinguals produced more vertical gestures when talking about Chinese time references with vertical spatial metaphors than (1) when talking about time conceptions in the English translations, and than (2) in the case of Chinese time references with no spatial metaphors. In Experiment 2, we showed that Chinese–English bilinguals preferred vertical gestures to lateral gestures when perceiving time references with vertical spatial metaphors. This bias towards vertical gestures still existed when they perceived them in English, though to a lesser extent. Nevertheless, there was no such bias towards vertical gestures when perceiving time references with no spatial metaphors. Results of Experiment 3 showed that English speakers had a different preference of temporal gestures from the Chinese–English bilinguals, and also indicated that the vertical tendency in Chinese–English bilinguals was not due to some general non-linguistic factors. We discuss these results in terms of production and perception of gestures, respectively.

With respect to gesture production, the findings suggest that the production of vertical gestures by Chinese–English bilinguals can be influenced by both

the habitual vertical conceptualisation of time and the online lexical form of the expressions they produce when speaking. On the one hand, participants' generation of vertical gestures for neutral time references in both Chinese (even without using vertical wording) and English indicate that they can employ a vertical conceptualisation of time (Boroditsky, 2001). As the number of vertical gestures for these neutral time references was not different between the two languages, the gesture production pattern implied that the vertical gesturing was likely to be influenced by a Chinese vertical time conceptualisation.

On the other hand, the extent to which vertical gestures were produced also depended on the linguistic choices. Firstly, participants produced more vertical gestures for time references with Chinese vertical spatial metaphors than for the English translations of Chinese vertical spatial metaphors. One may argue that the difference was due to the effect of the addressee, as speakers may take the perspective of the addressee and were aware that vertical gestures were not helpful when the addressee was English. If so, we would also find less vertical gesturing in the case of time references with neutral words, but there was no significant difference between the two languages in the number of vertical gestures produced with neutral words.

It could be argued that in Chinese, vertical spatial language is more strongly associated with time units such as weeks than with days (Chen, 2007), and thus wordlists could potentially confound the verticality of the stimulus with the time unit treated in the stimulus. However, the effect that we found in the production experiment is unlikely due to this, because, first, the between languages comparison of wordlists with vertical spatial metaphors showed that even for the same time units (e.g. last week), the number of vertical gestures was significantly higher in Chinese than in the English translations. Second, participants performed vertical gestures when using vertical wording to explain the vertical wordlists, and sometimes they also explained the same conceptions with non-vertical wording (e.g. "The expressions are *last week* and *next week*. So they are about two weeks, namely the *previous week* and the *following week/seven days ago*, and *seven days later*. Etc"). We further checked the temporal gestures of vertical wordlists which accompanied these cases of non-vertical wording. The proportion of vertical gesturing was significantly higher (90%) when accompanying vertical wording than when accompanying non-vertical wording (20%) (McNemar test,  $p = .039$ ). This indicates that even for identical time conceptions within a particular wordlist in the same language, verbally producing a vertical spatial metaphor for a time reference

immediately led to an increase in the production of vertical gestures. This piece of evidence, together with the findings from the within and between languages comparisons of the two types of time references provides strong evidence to support the fact that linguistic choices of vertical wording can also have an influence on participants' production of vertical gesturing.

Another alternative explanation for our findings could be related to the claim that the greater number of vertical gestures in Chinese context was superficial, because speakers simply produced vertical gestures so that the addressees could figure out the vertical spatial words to be reported in a led-to-believe end-of-game test. This is unlikely. First, the task was not a word-guessing game and participants were not instructed to use gestures at all. Second, the lexical form of "shàng" (above) alone does not necessarily result in many vertical gestures. We looked at a filler (shàng bān, on duty, xià bān, off duty) which also contained the identical morphemes of "shàng" and "xià". However, when describing the filler, participants far less often performed vertical gestures compared to that of "shàng zhōu" (last week) and "xià zhōu" (next week).<sup>6</sup> This fact also indicates that the vertical gesturing in time references were unlikely due to a matter of preference on the part of the participants or by a priming of the trials.

Our results are consistent with previous studies on forced gesturing about time. When participants were asked explicitly to point to space for time conceptions, Chinese speakers were more likely to point vertically for time conceptions with vertical spatial metaphors than for those with non-vertical spatial metaphors, or without any spatial metaphor (Fuhrman et al., 2011; Lai & Boroditsky, 2013). Furthermore, by comparing the temporal gestures for the same time conceptions produced by Chinese-English bilinguals, we provide new evidence for the effect of cross-linguistic choices on gesture production. These results are in line with a model that states that gestures are not only the result of a pre-determined language-specific conceptual scheme (vertical conceptualisations of time) (Boroditsky, 2001), but are also shaped by the interface between spatio-motoric thinking (imagistic representation) and speaking, in which spatial imagery is adjusted to fit the verbalisation (Kita & Özyürek, 2003). The production study provides unique new evidence supporting the Interface Hypothesis (Kita, 2000; Kita & Özyürek, 2003), and further contributes to the theory by suggesting that gesture production is dynamic and sensitive to linguistic encoding possibilities, even for abstract concepts and in a bilingual context.

With respect to the perception experiments, we found that the interpretation of temporal gestures was not only

influenced by general time conceptualisations but also by linguistic choices. First, for neutral time references, English speakers preferred lateral gestures whereas Chinese–English speakers also accepted vertical gestures. The differences are likely due to the fact that Chinese speakers can think of time vertically (Boroditsky, 2001). Secondly, in comparison to the lateral gesture plane, for Chinese–English speakers, the vertical gesture plane was preferred for time references with vertical spatial metaphors, but not for time references with neutral words. Since the extent to which English speakers preferred the vertical gestures was not different in both types of wordlists, the differences in the Chinese–English group were not caused by some general non-linguistic factors. Thus the perception of temporal gestures can also be affected by vertical spatial metaphors, in that sense being evidence that supports the findings of the production experiment.

Perhaps the most striking observation is the finding that Chinese–English speakers still preferred vertical gestures to lateral gestures for the English translations of time conceptions with vertical spatial metaphors. Since we did not find a correlation between the vertical tendency and L2 English proficiency, it may suggest that languages are co-activated in a bilingual speaker, even if only one language is contextually relevant (e.g. Brown & Gullberg, 2008; Wu & Thierry, 2010). In this case, the Chinese vertical spatial metaphors are also activated in a Chinese–English bilingual's mind, even in the English context. Future studies can focus on English learners of Chinese to find out whether this is due to an L1 mediated translations: if English speakers' vertical mapping in L2 Chinese is still activated in their L1 English, then it is likely that the vertical mapping is due to the activation of a stable representation of that mapping, since it is quite unlikely that English speakers translate the Chinese words into English when they are tested in the English mother tongue.

Interestingly, comparing the production and the perception studies on Chinese–English bilinguals, we saw that, although we used both implicit and explicit approaches to investigate gestures, the pattern of results from the production and the perception experiments show great similarity, especially in L1 Chinese. This parallel between the production and perception data seem to suggest that in addition to the stable conceptualisation, the online linguistic encoding possibilities of time conceptions also have an influence on the perception of temporal gestures.

Admittedly, in both experiments there was a linguistic input to which the participants had to react. Participants may be primed from the task instruction (vertical/non-vertical time wordlists), such that the pattern of results

may come from the implicit priming of the spatial morpheme (e.g. “above”) in the target wordlist (e.g. “above week”). In fact, participants often produced more than one temporal gesture when explaining a target wordlist in the production experiment. Within each Chinese wordlist, we further coded co-occurring spoken words accompanied with each temporal gesture (binary coding of co-occurring words: with/without vertical lexicon). In total, we obtained 588 lexicon-gesture tokens, which included 148 vertical words and 440 non-vertical words, and 101 vertical gestures and 487 non-vertical gestures. There was a significant effect of vertical lexicon on the vertical gesturing ( $\beta = 2.36$ , Wald  $\chi^2 = 32.15$ ,  $df = 1$ ,  $p < .0001$ ), even after controlling for wordlist-type and other factors. This indicated that within a wordlist (regardless of neutral or vertical priming), participants were more likely to produce vertical gestures when their co-occurring words were vertical lexicons. Furthermore, among the vertical gestures, 50 vertical gestures were produced with vertical lexicons whereas 51 vertical gestures were produced with non-vertical words. The fact that half of the vertical gestures were produced with non-vertical temporal expressions indicated that Chinese people can conceptualise time vertically, even when the wording did not have vertical spatial metaphors. Together, these additional analyses show that the pattern of results is not simply due to a priming effect,<sup>7</sup> but a consequence of both lexical and conceptual effects.

The population of Chinese–English bilinguals in Experiment 1 (production) was different from the population in Experiment 2 (perception) in terms of context in which they were tested (the Netherlands vs. China). One possible concern is that the Chinese–English bilinguals for the production experiment had been away from China for a certain period of time (Mean = 20 months) and might have been immersed in an “English-taught international programme”. Yet, in this study we did not find strong evidence supporting this concern, as the length of staying in the Netherlands was not found to be a significant factor ( $\beta = -0.068$ , Wald  $\chi^2 = .14$ ,  $df = 1$ ,  $p = .71$ ).

Additionally, the similar pattern of results exhibited in the production and perception studies not only suggests the effect of linguistic choice on the production and perception of co-speech gestures, but also provides new evidence for the interconnection between production and perception (Pickering & Garrod, 2013) from the perspective of gestures.

## Conclusions

In the present study we investigated whether and why Chinese–English bilinguals produce vertical gestures

about time and how different factors interact with each other in the bilingual language processing. We addressed these questions by investigating Chinese–English bilingual speakers’ production and perception of gestures for temporal expressions. The findings of this study contribute to theories accounting for the speech–gesture relationship, bilingual mental lexicons and embodied cognition. First, our production results support the claim that gestures are not only shaped by the language-specific conceptual schema, but also by the linguistic encoding possibilities (Kita & Özyürek, 2003). Moreover, we extend the Interface model, which was proposed and tested predominantly based on motion events, to abstract concepts such as time, and to a bilingual context. Second, we are the first to propose that the linguistic encoding possibilities influence the perception of gestures as well. Future studies can test this hypothesis by use of a more implicit approach. Furthermore, our gesture data provide evidence for the view that languages are co-activated in a bilingual (e.g. Brown & Gullberg, 2008; Wu & Thierry, 2010). Finally, this study also provides insight into Chinese speakers’ implicit and explicit understanding of time. If gestures are a visible embodiment of cognition (Hostetter & Alibali, 2008), then the production of vertical temporal gestures for time by Chinese–English bilinguals offers empirical support for the idea that Chinese speakers can employ a vertical conceptualisation of time (Boroditsky, 2001). To further explore this idea, future work can study different samples of Chinese speakers. For instance, we can study the temporal gestures by learners of Chinese as a second language to see whether their temporal gestures change after learning Chinese; we can also investigate the spatio-temporal reasoning of Chinese deaf signers, an atypical population of Chinese speakers *in* the Chinese culture, who differ from Mandarin speakers in spatial metaphors for time linguistically (Gu, Zheng, & Swerts, 2017). This sample may provide a unique opportunity to study the effect of linguistic force on spatial-temporal thinking within the culture. Additionally, these studies can examine to what extent the cross-linguistic differences in mental lexicons cause the differences in conceptualisation.

## Notes

1. Left-right spatio-temporal metaphors are actually absent from English speech. This lateral time axis is likely due to the reading/writing direction (see discussions in e.g. Casasanto & Jasmin, 2012). In Chinese, when talking about time in the lateral axis, “left” (左, zuǒ) and “right” (右, yòu) are only used together following a specific time point. It refers to “being earlier/later than a certain time point (around that time)”, e.g. “around one

o’clock” can be said as “一点左右”, *yī-diǎn zuǒ-yòu*, which literally means “one o’clock left right”.

2. Another account that can contribute to Chinese speakers’ vertical conceptualisation of time is the vertical writing direction in the old days (e.g. Chen, 2007; Fuhrman et al., 2011). This factor is also addressed in the analyses section.
3. When wordlists with “others” were excluded in the analysis, the results remained similar.
4. This model is equivalent to the generalised linear-mixed model (using a logistic link function and the probability density function for the logistic) for the binary outcome. As robustness checks, both a random effects count data regression and a 2x2 repeated measures ANOVA yield the same effects, when the dependent variable was coded as the number of vertical gestures for each wordlist-type.
5. The counter-balance design was still achieved in the remaining 79 participants (there were 39 participants for version A, in which 19 participants first did the Chinese task and the other 20 first did the English task; there were 40 participants for version B, in which 18 participants first did the Chinese task and the other 22 first did the English task).
6. One possible explanation is that “shàng zhōu” (last week) and “shàng bān” (on duty) have different levels of semantic transparency. Such a difference is essentially associated with the different word conceptions. Regarding Chinese morphological processing, different meanings of an ambiguous morpheme are activated during word recognition (e.g. Tsang & Chen, 2013; Tsang, Wong, Huang, & Chen, 2014). In the context of “shàng zhōu” (last week), the morpheme form “shàng” activates the lemma “above” more than that in the context of “shàng bān” (on duty), according to their corresponding context-consistent meaning. Therefore, given the same lexicon form of “shàng”, the ultimate activation of vertical spatial metaphors plays a role in shaping vertical gestures.
7. If this priming also took place when the Chinese participants produced gestures for the English counterpart, we would expect that Chinese participants with lower English proficiency were more likely to translate the neutral English word into Chinese, being more primed to produce/prefer vertical gestures than Chinese participants with higher English proficiency. However, we did not find any evidence for this, as the variable of English proficiency was not significant in both production and perception experiments. Additionally, according to a similar previous study on English speakers’ time gestures, an implicit priming of the spatial morpheme alone (e.g. “ahead”) is insufficient to elicit systematically congruent temporal gestures (Casasanto & Jasmin, 2012), because temporal gestures also reveal an implicit spatial conceptualization of time that may not be inferred from the language.

## Acknowledgements

We are very grateful to Fey Parrill, Bastien Boutonnet and all the anonymous reviewers for their constructive comments on earlier versions of the manuscript. We thank Yequi Zheng for

her assistance in statistics and useful comments. We also thank Hua Chen, Yuwen Wang, Yu Gu, Nanne van Noord, Hans Westerbeek and others who have assisted us in running experiments. Part of the work of this paper has been presented at TiGeR 2013, Chinese Linguistics in the Netherlands 2014, ISGS 6, LOT 2014 Nijmegen Summer School, and CogSci 2014 conferences.

### Disclosure statement

No potential conflict of interest was reported by the authors.

### Funding

The first author received financial support from The Netherlands Organization for Scientific Research, via NWO Promoties in de geesteswetenschappen [322-89-007], which is gratefully acknowledged.

### References

- Bender, A., & Beller, S. (2014). Mapping spatial frames of reference onto time: A review of theoretical accounts and empirical findings. *Cognition*, 132(3), 342–382. doi:10.1016/j.cognition.2014.03.016
- Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, 75(1), 1–28. doi:10.1016/S0010-0277(99)00073-6
- Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, 43(1), 1–22. doi:10.1006/cogp.2001.0748
- Boroditsky, L., & Gaby, A. (2010). Remembrances of times east: Absolute spatial representations of time in an Australian aboriginal community. *Psychological Science*, 21(11), 1635–1639. doi:10.1177/0956797610386621
- Brown, A., & Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1-L2 encoding of manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition*, 30(2), 225–251. doi:10.1017/S0272263108080327
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6(1), 3–5. doi:10.1037/e527772014-223
- Calbris, G. (2008). From left to right: Coverbal gestures and their symbolic use of space. In A. Cienki & C. Müller (Eds.), *Metaphor and gesture* (pp. 27–53). Amsterdam: John Benjamins.
- Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition*, 106(2), 579–593. doi:10.1016/j.cognition.2007.03.004
- Casasanto, D., & Bottini, R. (2014). Mirror reading can reverse the flow of time. *Journal of Experimental Psychology: General*, 143(2), 473–479. doi:10.1037/a0033297
- Casasanto, D., & Jasmin, K. (2012). The hands of time: Temporal gestures in English speakers. *Cognitive Linguistics*, 23(4), 643–674. doi:10.1515/cog-2012-0020
- Chen, J. Y. (2007). Do Chinese and English speakers think about time differently? Failure of replicating Boroditsky (2001). *Cognition*, 104(2), 427–436. doi:10.1016/j.cognition.2006.09.012
- Chui, K. (2011). Conceptual metaphors in gesture. *Cognitive Linguistics*, 22(3), 437–458. doi:10.1515/cogl.2011.017
- Cienki, A. (1998). Metaphoric gestures and some of their relations to verbal metaphorical expressions. In J. Koenig (Ed.), *Discourse and cognition: Bridging the gap* (pp. 189–204). Stanford, CA: CSLI Productions.
- Clark, H. (1973). Space, time, semantics, and the child. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 27–63). New York, NY: Academic Press.
- Cooperrider, K., & Núñez, R. (2009). Across time, across the body: Transversal temporal gestures. *Gesture*, 9(2), 181–206. doi:10.1075/gest.9.2.02coo
- Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS One*, 8(3), e57410. doi:10.1371/journal.pone.0057410
- de la Fuente, J., Santiago, J., Román, A., Dumitrache, C., & Casasanto, D. (2014). When you think about it, your past is in front of you: How culture shapes spatial conceptions of time. *Psychological Science*, 25(9), 1682–1690. doi:10.1177/0956797614534695
- Fuhrman, O., & Boroditsky, L. (2010). Cross-cultural differences in mental representations of time: Evidence from an implicit nonlinguistic task. *Cognitive Science*, 34(8), 1430–1451. doi:10.1111/j.1551-6709.2010.01105.x
- Fuhrman, O., McCormick, K., Chen, E., Jiang, H. D., Shu, D. F., Mao, S. M., & Boroditsky, L. (2011). How linguistic and cultural forces shape conceptions of time: English and Mandarin time in 3D. *Cognitive Science*, 35(7), 1305–1328. doi:10.1111/j.1551-6709.2011.01193.x
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9(3), 558–565. doi:10.3758/bf03196313
- Gu, Y., Zheng, Y., & Swerts, M. (2017). Does Mandarin Spatial metaphor for time influence Chinese deaf signers' spatio-temporal reasoning? Manuscript submitted for publication.
- Gullberg, M. (1998). *Gesture as a communication strategy in second language discourse: A study of learners of French and Swedish* [Travaux de l'institut de linguistique de Lund], 35. Linguistics and Phonetics, (Dissertation). doi:10.1017/S0272263100271052
- Hendricks, R. K., & Boroditsky, L. (2015). New space-time metaphors foster new mental representations of time. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th annual cognitive science society* (pp. 902–907). Austin, TX: Cognitive Science Society.
- Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review*, 15(3), 495–514. doi:10.3758/pbr.15.3.495
- Kellerman, E., & Van Hoof, A.-M. (2003). Manual accents. *IRAL – International Review of Applied Linguistics in Language Teaching*, 41(3), 251–269. doi:10.1515/iral.2003.011
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press. doi:10.1017/cbo9780511807572
- Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162–185). Cambridge, MA: Cambridge University Press. doi:10.1017/cbo9780511620850.011
- Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and Cognitive Processes*, 24(2), 145–167. doi:10.1080/01690960802586188

- Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48(1), 16–32. doi:10.1016/s0749-596x(02)00505-3
- Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2007). Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes*, 22(8), 1212–1236. doi:10.1080/01690960701461426
- Lai, V. T., & Boroditsky, L. (2013). The immediate and chronic influence of spatio-temporal metaphors on the mental representations of time in English, Mandarin, and Mandarin-English speakers. *Frontiers in Psychology*, 4, 1–10. doi:10.3389/fpsyg.2013.00142
- Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, 41(3), 841–849. doi:10.3758/brm.41.3.841
- Le Guen, O., & Balam, L. I. P. (2012). No metaphorical timeline in gesture and cognition among Yucatec Mayas. *Frontiers in Psychology*, 3, doi:10.3389/fpsyg.2012.00271
- Mashal, N., Andric, M., & Small, S. (2011). Motor and nonmotor language representations in the brain. In M. Faust (Ed.), *The handbook of the neuropsychology of language* (pp. 276–293). Oxford: Wiley-Blackwell. doi:10.1002/9781118432501.ch14
- Negueruela, E., Lantolf, J. P., Jordan, S. R., & Gelabert, J. (2004). The “private function” of gesture in second language speaking activity: A study of motion verbs and gesturing in English and Spanish. *International Journal of Applied Linguistics*, 14(1), 113–147. doi:10.1111/j.1473-4192.2004.00056.x
- Núñez, R., & Cooperrider, K. (2013). The tangle of space and time in human cognition. *Trends in Cognitive Sciences*, 17(5), 220–229. doi:10.1016/j.tics.2013.03.008
- Núñez, R., Cooperrider, K., Doan, D., & Wassmann, J. (2012). Contours of time: Topographic construals of past, present, and future in the Yupno Valley of Papua New Guinea. *Cognition*, 124(1), 25–35. doi:10.1016/j.cognition.2012.03.007
- Núñez, R. E., & Sweetser, E. (2006). With the future behind them: Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science*, 30(3), 401–450. doi:10.1207/s15516709cog0000\_62
- Özçalışkan, Ş. (2016). Do gestures follow speech in bilinguals’ description of motion? *Bilingualism: Language and Cognition*, 19(3), 644–653. doi:10.1017/s1366728915000796
- Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S. (2016). Does language shape silent gesture? *Cognition*, 148, 10–18. doi:10.1016/j.cognition.2015.12.001
- Özyürek, A. (2002). Speech-gesture relationship across languages and in second language learners: Implications for spatial thinking and speaking. In B. Skarabela, S. Fish, & A. H. Do (Eds.), *Proceedings of the 26th annual conference on language development* (pp. 500–509). Somerville, MA: Cascadilla Press.
- Özyürek, A., Kita, S., Allen, S. E., Furman, R., & Brown, A. (2005). How does linguistic framing of events influence co-speech gestures? Insights from crosslinguistic variations and similarities. *Gesture*, 5(1/2), 219–240. doi:10.1075/gest.5.1-2.15ozy
- Özyürek, A., Willems, R., Kita, S., & Hagoort, P. (2007). On-line integration of semantic information from speech and gesture: Insights from event-related brain potentials. *Journal of Cognitive Neuroscience*, 19(4), 605–616. doi:10.1162/jocn.2007.19.4.605
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108(1), 155–184. doi:10.1016/j.cognition.2008.02.007
- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36(4), 329–347. doi:10.1017/s0140525x12001495
- Saj, A., Fuhrman, O., Vuilleumier, P., & Boroditsky, L. (2014). Patients with left spatial neglect also neglect the “left side” of time. *Psychological Science*, 25(1), 207–214. doi:10.1177/0956797612475222
- Santiago, J., Lupáñez, J., Pérez, E., & Funes, M. J. (2007). Time (also) flies from left to right. *Psychonomic Bulletin & Review*, 14(3), 512–516. doi:10.3758/bf03194099
- Slobin, D. I. (1987). Thinking for speaking. In J. Aske, N. Beery, L. Michaelis, & H. Filip (Eds.), *Proceedings of the 13th annual meeting of the Berkeley Linguistics Society* (pp. 435–445). Berkeley, CA: Berkeley Linguistic Society.
- Stam, G. (2006). Thinking for speaking about motion: L1 and L2 speech and gesture. *IRAL – International Review of Applied Linguistics in Language Teaching*, 44(2), 145–171. doi:10.1515/iral.2006.006
- Tsang, Y. K., & Chen, H.-C. (2013). Early morphological processing is sensitive to morphemic meanings: Evidence from processing ambiguous morphemes. *Journal of Memory and Language*, 68(3), 223–239. doi:10.1016/j.jml.2012.11.003
- Tsang, Y. K., Wong, A. W., Huang, J., & Chen, H. (2014). Morpho-orthographic and morpho-semantic processing in word recognition and production: Evidence from ambiguous morphemes. *Language, Cognition and Neuroscience*, 29(5), 543–560. doi:10.1080/01690965.2013.790554
- UCLES. (2001). *Quick placement test*. Oxford: Oxford University Press and University of Cambridge Local Examinations Syndicate.
- Wu, Y. J., & Thierry, G. (2010). Chinese–English bilinguals reading English hear Chinese. *Journal of Neuroscience*, 30(22), 7646–7651. doi:10.1523/jneurosci.1602-10.2010