Does Language Shape the Production and Perception of Gestures?
A Study on late Chinese-English Bilinguals' Conceptions about Time

Yan Gu (yan.gu@tilburguniversity.edu)
Lisette Mol (l.mol@tilburguniversity.edu)
Marieke Hoetjes (m.w.hoetjes@tilburguniversity.edu)
Marc Swerts (m.g.j.swerts@tilburguniversity.edu)
Tilburg center for Cognition and Communication (TiCC), Tilburg University, the Netherlands

Abstract

Does language influence the production and perception of gestures? The metaphorical use of language in representing time is deeply interlinked with actions in space, such as gestures. In Chinese, speakers can talk and gesture about time as if it were horizontal, sagittal, or *vertical*. In English, speakers rarely employ the vertical plane. Two experiments showed that the verbal use of vertical spatial metaphors had an online influence on the production and perception of gestures by late Chinese-English bilinguals. Participants produced more vertical gestures when talking about time references by use of vertical spatial metaphors, e.g. 'shàng-zhōu' (literally: 'above week', meaning 'last week'), and they preferred vertical gestures to horizontal gestures when perceiving time references with vertical spatial metaphors. Gestures are not only shaped by the language specific conceptualisation, but are also sensitive to the changes in linguistic choices, both in production and perception.

Keywords: gesture; space; time; metaphor; Chinese

Introduction

When representing abstract conceptions, people often recruit metaphors from more concrete domains, for instance, 'time is space' (Casasanto & Boroditsky, 2008). In all cultures people use space to represent time (Boroditsky, 2000). For example, English speakers can use spatial metaphors to talk about time such as 'The future lies not too *far ahead*', and 'The day has been *long*'. In English, as well as in many other languages, speakers often conceptualise the past at the back（or leftwards), and the future at the front (or rightwards) (Calbris, 2008; Clark, 1973). However, this is not the case for all languages: for example, speakers of Aymara have the concept that 'The future is *behind*' (Núñez & Sweetser, 2006). The spatiotemporal thought can also be expressed in speakers' co-speech metaphoric gestures. Typically, English speakers produce horizontal and sagittal gestures to indicate timelines (Casasanto & Jasmin, 2012).

Interestingly, Chinese speakers can *additionally* talk about time vertically by employing vertical spatial metaphors of '上' (shàng: above) and '下' (xià: below) to indicate the time conceptions of early and late. For example, '上周' (shàng zhōu) can literally be translated as 'above week', which means 'last week', while '下下周' (xià xià zhōu) as 'below below week', referring to 'the week after next week'. Due to the differences in use of spatial metaphors in time conceptions, Boroditsky (2001) argues that Chinese speakers may have a different conceptualisation of time than English speakers. Her argument is based on Slobin's (1987) 'thinking-for-speaking' hypothesis, which proposes that speakers' habitual thinking in a language will influence their cognitive attention and speech production - that is, the choices of what to focus on and how to say it. Specifically in Chinese, Boroditsky (2001) believes that the use of vertical spatial metaphors to talk about time shapes Chinese speakers' language-specific conceptual schema.

Interestingly, Chinese speakers can also gesture about time vertically. For instance, they can point upward to indicate the concept of an earlier time such as 'last week' or point downward to indicate that of a later time such as 'next week' (Chui, 2011; Gu, Mol, Hoetjes, & Swerts, 2013). One possible explanation to account for such a phenomenon could be embodiment, which proposes that conceptual representations are largely grounded in a sensorimotor medium motor, and representations are activated and often instantiated in the forms of gestures. However, given the language speakers are Chinese, one cannot preclude that Chinese lexical representations are activated which in turn influence the gestural patterns. Then the question arises whether the vertical gesturing about time results from Chinese speakers having different time conceptualisations in general (Boroditsky, 2001), or whether there is an additional online effect of how time is linguistically encoded in Chinese. In other words, the vertical gesturing about time conceptions can be due to (a) Chinese speakers' vertical thinking of time (a language-specific conceptual schema); (b) the linguistic choices of the vertical words which drive the gestures; or (c) both the conceptual schema and the current linguistic choices. If the first explanation holds, then Chinese speakers will also gesture vertically when speaking English. That means their vertical gesturing will be unaffected by the language itself. However, if the vertical gesturing is not merely caused by the general conceptualisation of time, then it is likely to be immediately affected by the language.

According to Kita & Özyürek's (2003) Interface Hypothesis, gestures not only reflect imagistic (spatio-motoric) representations of events, but also aid thinking-for-speaking. In their production model, the generation of a gesture is modulated by two forces: the spatio-motoric thinking from the Working Memory, and the linguistic encoding possibilities from the Message Generator, both of which interact with each other. The ultimate content of a gesture is determined by the Action Generator, which takes into account the two forms of thinking, such that gestures are adjusted to fit the verbalisation.

For instance, Kita & Özyürek (2003) found that a scene of a 'Rolling Event' can be expressed as 'rolling down' in English, with manner and path conflated into a single clause. Therefore, this information tends to be conflated in

gesture too, with one gesture expressing both manner and path. However, in Turkish and Japanese, the manner and path have to be expressed separately in two clauses (i.e., 'move down, in a rolling fashion'), and consequently tend to be accompanied by two separate gestures. This suggests that gestures are coordinated to speech online, and the linguistic encoding possibilities influence gesture production *during formulation* (see hypothesis b above).

This cross-linguistic difference in the description of events also influences one's perception. For instance, as revealed by an eye-tracking study, when participants were instructed to watch a motion event in the intent of later description, native speakers of Greek (a verb-framed language similar to Turkish) were significantly more likely to focus on path over manner than native speakers of English. However, there were no such differences when they were simply told to watch the clip. The results showed that the verbal task has an effect on the gazing of motion and path, and language can guide event perception (Papafragou, Hulbert, & Trueswell, 2008).

So far, evidence supporting the Interface Hypothesis predominantly has come from studies of motion event descriptions and from monolingual speakers (e.g., Kita, Özyürek, Allen, Brown, Furman, & Ishizuka, 2007; Özyürek, Kita, Allen, Furman, & Brown, 2005). It is found that languages are co-activated in a bilingual speaker (Wu & Thierry, 2010). In order to find out what shapes Chinese speakers' vertical gesturing about time, and also to test the model further, we study the abstract concept of time in a bilingual context. Firstly, as time in Chinese can be expressed by use of vertical spatial metaphors or no spatial metaphor, and secondly time conceptions with Chinese vertical spatial metaphors are encoded linguistically differently in English, the Interface Hypothesis will predict that: (1) the gesture production during the verbal use of Chinese vertical spatial metaphors will be different from that during the verbal use of neutral words; (2) the production of gestures in Chinese will be different from that of the paraphrased translation in English. However, if Chinese-English bilinguals apply the same time conceptualisations in L1 and L2, these differences are not predicted.

Moreover, an increasing number of studies show that the production and the perception of speech and gesture are interconnected. For instance, perceiving gestures automatically activates the brain areas involved in producing these corresponding actions (cf. Mashal, Andric, & Small, 2012; Hostetter & Alibali, 2008). Additionally, studies have shown that there is an on-line integration of semantic information from speech and gesture (e.g., Özyürek, Willems, Kita, & Hagoort, 2007). Therefore, it is legitimate to test whether the *perception* of vertical gestures about time by the Chinese-English bilinguals is also influenced by the linguistic encoding possibilities besides the influence from the pre-determined conceptual schema.

In the current study we aim to explore whether the verbal expression of time conceptions with spatial metaphors affects Chinese-English bilinguals' production and perception of gestures. In two studies on bilinguals' vertical conceptualisation of time, we demonstrate that language influences both production and perception.

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 'yesterday', and 'tomorrow' are spatial expressions with reference to time)? (2) Given the fact that English and Chinese speakers may think of time differently (Boroditsky, 2001), will late Chinese-English bilinguals produce more vertical gestures in Chinese than in English? (Bilinguals are defined as *sequential bilinguals*, who first acquired Chinese as the L1 and then English as the L2.)

Method

Participants

46 late Chinese-English bilinguals (L1 = Chinese; 35 F and 11 M; mean age = 24 yrs) were paid for participation. They were students from Tilburg University (Netherlands), who originally came from China. Their English proficiency was between intermediate and advanced.

Stimuli

Eleven Chinese wordlists were constructed for a word definition task, four of which were relevant to the current study. They consisted of words that conveyed time conceptions (see Table 1). Time references (1) and (2) contained words with vertical spatial metaphors ('上'/shàng, above, and '下'/xià, below) to indicate the time conceptions of 'early' and 'late'. By contrast, words in (3) and (4) did not contain explicit lexical references to vertical ('shàng' or 'xià'), or sagittal ('前'/qián, front or '后'/hòu, back) space dimensions. The English version was a translation of Chinese. Note that wordlists (1) and (2) were translated into English to express the same concepts but with neutral words. The reason for this is twofold: first, English has few comparable vertical spatial metaphors for time conceptions; secondly, we can examine whether there is a trace of the L1 vertical thinking of time in the L2.

Table 1: Wordlists of targeted time referents.

	Chinese	English
(1)	上周, 下周 shàng zhōu, xià zhōu	last week, next week
(2)	上辈子, 下辈子 shàng bèi zi, xià bèi zi	previous life, next life
(3)	昨天,今天, 明天 zuó tiān, jīn tiān, míng tiān	yesterday, today , tomorrow
(4)	早晨, 晌午, 傍晚, 深夜 zǎo chén, shǎng wǔ, bàng wǎn, shēn yè	morning, noon, evening, late at night

Procedure

Each participant was paired with an addressee and tested individually. The experiment was ostensibly set up as a language and memory experiment, 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. There were wordlists in the centre of the monitor that were only seen by the speaker. The speaker was informed that he/she would have to remember the wordlists shortly after having seen them twice, find the relationship between the words within a list and explain the definitions of the words in the list to the addressee as logically as possible. The addressee was told to remember the speaker's descriptions as much as possible for the 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 testing order of the languages was counterbalanced and the interval between the two tests was approximately ten days. The addressees were native speakers of Chinese for the Chinese task, and English-Dutch bilinguals for the English task. The entire experiment was videotaped with participants' consent. Participants were not instructed to perform gestures in the task. A post-experiment questionnaire showed that participants were not aware that the purpose of the study was to investigate speakers' gestures.

Coding and Measurements

The temporal gestures (gestures performed to represent time conceptions) accompanying the speech that described the target wordlists were annotated in ELAN. The planes of gestures were categorised as vertical, horizontal, or sagittal, and the directionality of each plane was also indicated. A participant could explain a wordlist by using temporal gestures from several planes, which were coded in each plane accordingly. When a participant produced repetitive temporal gestures within a plane for a specific word or words in the wordlist, this gesturing was only counted once. For each participant, the number of vertical gestures produced for each of the four wordlists was counted. Thus the wordlist-type (vertical and neutral) and the language factors (Chinese and English) were independent variables, and the number of vertical gestures was the dependent variable.

Four participants did not produce any gesture in Chinese or in English for the four target wordlists and one participant did not finish the experiment. Data from these five people were excluded from the analysis.

Results and Discussion

Participants produced a total of 328 (41×4×2) target wordlists for both languages, 269 (82.0%) of which were accompanied by temporal gestures (128 (78.0%) for Chinese and 140 (85.4%) for English) (see Figure 1).

Among the four wordlists for each participant, the average number of temporal gestures in English (M = 3.51, SD = .71) was significantly higher than that in Chinese (M = 3.14, SD = 1.06), t (40) = 1.95, p = .058, two tailed, r = .30. This is to be expected, because late bilingual speakers usually gesture more often in their second language than in their first language (Gullberg, 1998).

A 2×2 repeated measures ANOVA with wordlist-type (vertical and neutral) and language (Chinese and English) as within subject variables showed that there was a main effect of wordlist-type on the number of vertical gestures, F (1, 40) = 14.30, p < .001, η_p^2 = .26, but no main effect of language, F (1, 40) = 1.34, p = .18, η_p^2 = .03. There was a significant interaction between language and wordlist-type (vertical and neutral), F (1, 40) = 10.94, p = .002, η_p^2 = .22.

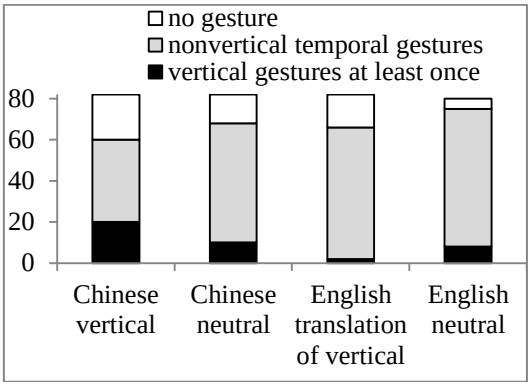

Figure 1: The number of wordlists that were accompanied by gestures for the four types of time references (82 for each type). The non-vertical temporal gestures included gestures in both horizontal and sagittal planes.

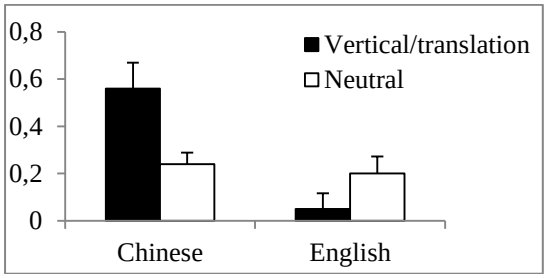

Figure 2: The number of vertical gestures for vertical and neutral wordlists. Note in English vertical wordlists were a translation of Chinese time concepts in neutral words.

Follow-up *post hoc* analyses (with Bonferroni α level corrected to 0.025) revealed that the Chinese wordlists with vertical spatial metaphors (M = .56, SD = .71) were accompanied by a significantly higher number of vertical gestures than the neutral Chinese words (M = .24, SD = .43), t (40) = 2.81, p = .008, two tailed, r = .41, 95% CI = (.09, .55) (see Figure 2). This indicated that speaking time references with vertical spatial metaphors has 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, t (40) = 1.64, p = .11, two tailed, r = .06, 95% CI = (-.03, .32). This suggests that the production of vertical temporal gestures is sensitive to the linguistic choices.

As for the comparisons between languages, firstly, the number of vertical gestures for wordlists with vertical spatial metaphors was significantly higher in Chinese (M = .56, SD = .71) than in the English translation (M = .05, SD = .31), t (40) = 4.40, p < .001, r = .57, 95% CI = (.28, .75). Figure 3 shows a participant producing vertical gestures in Chinese and horizontal ones in English.

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 should expect the same pattern in the case of wordlists with neutral words. However, for the neutral wordlists, there was no significant difference in the number of accompanying vertical gestures between the two languages (Chinese: M = .24, SD = .43; English M = .20, SD = .46), t (40) = .57, p = .57, two tailed, r = .09, 95% CI = (-.12, .22).

Why did the temporal gestures of the two types of time references (vertical and neutral) display such differences across languages? It is likely that apart from the thinking from spatio-motoric cognition, there is another force – the force from the linguistic encoding possibilities, resulting in the increase of producing vertical gestures for the time references with vertical spatial metaphors in Chinese.

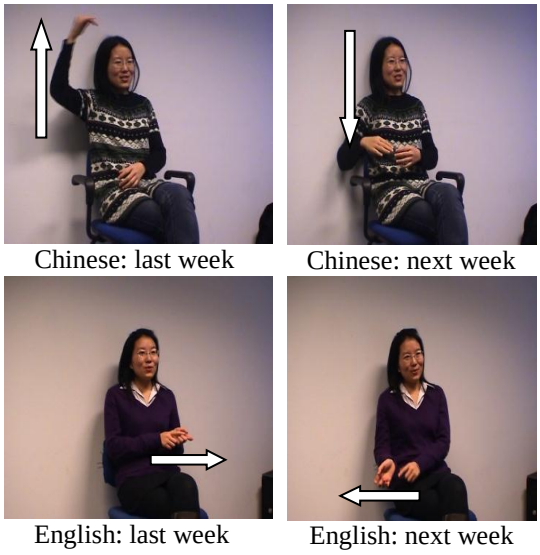

Figure 3: Gestures of *last week* and *next week* by a participant in Chinese and in English.

Perception Experiment

In the perception experiment, 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 words which do not have such an explicit spatial indicator? (2) Will there be perceptual differences in languages in that respect between Chinese and English? If one's perception of gesture is also sensitive to linguistic choices, then firstly Chinese-English bilinguals are likely to prefer vertical gestures to horizontal ones for Chinese time references with vertical spatial metaphors. Also, Chinese speakers are likely to prefer vertical gestures more 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 vertical wordlists were translated into neutral words). To test these hypotheses, Chinese-English bilinguals were asked to do a rating task.

Method

Participants

109 late Chinese-English bilinguals (L1 = Chinese; 52 F and 57 M; mean age = 18 yrs) from Nanjing University, China participated in the first part of the experiment and 90 participants came back for the second part. They were paid for their participation.

Stimuli

Thirty items were constructed, consisting of eight target items of time references and twenty-two fillers. Each item consisted of a sentence which was followed by a silent video clip of a seated person (visible from shoulders to upper legs) who made specific gestures. 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 would show an actor who first pointed to his left side and then right side (horizontal gesture plane) or, in a counterbalanced version, who pointed upward and downward (vertical gesture plane) to indicate the time conceptions of 'last week' and 'next week'.

Table 2: Time references with vertical spatial metaphors and time references with neutral words.

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

Among the eight items in Table 2, time references (1), (2), (3) and (4) contained vertical spatial metaphors and time references (5), (6), (7) and (8) were literally neutral (no spatial metaphors) in Chinese. Half of the time references [(1) & (2); (5) & (6)] were performed in the vertical gesture plane and the other half [(3) & (4); (7) & (8)] were performed in the horizontal plane. These references were counterbalanced in gesture planes by creating a second version, in which, items [(1) & (2); (5) & (6)] were performed in the horizontal plane whereas the other four were performed in the vertical plane.

Additionally, half of the fillers (unrelated to time conceptions) had incongruent gestures (of varying degree), with the sentence instructions and the other half were congruent. The text was created in Chinese and the English text was a translation of the Chinese one in neutral words. All tests were conducted via Qualtrics.

Procedure

Participants took part twice in the experiment, once in Chinese and once in English, with an interval of one week. In the first test, they were randomly assigned to one of the two versions to fill out a 1-7-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'. The sequence of the languages tested was counterbalanced. Data from 30 participants were excluded because they either did the two tests in the same language or did not finish the task.

Results and Discussion

First, there was no significant difference between the two versions, F (1, 77) = .031, p = .86, η_p^2 = .00, so data from the two versions were merged. A 2×2×2 repeated measures ANOVA with wordlist-type (vertical and neutral) × gesture plane (horizontal and vertical) × language (Chinese and English) as within subject factors, and rating scores as dependent variable revealed that there were main effects of word-type, F (1, 78) = 37.11, p < .001, η_p^2 = .32 and gesture plane, F (1, 78) = 23.65, p < .001, η_p^2 = .22, but there was no main effect of language, F (1, 78) = 1.83, p = .18, η_p^2 = .02. Furthermore, there was a significant interaction between wordlist-type and gesture plane, F (1, 78) = 17.08, p < .001, η_p^2 = .18, a significant interaction between wordlist-type and language, F (1, 78) = 4.26, p = .042, η_p^2 = .052, and a marginally significant interaction between word-type × gesture plane × language, F (1, 78) = 3.39, p = .07, η_p^2 = .04 (see Figure 4).

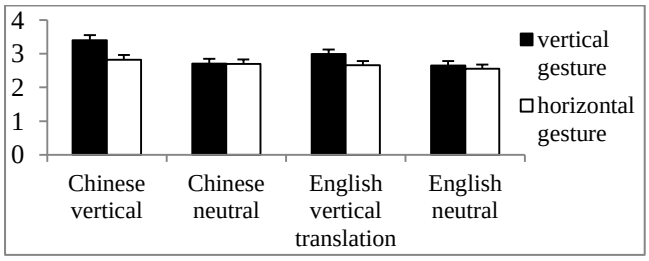

Figure 4: The mean rating scores of the two types of time references (words with vertical spatial metaphors and with neutral words) in Chinese and in English.

In order to understand these interactions, follow-up *post hoc* analyses (with Bonferroni α level corrected to 0.025) were conducted. Firstly, for Chinese wordlists containing vertical spatial metaphors (Chinese vertical), participants preferred vertical gestures (M = 3.40, SD = 1.38) to horizontal gestures (M = 2.80, SD = 1.27), t (78) = 5.53, p < .001, r = .53, 95% CI = (.37, .79). One may immediate interpret this as Chinese speakers' general preference of vertical gestures over horizontal 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 that for 'Chinese neutral' wordlists (M = 2.70, SD = 1.17), t (78) = 8.20, p < .001, r = .68, 95% CI = (.52, .86), but they did not rate significantly different in horizontal gestures for the two type of wordlists, t (78) = 1.25, p = .21, r = .14, 95% CI = (-.07, .31) (see Figure 4). Assuming that the neutral and metaphoric time references share the same conceptualisation of time within a language, the major differences between the two are the differences in linguistic choices. This indicated that the vertical spatial metaphor of time referents had an influence on perceiving the vertical temporal gestures.

With respect to English, the pattern of results was quite similar to that in Chinese. For the English translation of Chinese vertical spatial metaphors (English vertical translation), participants also preferred vertical gestures (M = 2.99, SD = 1.32) to horizontal 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 preference for wordlists without spatial metaphors (English neutral), either. Besides, vertical gestures for 'English vertical translation' (M = 3.40, SD = 1.38) was rated by a significantly higher score than vertical gestures for the 'English neutral' wordlists (M = 2.70, SD = 1.17), t (78) = 8.20, p < .001, r = .68, 95% CI = (.16, .53) (see the right half of Figure 4). The similar pattern of results from the L1 and L2 suggests that Chinese speakers still thought in their first language when perceiving the English words, and the language in which they perceived appeared to have little influence on the perception of the gestures. This also supports Wu and Thierry (2010)'s proposal that languages are co-activated in a bilingual speaker.

However, 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 that when perceiving the English translation of the wordlists (M = 2.99, SD = 1.32), t (78) = 2.05, p = .044, two tailed, r = .23, 95% CI = (.11, .81), but this 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 in Figure 4). As for the rating score for the horizontal gestures, the difference across the two languages was not significant for either type of wordlist (white bars). In short, only for wordlists with vertical spatial metaphors were more preferred in Chinese as opposed to English. The results indicated that the English translation of Chinese vertical spatial metaphors weakened the linguistic influence on gesture perception. This raises an interesting question as to what extent these cross-linguistic differences in mental lexicons will cause the differences in conceptualisation. Future work should study the correlation between gesture production, length of stay, and level of L2 proficiency.

To sum up, the results showed that firstly, in Chinese, the bilingual speakers preferred the vertical gestures more for time references with vertical spatial words than for neutral words. Secondly, when perceiving time references with vertical spatial metaphors, they preferred vertical gestures to horizontal gestures. Additionally, participants preferred vertical gestures for time references containing vertical spatial metaphor over vertical gestures for the English translation.

Interestingly, the results of the perception experiment displayed a similar pattern to that of the production,

especially in the L1 Chinese. The parallel between production and perception data suggests that the linguistic encoding possibilities of time conceptions also have an influence on the perception of temporal gestures. Superficially, when having the linguistic input, one's ideas or concepts of a certain entity are activated and the corresponding mental stimulations are created, including possible *co-thought gestures* based on the action schemata. These gestures are likely to be influenced by the linguistic choices, and adjusted to aid for the understanding of the gesture input. Finally, there is an integration of the semantic information from both inputs.

General Discussion and Conclusion

This study investigated late Chinese-English bilingual speakers' production and perception of temporal gestures for time conceptions. In the production experiment, first of all, participants' performances of vertical gestures for time references in both Chinese and English indicated that they can employ a vertical conceptualisation of time. The results from the comparisons of the two types of time references within and between languages showed that the extent to which vertical gestures were produced was also dependent on the linguistic choices. Therefore, gestures are shaped by the online interface between spatio-motoric thinking and speaking, in which spatial imagery is adjusted to fit the verbalisation, rather than by the pre-determined language-specific conceptual schemas (vertical conceptualisations of time) *only* (Boroditsky, 2001). The study provides unique new evidence for the Interface Hypothesis (Kita & Özyürek, 2003) and suggests that gesture production is dynamic and sensitive to linguistic encoding possibilities.

Additionally, linguistic choices seem to also influence the perception of gestures. For example, in comparison to the horizontal gesture plane, the vertical gesture plane was more preferred for time references with vertical spatial metaphors. Thus the perception of temporal gestures is affected by the vertical spatial metaphors, similarly as what we have found for production. This similar patterns exhibited in both studies not only provide insight into the immediate and chronic influence of a language on the production and perception of gestures, but also suggest that there is an interconnection between perception and production. Future work can research how the production mechanisms are used to predict perception of gestures.

Acknowledgements

The first author received financial support from NWO Promoties in de geesteswetenschappen (322-89-007), which is greatly acknowledged. We thank the reviewers' for their useful comments. We are grateful to all the participants and colleagues for their participation in the experiments. We thank Hua CHEN for her help with the perception experiments in Nanjing University, China.

References

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

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

Calbris, C. (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.

Casasanto, D. & Jasmin, K. (2012). The hands of time: Temporal gestures in English speakers. *Cognitive Linguistics*, *23*(4), 643-674.

Chui, K. (2011). Conceptual metaphors in gestures. *Cognitive Linguistics*, *22*(3), 437-458.

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: Academic Press.

Gu, Y., Mol, L., Hoetjes, M.W., & Swerts, M. (2013). What can Chinese speakers' temporal gestures reveal about their conception of time? *Proceedings of TiGeR 2013 conference*. Tilburg, Netherlands.

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.

Hostetter, A. B. & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin and Review*, *15*(3), 495-514.

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.

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.

Mashal, N., Andric, M., & Small, S.L. (2012). Motor and non-motor language representations in the brain. In: Faust, M. (Ed.), *The Handbook of the Neuropsychology of Language* (pp. 276-293). Oxford, England: Wiley-Blackwell.

Núñez, R., & Sweetser, E. (2006). With the future behind them: Convergent evidence from Aymara language and gesture in the cross-linguistic comparison of spatial construals of time. *Cognitive Science*, *30*(3), 401-450.

Özyürek, A., Kita, S., Allen, S., Furman, R., & Brown, A. (2005). How does linguistic framing of events influence co-speech gestures? Insights from cross-linguistic variations and similarities. *Gesture*, *5*(1/2), 219-240.

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, *108*(1), 155-184.

Slobin, D. (1987). Thinking for speaking. *Proceedings of the Berkeley Linguistic Society*, *13*, 435-445.

Wu, Y.J. & Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. *Journal of Neuroscience, 30* (22), 7646-7651.