

Time flows vertically in Chinese

Li, Yang; Oppenheim, Gary; Thierry, Guillaume

Brain and Cognition

DOI:

[10.1016/j.bandc.2023.106057](https://doi.org/10.1016/j.bandc.2023.106057)

Published: 01/08/2023

Publisher's PDF, also known as Version of record

[Cyswllt i'r cyhoeddiad / Link to publication](#)

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):

Li, Y., Oppenheim, G., & Thierry, G. (2023). Time flows vertically in Chinese. *Brain and Cognition*, 170, Article 106057. <https://doi.org/10.1016/j.bandc.2023.106057>

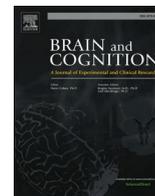
Hawliau Cyffredinol / General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Time flows vertically in Chinese

Yang Li^{a,b,*}, Gary Oppenheim^{b,c}, Guillaume Thierry^{b,d}

^a Faculty of Sport, Health, and Social Sciences, Solent University, UK

^b Department of Psychology, School of Human and Behavioural Sciences, Bangor University, Bangor, Wales, UK

^c Department of Psychology, University of Texas at Austin, USA

^d Faculty of English, Adam Mickiewicz University, Poland

ARTICLE INFO

Keywords:

Spatiotemporal metaphors
Vertical axis
Event-related potentials
N400
Semantics

ABSTRACT

Speakers of Mandarin Chinese are thought to conceptualise time along the vertical axis—as evidence for metaphor embodiment—but the extant behavioural evidence remains unclear. Here, we used electrophysiology to test space–time conceptual relationships implicitly in native speakers of Chinese. We employed a modified arrow flanker task, in which the central arrow in a set of three was replaced by a spatial word (e.g., 上–‘up’), a spatiotemporal metaphor (e.g., 上个月–‘last month’, literally ‘up month’) or a non-spatial temporal expression (e.g., 去年 –‘last year’, literally ‘gone year’). N400 modulations of event-related brain potentials served to measure the level of perceived congruency between semantic word content and arrow direction. Critically, we tested whether N400 modulations expected for spatial words and spatial temporal metaphors would generalise to non-spatial temporal expressions. In addition to the predicted N400 effects, we found a congruency effect of a similar magnitude for non-spatial temporal metaphors. On the basis of direct brain measurements indexing semantic processing, and in the absence of contrastive behavioural patterns, we demonstrate that native speakers of Chinese conceptualise time along the vertical axis, and thus have embodied spatiotemporal metaphors.

1. Introduction

Humans frequently use certain gestures and linguistic forms to anchor the temporality of events in the spatial domain (e.g., Boroditsky, Fuhrman & McCormick, 2011; Gu, Zheng & Swerts, 2019; Lai & Boroditsky, 2013). In English-speaking cultures, for example, people often figuratively refer to the front space when stating that they *look forward* to an event set in the future, or they will *look back* on a memorable life event with nostalgia. Such spatiotemporal metaphors serve as concrete reference in order to conceptualise time, an abstract concept.

Whilst most cultures and languages represent time in terms of space, there is considerable variation in how this relationship manifests. For example, speakers of Aymara and Moroccan will *look back* (rather than forward) to a future event (de la Fuente, Santiago, Román, Dumitrache & Casasanto, 2014; Núñez & Sweetser, 2006), and in some languages, such as Mandarin Chinese, the same orientation (e.g., front) can refer either to the future (e.g., *zhan-wang*– ‘looking into the future’, literally, ‘unfold-to gaze’) or the past (e.g., *qian-tian* – ‘the day before yesterday’, literally, ‘front day’). Moreover, time can be represented along different axes within the same language: for instance, English uses both sagittal (front-back) and horizontal (left–right) axes, while Chinese uses sagittal

and vertical axes (up-down, e.g., Boroditsky, Fuhrman & McCormick, 2011; Gu et al., 2019; Lai & Boroditsky, 2013; Boroditsky, 2001) to represent time.

Thus, an interesting question is whether time conceptualization engaging spatial representation relates to the way in which language refers to time. Several renowned studies have suggested that Chinese-English bilinguals conceptualize time vertically (e.g., Boroditsky, Fuhrman & McCormick, 2011; Lai & Boroditsky, 2013; Boroditsky, 2001; Chen, 2007; Fuhrman et al. 2011; Miles, Tan, Noble, Lumsden & Macrae, 2011; Yang & Sun 2016). For example, using a spatial priming paradigm, Boroditsky (2001) showed that English native speakers were faster to respond to temporal relationships following a horizontal prime (left/right), whereas Chinese native speakers were faster after experiencing a vertical prime. While this result was consistent with observations from follow-up experiments (Miles, Tan, Noble, Lumsden & Macrae, 2011; Yang & Sun 2016), others have failed to replicate these effects (Chen, 2007; January & Kako, 2007; Tse & Altarriba, 2008), their authors arguing instead that Chinese-English bilinguals adhere to a human cognitive universal of time being represented horizontally by default and that the use of the vertical axis for the representation of time is conceptually redundant. While cross-linguistic variability in

* Corresponding author at: Faculty of Sport, Health and Social Science, Solent University, UK.

E-mail address: yang.li@solent.ac.uk (Y. Li).

<https://doi.org/10.1016/j.bandc.2023.106057>

Received 25 December 2022; Received in revised form 24 May 2023; Accepted 12 June 2023

Available online 26 June 2023

0278-2626/© 2023 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

spatiotemporal metaphors could invite speculation about linguistic relativity, experimental invariance across speakers of those languages could instead suggest an underlying cognitive universal that persists despite linguistic variation (cf. e.g., Hall, Mayberry & Ferreira, 2013).

Here, we examined whether using the vertical axis as temporal reference is indeed semantically redundant or instead is a core feature of time conceptualisation in Chinese, that is, a preferred spatial representation against horizontal reference axes. Importantly, instead of depending solely on behavioural responses from participants asked to verbalise mental representations of time and space, we used event-related brain potentials (ERPs) to measure congruency between minimal temporal and spatial information delivered in a highly controlled fashion. This is an important new step because ERP measures index implicit information processing, without requiring participants' explicit appraisal or metacognitive awareness of the process at work. To do this, we presented participants with stimuli involving both spatial and temporal information in the form of written words flanked by either vertical arrows (Experiment 1; Fig. 1A) or horizontal arrows (Experiment 2; see Fig. 1B).

To quantify the congruency between verbal and spatial information conveyed by a word-arrow composite stimulus, we measured the mean amplitude of the N400 component, a well-known index of semantic integration (Kutas & Federmeier, 2011; Kutas & Hillyard, 1984a; Kutas & Hillyard, 1984b; Kutas, Lindamood & Hillyard, 1984). Chinese participants indicated the direction of the arrows flanking a verbal expression in three experimental conditions (see Table 1): *spatial words* (e.g., 上, *shang* – 'up' and 下, *xia* – 'down'), *spatiotemporal metaphors* (i.e., 上个月, *shang-ge yue* – 'last month', literally 'up month', and 下个月, *xia-ge yue* – 'next month', literally 'down month'), or *temporal expressions* (i.e., 去年, *qu nian* – 'last year', literally 'gone year' and 明年, *ming nian* – 'next year', literally 'bright year').

We hypothesised that, if Chinese speakers conceptualise time vertically, then we should observe verbal-spatial interference between temporal expressions and the direction of vertical arrows in Experiment 1. Assuming that time is conceptualized along the vertical axis in Chinese, greater N400 amplitudes should be recorded in response to incongruent than congruent word-arrow configurations along the vertical axis. We expected congruency effects to appear for stimuli that contain the characters for 'up' and 'down' (i.e., 上个月–'last month' and 下个月–

'next month'), and if Chinese speakers represent time along the vertical axis more generally, then time expressions that lack these directional characters (i.e., 去年 –'last year' or 明年 –'next year') should elicit the same modulation.

We also conducted a version of this experiment using horizontal rather than vertical arrows to serve as a control. Predictions for the horizontal arrows in Experiment 2 were markedly different: Whilst N400 effects could be expected for 左–'left' and 右–'right' when flanked by incongruent versus congruent arrows, arrow direction was unrelated to either the meaning of the spatial temporal metaphor or that of the non-spatial temporal expressions of Chinese, and thus no N400 modulation should be expected in these cases.

2. Materials and methods

2.1. Participants

Twenty-three native Chinese speakers took part in both Experiments 1 and 2. Data from one participant were excluded due to poor electro-physiological recording quality relating to excessively high impedances, reducing N to 22. All participants received cash or course credits for their participation. The experiment was approved by the ethics committee of the School of Human and Behavioural Sciences at Bangor University.

2.2. Design & stimuli

In Experiment 1 (vertical arrows), the critical words were spatial words 上 (*shang*) – 'up' and 下 (*xia*) – 'down', spatiotemporal metaphors 上个月 (*shang-ge yue*) – 'last month' (literally 'up month') and 下个月 (*xia-ge yue*) – 'next month' (literally 'down month'), and non-spatial temporal expressions 去年 (*qu nian*) – 'last year' (literally 'gone year') and 明年 (*ming nian*) – 'next year' (literally 'bright year'; see Table 1). They were surrounded by dark red arrows set above and below each term, both pointing either up or down (Fig. 1A). In Experiment 2 (horizontal arrows), the critical words were spatial words 左 (*zuo*) – 'left' and 右 (*you*) – 'right', and the same spatiotemporal metaphors and temporal expressions used in Experiment 1. Expressions were flanked by dark red arrows on either side, both pointing to the left or to the right

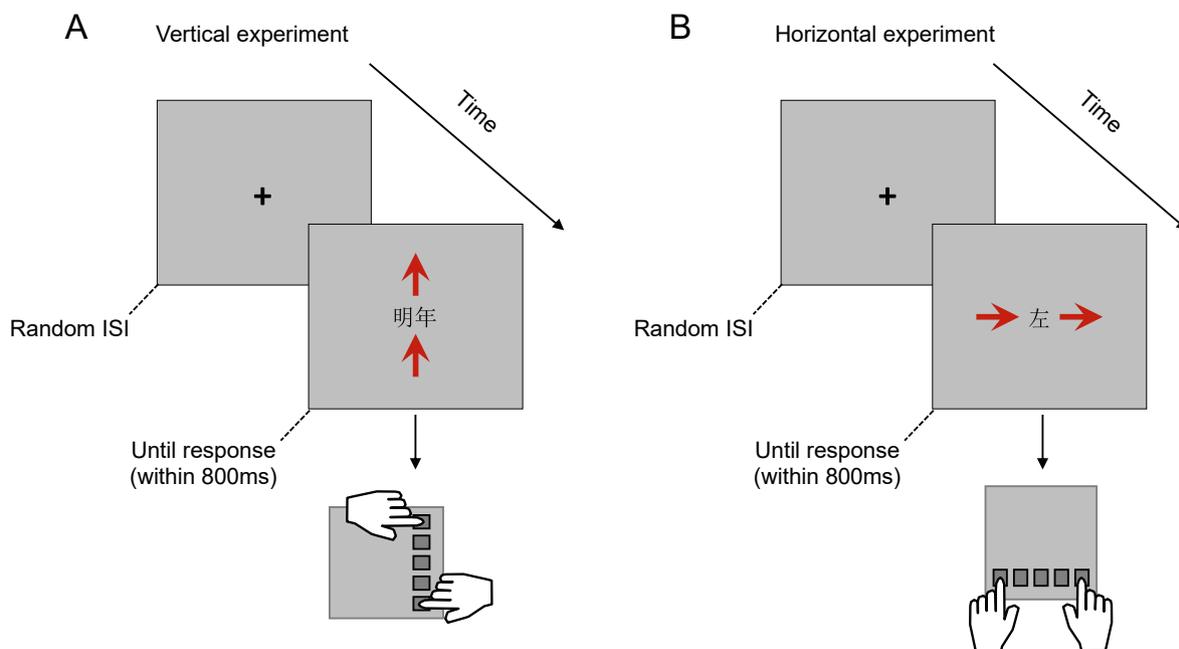


Fig. 1. Sample of stimulus streams in the vertical (A) and the horizontal (B) experiments. Participants reported the direction of flanking arrows as quickly as possible when words in Chinese did not depict food items.

Table 1
Experimental conditions featured in the two experiments.

Experiment	Conditions	Chinese	Pinyin	Translation	Literal translation	Arrow direction / congruency	
Vertical experiment	Spatial words	上	shang	up	up	up / yes	down / no
		下	xia	down	down	up / no	down / yes
	Spatiotemporal metaphors	上个月	shang-ge yue	last month	up month	up / yes	down / no
		下个月	xia-ge yue	next month	down month	up / no	down / yes
	Temporal words	去年	qu nian	last year	gone year	up / ?	down / ?
明年	ming nian	next year	bright year	up / ?	down / ?		
Horizontal experiment	Spatial words	左	zuo	left	left	left / yes	right / no
		右	you	right	right	left / no	right / yes
	Spatiotemporal metaphors	上个月	shang-ge yue	last month	up month	left / -	right / -
		下个月	xia-ge yue	next month	down month	left / -	right / -
	Temporal words	去年	qu nian	last year	gone year	left / -	right / -
明年	ming nian	next year	bright year	left / -	right / -		

(Fig. 1B). There were also 78 filler words (1–3 characters in length) corresponding to food items (types of food, fruit, vegetable, drinks), which were also surrounded by arrows as described above.

2.3. Procedure

Participants filled in a language background and a reading/writing habits questionnaire whilst the 64-channel Ag/AgCl electrophysiological cap was being fitted. They sat in front of a 19-inch CRT monitor, placed at a distance of 100 cm in a sound-proofed room. Each critical stimulus surrounded by arrows in one of two orientations was presented 36 times in the horizontal experiment (across four blocks) and 36 times in the vertical experiment (across four blocks). The order of the blocks was pseudo-randomised between participants and counterbalanced. Each trial started with a fixation cross for a duration of between 200 and 300 ms (duration was randomly selected in the interval in steps of 10 ms), followed by the expression flanked by two arrows for a maximum duration of 800 ms during which participants were required to indicate the direction of the flanking arrows, whilst keeping their gaze on the centre, by pressing one of two buttons on a response box set under their left and right index fingers. Each experiment also included filler trials featuring food expressions, for which participants were instructed to withhold their response (as to ensure systematic semantic processing of the expression at fixation). The instruction to participants was thus to indicate the direction of the arrows by pressing corresponding keys when and only when the expression presented at fixation did not refer to food or drinks. Participants responded using the left and right index fingers which were mapped consistently on the horizontal axis for the horizontal experiment but remapped on the sagittal axis in the vertical experiment by rotating the response box on their lap. In the horizontal experiment, response mapping was intuitively congruent with arrow direction in all cases (right finger - right arrow) but in the vertical experiment, response mappings were counterbalanced between participants: half of the participants used their right finger to report arrows pointing up and the other half used their right index finger to report arrows pointing down (see Fig. 1).

2.4. EEG data recording and processing

Electrophysiological data were recorded at a rate of 1 kHz from 64 Ag/AgCl electrodes according to the extended 10–20 convention. The electroencephalogram (EEG) was filtered online using a bandpass filter with cut off values of 0.05 Hz low pass, 200 Hz high pass and an accuracy of 0.15 nV/LSB. It was referenced to electrode Cz and the impedances were kept below 5 k Ω . The electroencephalogram (EEG) was downsampled to 250 Hz offline and filtered with a 0.1 Hz (24 dB/oct) high pass. The data was scanned for abnormalities and artefact-containing segments (i.e., muscle artefacts) before being re-referenced to the global average reference (excluding electrooculogram channels). Independent component analysis (ICA) correction was performed focusing on blinks, eye movements, and muscle activity. We then

applied a 30 Hz (24 dB/oct) low-pass filter and extracted epochs ranging from -200 to 1000 ms from continuous data. Baseline correction was applied relative to -200 to 0 ms. Epochs with activity exceeding ± 100 μ V at any electrode site apart from the electrooculogram channels within each epoch window were discarded. N400 mean amplitudes were extracted in the typical time window of 300–500 ms after stimulus onset at electrodes of predicted maximal sensitivity (i.e., FC1, FCz, FC2, C1, Cz, C2, CP1, CP2, CPz; Kutas & Fedemeier, 2000; Kutas & Hillyard, 1984a; Kutas & Hillyard, 1984b; Kutas, Lindamood & Hillyard, 1984). ERP amplitudes were analysed by means of a 3 (Condition) * 2 (Congruency) repeated-measures ANOVA in both experiments. The factor Condition had three levels {spatial word, spatiotemporal metaphor, temporal word} and the factor Congruency had two levels {congruent, incongruent}.

2.5. Behavioural data analysis

We used linear mixed effects regression to analyse latencies for correct responses via the lme4 package (Bates et al., 2021) in R, using likelihood ratio testing to estimate p-values via the afex::mixed function (Singmann, Bolker, Westfall & Aust, 2017). We first excluded incorrect responses and improbably fast response times (RTs < 200 ms), and, following a Box-Cox test, we applied an inverse transformation ($-100000 * RT^{-1}$) to the remaining RTs to correct a deviation from normality. Regression models contained two centered fixed effects and their interaction: (1) *Stimulus-flanker congruency*, a binomial contrast with levels {incongruent, congruent}, and (2) *Orthographic stimulus*, a three-level Helmert-coded factor consisting of two orthogonal contrasts: (i) *Contains time word* {no, yes}, contrasting vertical words with the mean of the temporal and spatiotemporal stimuli, and (ii) *Contains spatiotemporal metaphor* {no, yes}, contrasting the non-spatial temporal stimuli (i.e., bright/gone year) with the spatiotemporal stimuli (i.e., up/down month). Models also included by-participants maximal random effects structures, omitting the estimation of correlations between random effects to facilitate convergence (Barr, Levy, Scheepers & Tily, 2013). Error rate analyses apply the same procedures to logistic mixed effects regression by including a binomial link function.

3. Results

3.1. Experiment 1: Word categorization with vertical flankers

3.1.1. Behavioural results

As illustrated in Fig. 2A, participants were, overall, significantly slower to classify words flanked by incongruent distractors than those flanked by congruent distractors ($\beta = -13.40$, SE = 2.14, $\chi^2_1 = 22.48$, $p < 0.001$). The main effect of the three-level orthographic stimulus factor also reached significance ($\chi^2_2 = 9.42$, $p = 0.009$), reflecting (1) significantly slower responses for the directional terms than for the temporal stimuli ($\beta = -3.09$, SE = 1.20, $t = -2.57$) and (2) marginally slower responses for temporal stimuli that contained spatiotemporal metaphors

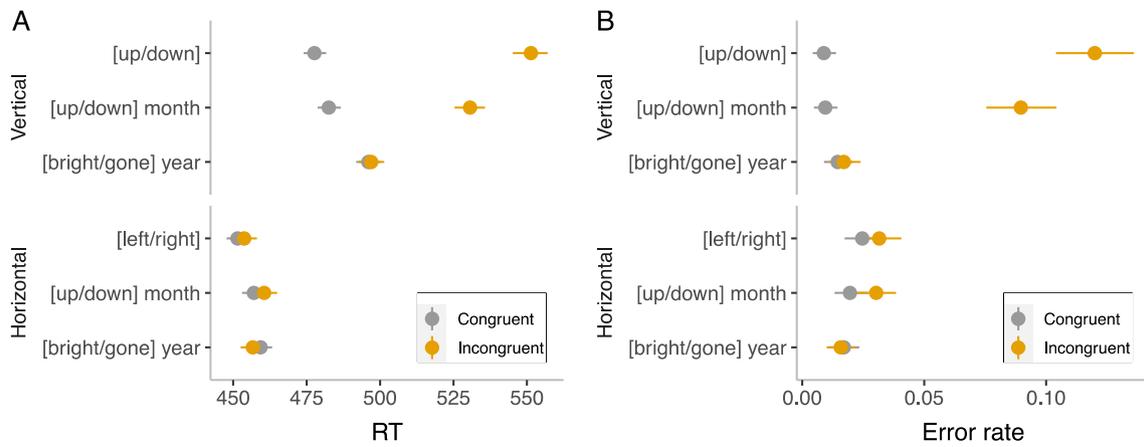


Fig. 2. (A) Mean response times and error rates (B) of spatial words, spatiotemporal metaphors and temporal words flanked by vertical distractors in Experiment 1 (top) or horizontal distractors in Experiment 2 (bottom). Error bars represent bootstrapped confidence intervals.

than those that did not ($\beta = 2.67$, $SE = 1.34$, $t = 1.99$). The overall congruency by stimulus interaction also reached significance ($\chi^2_2 = 41.70$, $p < 0.001$), reflecting significantly stronger congruency effects for (1) the directional terms than for the temporal stimuli ($\beta = 15.84$, $SE = 2.72$, $t = 5.83$), and (2) temporal stimuli that contained spatiotemporal metaphors than those that did not ($\beta = -16.07$, $SE = 2.71$, $t = -5.94$). Re-fitting models to restricted datasets for each level of the orthographic stimulus factor showed significant congruency effects for both the spatial terms ($\beta = -24.11$, $SE = 3.79$, $\chi^2_1 = 23.03$, $p < 0.001$) and spatiotemporal metaphors ($\beta = -16.40$, $SE = 2.79$, $\chi^2_1 = 20.81$, $p < 0.001$) but not non-spatial temporal terms ($\beta = -0.31$, $SE = 1.12$, $\chi^2_1 = 0.07$, $p = 0.79$).

Accuracy measures, also illustrated in Fig. 2B, produced very similar results. Participants were, overall, significantly more error-prone when classifying words flanked by incongruent distractors than those flanked by congruent distractors ($\beta = -1.56$, $SE = 0.28$, $\chi^2_1 = 20.54$, $p < 0.001$). The main effect of the three-level orthographic stimulus factor also

reached significance ($\chi^2_2 = 14.14$, $p < 0.001$), reflecting (1) significantly more error-prone responses for the directional terms than for the temporal stimuli ($\beta = -0.47$, $SE = 0.19$, $t = -2.50$) and (2) significantly more error-prone responses for temporal stimuli that contained spatiotemporal metaphors than those that did not ($\beta = 0.67$, $SE = 0.22$, $t = 3.09$). The overall congruency by stimulus interaction also reached significance ($\chi^2_2 = 34.64$, $p < 0.001$), reflecting significantly stronger congruency effects for (1) the directional terms than for the temporal stimuli ($\beta = 1.53$, $SE = 0.35$, $t = 4.40$), and (2) temporal stimuli that contained spatiotemporal metaphors than those that did not ($\beta = -2.19$, $SE = 0.42$, $t = -5.21$). Re-fitting models to restricted datasets for each level of the orthographic stimulus factor showed significant congruency effects for both the spatial terms ($\beta = -2.75$, $SE = 0.39$, $\chi^2_1 = 29.04$, $p < 0.001$) and spatiotemporal metaphors ($\beta = -2.36$, $SE = 0.37$, $\chi^2_1 = 24.61$, $p < 0.001$) but not non-spatial temporal terms ($\beta = -0.11$, $SE = 0.33$, $\chi^2_1 = 0.11$, $p = 0.74$).

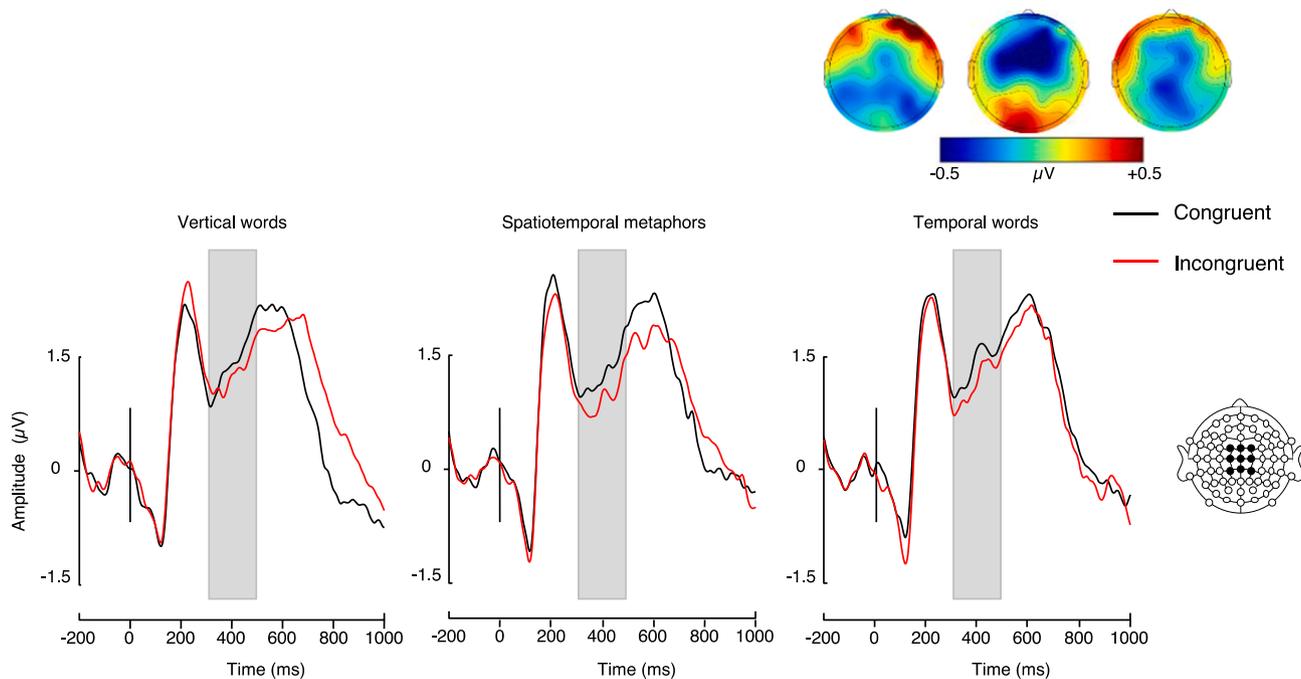


Fig. 3. ERPs depict the signal from a linear deviation of nine electrodes in Experiment 1. The highlighted time window shaded in light grey is 300–500 ms where mean N400 amplitudes were computed. The topographies represent mean voltage differences between incongruent and congruent conditions over the duration of the N400 analysis window and laid out in the same order as the ERPs presented below.

3.1.2. ERP results

The repeated measures ANOVA on N400 mean amplitudes revealed a significant main effect of congruency, $F(1, 21) = 8.85$, $P = 0.01$, $\eta^2p = 0.3$, such that incongruent conditions elicited more negative N400 amplitudes than congruent ones (see Fig. 3). There was no main effect of the three-level orthographic stimulus factor, $F(2, 42) = 2$, $P = 0.15$, $\eta^2p = 0.09$, and there was no interaction between congruency and orthographic stimulus, $F(2, 42) = 0.3$, $P = 0.75$, $\eta^2p = 0.01$.

3.2. Experiment 2: Word classification with horizontal flankers

3.2.1. Behavioural results

Applying the same regression models as in Experiment 1, neither the main effects of stimulus-flanker congruency nor the orthographic stimulus factor approached significance in the RT analysis, nor did their interaction (all $p > 0.1$; see Fig. 2A). In the error analyses, only the overall orthographic stimulus factor reached significance ($\chi^2_2 = 6.14$, $p = 0.046$; all other $p > 0.1$).

3.2.2. ERP results

We found a significant main effect of the three-level orthographic stimulus factor on mean N400 amplitude, $F(2, 42) = 3.41$, $P = 0.04$, $\eta^2p = 0.14$. *Post hoc* pairwise comparisons showed that spatiotemporal orthographic stimuli elicited more negative N400 amplitudes overall than the spatial word condition ($P = 0.03$). There was no significant main effect of congruency, $F(1, 21) = 2.6$, $P = 0.12$, $\eta^2p = 0.11$, or interaction between orthographic stimulus and congruency, $F(2, 42) = 0.74$, $P = 0.48$, $\eta^2p = 0.03$; see Fig. 4).

4. Discussion

This study investigated whether native speakers of Chinese tend to conceptualize time along the vertical axis, and how such tendency compares with conceptualisation along the horizontal (transversal) axis. Beyond the anticipated congruency effects for ‘up’ and ‘down’ words (上 and 下), Experiment 1 showed N400 modulations not only for spatiotemporal metaphors that contain these direct spatial references to

verticality, but also for purely temporal expressions that do not imply any spatial directional information. Experiment 2, in contrast, did not elicit any such modulations from flankers along the horizontal axis, even in the case of simple correspondence between left and right arrows and spatial terms.

Firstly, our results are consistent with studies suggesting that Chinese-English bilinguals represent time along the vertical axis (e.g., Boroditsky, Fuhrman, & McCormick, 2011; Lai & Boroditsky, 2013; Boroditsky, 2001) in cases where spatiotemporal metaphors are directly relevant for the task and stimulus at hand (i.e., ‘last month’ and ‘next month’, which contain the up and down characters in Chinese). Furthermore, participants displayed similar congruency effects when processing temporal metaphors that do not specify a spatial orientation (e.g., ‘last year’ and ‘next year’). A conservative explanation for this is that native Chinese speakers conceptualize time along the vertical axis, that is, they generalise the axis used for space and spatiotemporal metaphors of Chinese to temporal expressions that do not refer to verticality or space.

In several published studies, authors have argued that native speakers of Mandarin Chinese may organise time along the vertical axis because of the writing direction of Chinese (e.g., Bergen & Chan, 2012; Fuhrman & Boroditsky, 2010). However, this writing direction was abandoned by modern China in 1956 —nearly half a century before our participants were born. And indeed, data from questionnaires confirmed that our participants seldom wrote or read Chinese vertically (writing: Mean = 1.74, SD = 1.01; reading: Mean = 2.57, SD = 1.83; scale range from 1 to 10). Therefore, it seems implausible that Chinese speakers born in this century would have developed a vertical conception of time in response to their minimal direct experience with a vertical writing system (e.g., Miles, Tan, Noble, Lumsden & Macrae, 2011; Yang & Sun, 2016).

It is interesting to consider why we found no congruency effect along the horizontal axis (Experiment 2). First, it must be noted that the design of this study does not allow us to properly compare the horizontal manipulations to the vertical because the experiments varied in three ways: (a) they used different spatial words/characters as stimuli: 上 – ‘up’ and 下 – ‘down’ in Experiment 1, versus 左 – ‘left’ and 右 – ‘right’ in

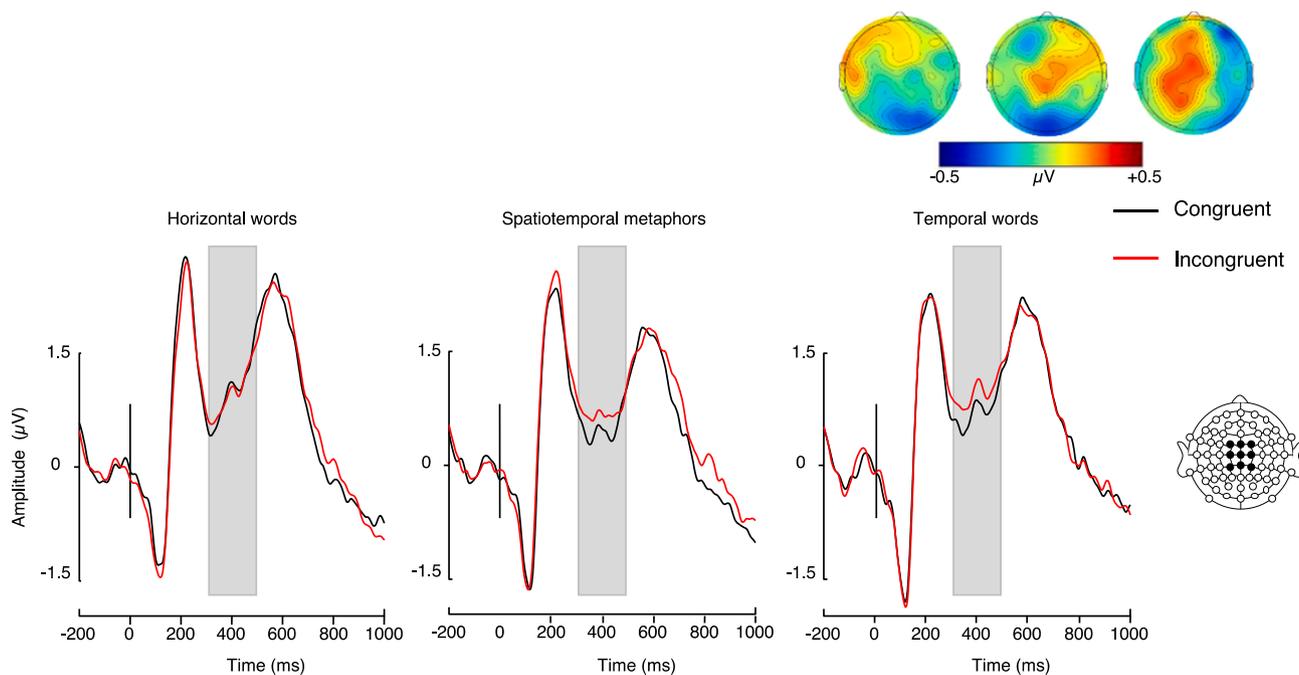


Fig. 4. ERPs depict the signal from a linear deviation of nine electrodes in Experiment 2. The highlighted time window shaded in light grey is 300–500 ms were mean N400 amplitudes were computed. The topographies represent mean voltage differences between incongruent and congruent conditions over the duration of the N400 analysis window and laid out in the same order as the ERPs presented below.

Experiment 2; (b) flanking arrows differed in their locations relative to the primary stimulus and participants' field of view: within the nasal field of view in Experiment 1 (immediately above and below the words), versus peripherally in Experiment 2 (to the left and right of the words); (c) stimulus–response congruency: Experiment 1 required participants to remap vertical arrows to transverse responses (e.g., upward arrow → left hand) whereas Experiment 2 enabled participants to directly map transverse arrows to transverse responses (e.g., left arrow → left hand). We have presented these data as two separate experiments because having multiple differences between their designs cautions against attributing apparent differences in the results to any specific manipulation. The drawback of this approach is that although Experiment 1 provides evidence for Chinese speakers' vertical representation of time, we cannot fully exclude the possibility that they may also represent time along a transverse axis.

Note that the lack of a congruency effect in the case of left / right spatial words flanked by left / right arrows is not that surprising, given the additional insights provided by the behavioural data analysis: Indeed, response times in Experiment 2 were extremely fast and non-discriminative between experiment conditions. This is very likely due to the directness of the mapping between arrow direction and response side (cf. (c) above) and probably indicates that processing of arrow direction was too fast to reflect any priming from the processing of word meaning. In other words, arrow direction was determined before word meaning could interfere with button press decision.

Although we have described the N400 component as an index of semantic integration, it could also reflect interference suppression in executive tasks. We are unable to empirically determine whether the observed N400 modulations in Experiment 1 arose from integration or suppression, but this research is mainly interested in the modulations' ultimate source. In either case, the ultimate source of the effect is a conflict between the *meaning* of the temporal expression and the arrows' orientation in space.

In summary, the findings of the current study are consistent with the general assumption of Linguistic Relativity (Whorf, 1956), suggesting that language structure (in this case spatiotemporal metaphors of Chinese) can shape individuals' conceptualisations (in this case the representation of time along the vertical axis). Critically, the experiment also revealed that Chinese-English bilinguals conceptualize time vertically, in case of concepts that are labelled using expressions that contain no reference to direction in space. These findings of generalization to verbal forms that do not directly reference space are novel and extend linguistic relativity effects beyond strict relationships between labels and the concepts they refer to. Here, temporal concepts that do not refer to verticality (last year, literally “gone year” and next year, literally “bright year” in Chinese) appear to have inherited conceptual properties of similar temporal expressions that correspond to spatial–temporal metaphors in Chinese. This stands in contrast with the theory that labels entertain a strong and exclusive relationship with corresponding semantic representations, namely, the Label-feedback Hypothesis (Lupyan, 2012). Whilst our findings are not incompatible with Lupyan's hypothesis, they offer a path towards generalized linguistic relativity effects.

Along those lines, we note that our specific results are equally compatible with the less controversial claim that thought shapes language. Conceptual representations obviously drive speakers' choice of content words and their more deliberate use of metaphor, so it would not be a stretch to assume that some cross-cultural conceptual variation could give rise to cross-linguistic variation in idiomatic *linguistic* metaphors. But until now, the empirical evidence for even such variation in spatiotemporal conceptual representations was remarkably elusive, leading to claims that humans consistently represent the orientation of time despite inconsistency in linguistic metaphors for it. Using a direct measure of brain activity in this study has allowed us to identify a vertical orientation for temporal representations that would have been invisible from behavioural measures. Regardless of the direction of

causality, this ERP evidence argues for the nonindependence of linguistic metaphors from conceptual representations.

Nevertheless, we acknowledge that the evidence afforded by this study regarding conceptualisation of time along the horizontal axis is insufficient and will likely require further investigation. In addition, it is also unknown what the results would look like if the experiment did not feature any spatial term, or any spatiotemporal metaphor. Indeed, one could imagine that the effects found for non-directional temporal metaphors such as 去年 – ‘last year’ and 明年 – ‘next year’ may depend upon presentation of spatial terms in the same experimental contexts.

Future studies will need to address some of the limitations of the present experiment, namely, the limited comparability of stimuli and response mappings across axis orientations as well as the contextual effects of spatial expression on temporal expression processing. One suggestion would be to implement the same paradigm without any mention of metaphors or spatial terms but rather based on content words, possibly cross-linguistically, that imply temporal orientation without calling upon the metaphors themselves.

CRedit authorship contribution statement

Yang LI: Conceptualization, Data curation, Formal analysis, Writing - original draft, Writing - review & editing. **Gary Oppenheim:** Conceptualization, Formal analysis, Writing - original draft, Writing - review & editing. **Guillaume Thierry:** Conceptualization, Formal analysis, Writing - original draft; Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We thank Aina Casaponsa and Manon Jones for the contribution to the discussion of the results. GT was supported by a grant from the Polish National Agency for Academic Exchange under the NAWA Chair Programme (PPN/PRO/2020/1/00006)

References

- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., Fox, J., Bauer, A. & Krivitsky, P. (2021). lme4: Linear mixed-effects models using Eigen and S4. R package version, 1(1–27), 1. <http://CRAN.R-project.org/package=lme4>.
- Bergen, B., & Chan Lau, T. T. (2012). Writing Direction Affects How People Map Space Onto Time. *Frontiers in Psychology*, *3*(109). <https://doi.org/10.3389/fpsyg.2012.00109>
- Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, *43*(1), 1–22. <https://doi.org/10.1006/cogp.2001.0748>
- Boroditsky, L., Fuhrman, O., & McCormick, K. (2011). Do English and Mandarin speakers think about time differently? *Cognition*, *118*(1), 123–129. <https://doi.org/10.1016/j.cognition.2010.09.010>
- Chen, J.-Y. (2007). Do Chinese and English speakers think about time differently? Failure of replicating Boroditsky (2001). *Cognition*, *104*(2), 427–436. <https://doi.org/10.1016/j.cognition.2006.09.012>
- 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. <https://doi.org/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.

- Fuhrman, O., McCormick, K., Chen, E., Jiang, H., Shu, D., Mao, S., et al. (2011). How linguistic and cultural forces shape conceptions of time: English and Mandarin time in 3D. *Cognitive Science*, 35(7), 1305–1328. <https://doi.org/10.1111/j.1551-6709.2011.01193.x>
- Gu, Y. A. N., Zheng, Y., & Swerts, M. (2019). Having a different pointing of view about the future: The effect of signs on co-speech gestures about time in Mandarin–CSL bimodal bilinguals. *Bilingualism: Language and Cognition*, 1–12. <https://doi.org/10.1017/S1366728918000652>
- Hall, M. L., Mayberry, R. L., & Ferreira, V. S. (2013). Cognitive constraints on constituent order: Evidence from elicited pantomime. *Cognition*, 129(1), 1–17.
- January, D., & Kako, E. (2007). Re-evaluating evidence for linguistic relativity: Reply to Boroditsky. *Cognition*, 104(2), 417–426.
- Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences*, 4(12), 463–470. [https://doi.org/10.1016/s1364-6613\(00\)01560-6](https://doi.org/10.1016/s1364-6613(00)01560-6)
- Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*, 62, 621–647. <https://doi.org/10.1146/annurev.psych.093008.131123>
- Kutas, M., & Hillyard, S. A. (1984a). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307(5947), 161–163. <https://doi.org/10.1038/307161a0>
- Kutas, M., & Hillyard, S. A. (1984b). Event-related brain potentials (ERPs) elicited by novel stimuli during sentence processing. *Annals of the New York Academy of Sciences*, 425(Jun), 236–241. <https://doi.org/10.1111/j.1749-6632.1984.tb23540.x>
- Kutas, M., Lindamood, T., & Hillyard, S. (1984). Word expectancy and event-related brain potentials during sentence processing. In S. Kornblum, & J. Requin (Eds.), *Preparatory states and processes* (pp. 217–237). Hillsdale Lawrence Erlbaum.
- 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, 142. <https://doi.org/10.3389/fpsyg.2013.00142>
- Lupyan, G. (2012). Linguistically modulated perception and cognition: The label feedback hypothesis. *Frontiers in Psychology*, 3(54). <https://doi.org/10.3389/fpsyg.2012.00054>
- Miles, L. K., Tan, L., Noble, G. D., Lumsden, J., & Macrae, C. N. (2011). Can a mind have two time lines? Exploring space–time mapping in Mandarin and English speakers. *Psychonomic Bulletin & Review*, 18(3), 598–604. <https://doi.org/10.3758/s13423-011-0068-y>
- 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. <https://doi.org/10.1207/s15516709cog0000.62>
- Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2017). afex: Analysis of factorial experiments. *R package version*, 0.18-0. <http://cran.r-project.org/package=afex>.
- Tse, C.-S., & Altarriba, J. (2008). Evidence against linguistic relativity in Chinese and English: A case study of spatial and temporal metaphors. *Journal of Cognition and Culture*, 8(3–4), 335–357.
- Yang, W., & Sun, Y. (2016). A monolingual mind can have two time lines: Exploring space-time mappings in Mandarin monolinguals. *Psychonomic Bulletin & Review*, 23(3), 857–864. <https://doi.org/10.3758/s13423-015-0964-7>
- Whorf, B. (1956). *Language, thought, and reality*. Cambridge, MA: MIT Press.