

How verbal labels modulate early Cantonese-English bilinguals' perception and categorization of motion events

International Journal of Bilingualism

1–16

© The Author(s) 2025



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/13670069251386699

journals.sagepub.com/home/ijbYi Wang¹  and Li Wei² 

Abstract

Aims and objectives/purpose/research questions: The current study investigates how strongly verbal labels (i.e., motion verbs) affect early Cantonese-English bilinguals' perception and categorization of similar events. Specifically, it examines the degree to which bilinguals' performance aligns with that of L1 speakers of each language when their access to language ranges from maximal to minimal during task performance.

Design/methodology/approach: A total of 228 participants were recruited and categorized into different language contexts (Cantonese or English). Employing a triad-matching paradigm, we assessed the performance of monolingual and bilingual speakers in making similarity judgements under two distinct conditions: a verbal condition that required explicit language use (Experiment 1) and a nonverbal condition subjected to verbal interference (Experiment 2).

Data and analysis: Two types of measurements were utilized to assess participants' cognitive behaviour: an explicit measurement of categorical preferences and an implicit measurement of reaction time. Mixed-effects models were constructed to compare the performance of bilinguals with that of monolinguals for each language.

Findings/conclusions: Bilinguals displayed categorization preferences and processing efficiency akin to English speakers when language access was available during task performance, suggesting a unified mode of thinking regardless of the language in use. However, the observed language effects on categorization disappeared when the use of language was blocked by concurrent verbal interference.

Keywords

Language and cognition, motion event processing, early Cantonese English bilinguals, label-feedback hypothesis, cognitive restructuring

¹Cardiff University, UK; University College London, UK

²University College London, UK

Yi Wang is also affiliated to School of Psychology, HSE University, Moscow, Russia.

Corresponding author:

Yi Wang, Cardiff University, Cardiff, UK, CF10 3EU, London, UK.

Email: wangy471@cardiff.ac.uk

Introduction

The question of how language influences cognition, also known as the linguistic relativity effect, continues to be a subject of debate (Slobin, 1996; Whorf, 1956; Wolff & Holmes, 2011). A growing body of empirical evidence suggests that language, particularly language-specific labels, can exert either transient and immediate (Athanasopoulos et al., 2015; Bylund & Athanasopoulos, 2017; Montero-Melis et al., 2016) or habitual and durable effects (Casasanto & Boroditsky, 2008) on various cognitive processes, including categorization, visual discrimination, perception, and recognition memory. For instance, language's impact on cognition is most prominent when linguistic labels are employed as tools or strategies to facilitate perception and category formation (Gennari et al., 2002; Goldstone & Hendrickson, 2010; Lupyan, 2012). However, these effects diminish when access to language is minimized, suppressed, or blocked by verbal interference (Papafragou et al., 2008; Trueswell & Papafragou, 2010; Winawer et al., 2007). Lupyan (2012) explains this pattern through the label-feedback hypothesis, which suggests that linguistic labels influence cognitive processing.

In light of the nuanced way language affects cognition, one possible theoretical perspective is that linguistic labels actively influence cognitive processes by systematically activating their associated perceptual features (Lucy, 2016; Lupyan, 2012; Thierry, 2016). This raises questions about how bilingual speakers with typologically different languages behave when they have access to two different sets of linguistic labels. Recent evidence indicates that bilinguals can flexibly switch between different thinking patterns based on the test language (e.g., linguistic priming: Montero-Melis et al., 2016; language context: Kersten et al., 2010; Wang & Li, 2019; language of interference: Athanasopoulos et al., 2015), while other research suggests that bilinguals' cognitive behaviour is driven by the common linguistic pattern shared by both languages (Ameel et al., 2009; Filipović, 2020; Wang & Li, 2019). These varying findings may be attributed to factors such as the specific language structures under study, methodological approaches, and types of bilingualism (see Filipović, 2019; Lucy, 2016; Wang & Li, 2022b for detailed reviews).

While existing studies highlight the link between language and cognition, the extent to which verbal labels influence bilingual cognitive processing remains underexplored. Moreover, it is unclear how closely bilinguals' perception patterns resemble those of monolinguals in each language, and importantly, whether such convergence is influenced by different degrees of language involvement during task performance. As part of a series of studies on motion event cognition, our current study aims to bridge this gap by systematically manipulating different immediate language conditions (i.e., linguistic priming and verbal interference) to investigate how cross-linguistic variations in abstract grammatical features, especially motion verbs, influence the perception and categorization of similar events in early bilinguals of Cantonese (equipollent-framed) and English (satellite-framed). Investigating how bilinguals navigate the subtle typological differences between L1 and L2 provides new insights into the cognitive flexibility in bilingual minds. Furthermore, this study introduces reaction times (RTs) as a measure in similarity judgement tasks, offering a novel way to assess the implicit and automatic processing of motion concepts. Unlike categorical perception, which focuses on discrete classification, RTs capture fine-grained differences in cognitive effort, potentially revealing whether language-thought interface operates at a conscious, strategic level or an automatic, deeply ingrained level. This contributes to broader theoretical discussions on the nature of linguistic relativity effects.

Theories accounting for the language-thought interface: the label-feedback hypothesis

Regarding the impact of verbal labels on cognition, empirical evidence suggests that language and other perceptual cognitive processes (i.e., actions, visual, auditory, olfactory and other

attributes) are part of a distributed interactive system, where they interact and influence each other (Lupyan, 2012; Pulvermüller et al., 2005). For example, action-related words such as ‘pick’ and ‘kick’ stimulate both motor systems and semantic networks. Similarly, the label-feedback hypothesis posits that perceiving a stimulus automatically co-activates its corresponding verbal label, which then in turn affects perceptual features, such as colour, shape, texture or movement (Lupyan, 2012). For instance, seeing the colour blue automatically activates the verbal label ‘blue’ in a bottom-up manner. At the same time, hearing or thinking about the word blue can influence how we perceive the colour in a top-down process (Thierry et al., 2009; Winawer et al., 2007). The repeated act of assigning verbal labels to objects, actions, or sensory experiences reinforces the connection between language and perception. This process is mediated by the phonological loop – a cognitive mechanism that helps retain and process verbal information. Overt language use or linguistic priming can enhance the feedback between labels and categories, aiding perception and category formation based on specific linguistic features. Conversely, concurrent verbal interference might disrupt this loop and reduce label activation, potentially diminishing language’s impact on cognition. Illustratively, Winawer et al. (2007) found that Russian speakers, whose language distinguishes between lighter blues (*goluboy*) and darker blues (*siniy*), responded faster to different colour categories (light vs. dark blue) than within the same category (both light or both dark blue) when covert language use was allowed in colour discrimination. This processing advantage, however, vanished under verbal interference. This occurs because when observers are engaged in language to accomplish a verbal interference task (i.e., repeating nonsense words or strings of digits), their ability to rely on linguistic resources for colour perception and discrimination is reduced. This suggests that language can significantly influence perceptual decisions in real time, with its influence modulated by the presence or absence of verbal labels.

Motion event encoding and conceptual representations of motion

While the label-feedback hypothesis has been applied in the domains of colour (Lupyan et al., 2020; Winawer et al., 2007), grammatical gender (Sato et al., 2020), and grammatical aspect (Athanasopoulos & Bylund, 2013), our understanding of how the semantics of motion verbs influence information retrieval and category formation remains limited.

World languages describe motion events differently (Slobin, 2006; Talmy, 2000). For instance, in satellite-framed languages like German and English, the manner of motion (‘walk’) is typically encoded in the main verb, and the path of motion (‘up’) in a satellite. In contrast, verb-framed languages such as Spanish and French encode the path of motion in the main verb, with the manner either omitted or expressed in peripheral phrases (e.g., ‘A boy ascends a hill, running’).

While this binary typology of S- and V-framed languages has been successfully applied in analysing most Indo-European languages, it fails to accommodate serial-verb languages like Chinese and Thai, where manner and path are integrated into compound verb forms. For example, Cantonese is a serial-verb language and the serial-verb constructions in Cantonese can take up two or more components of equal grammatical status (Matthews & Yip, 2011). For instance, the path-of-motion 翻 (return) and 入 (enter) in (1) can appear together with manner 行 (walk) as a verb compound. Meanwhile, path components in (2) and (3) can also stand alone as single verbs, independent of manner elements. Therefore, serial-verb languages warrant recognition as a distinct third category, termed equipollently framed languages (E-languages), where both manner and path are expressed through elements of equal significance (Slobin, 2006):

- 1) 佢 行 咗 翻 入 教室
(Manner + Path compound verbs)
S/he walk ASP return enter classroom
'S/he walked back into the classroom'.
- 2) 佢 翻 咗 教室 (Manner-only construction)
S/he return ASP classroom
'S/he returned to the classroom'.
- 3) 佢 入 咗 教室 (Path-only construction)
S/he enter ASP classroom
'S/he entered the classroom'.

Recent studies show that Cantonese, due to its frequent use of path verbs, tends to omit manner-of-motion descriptions more often than S-languages like English (Matthews & Yip, 2011; Wang & Li, 2019; Yiu, 2014). At the conceptual level, speakers of S-languages tend to be more manner-oriented and respond more quickly in making manner-match decisions (Ji, 2019; Ji & Hohenstein, 2018), while E-language speakers are both manner- and path-oriented (Wang & Li, 2019, 2021, 2022a). These patterns align with Slobin's (2006) manner-salience hypothesis, which posits that spatial concepts in finite verbs are more readily accessible than those in nonfinite forms, due to their greater grammatical prominence. Thus, S-languages are characterized as 'high-manner-salient' because they typically encode manner information in main verbs, whereas E-languages are considered both 'manner-and-path salient', given their equal grammatical status.

Language-thought interface in early and late bilinguals

The influence of language on cognition is adaptable and context-dependent (Lupyan, 2012). For bilinguals, a critical question is how acquiring labels in a second language alters the way they categorize and interpret concepts. Research indicates that bilinguals exhibit distinct linguistic and cognitive characteristics. Moreover, variations in bilingualism (such as early vs. late acquisition, balanced vs. unbalanced proficiency, and simultaneous vs. sequential learning) significantly shape the integration of multiple languages within the bilingual mind (Filipović, 2019).

Kersten et al. (2010) studied how Spanish-English bilinguals categorize alien motion, focusing on the contrast between manner and path in English- or Spanish-based contexts. Using a triads-matching paradigm, they found that late bilinguals alternated their thought patterns depending on the test language, whereas early bilinguals consistently aligned with English monolinguals. Similarly, Lai et al. (2014) observed that late English-Spanish bilinguals, when tested in a Spanish context, were more likely to focus on path-of-motion for similarity judgements than those in an English context. However, early bilinguals showed uniform thinking patterns regardless of the test language. Wang and Li (2019) investigated how early Cantonese-English bilinguals categorized events using either only English or both Cantonese and English. They found that bilinguals uniformly adopted an English-based approach, regardless of the language prompts. Athanasopoulos et al. (2015) further reported that verbal interference has selective effects on perception and categorisation of abstract concepts for late bilinguals. It found that interference in Language A tends to induce Language-B congruent behaviours. When the verbal interference switches to Language B, participants shift their preferences to the uninterrupted language accordingly.

Together, these findings suggest that early bilinguals are more likely to develop a merged cognitive mode, integrating features of both languages 'to maximize common ground' (Filipović &

Hawkins, 2018). Late bilinguals, however, tend to use distinct mechanisms for language processing and are more likely to be influenced by immediate language contexts (Athanasopoulos et al., 2015; Montero-Melis et al., 2016).

The present study

This study examines how varying linguistic descriptions of motion affect the cognitive processing of motion in early Cantonese-English bilinguals under two scenarios: a verbal condition involving explicit language use (Experiment 1) and a nonverbal condition with verbal interference (Experiment 2). The specific research questions are as follows:

1. How do Cantonese and English monolinguals categorize and conceptualize motion events when their access to verbal labels ranges from maximal to minimal?
2. How do early Cantonese-English bilinguals categorize and conceptualize motion events when their access to verbal labels ranges from maximal to minimal? Do they exhibit a converged mode of thinking regardless of the language used, or do they adopt different thinking patterns aligned with each language?

Experiment 1 aims to test whether recent linguistic exposure modulates perception and cognition, and how bilinguals integrate both languages during cognitive processing. We hypothesize that when verbalization is available, English monolinguals will rely more on manner-of-motion for similarity judgements and respond more quickly to manner-match than their Chinese counterparts (Ji & Hohenstein, 2018; Wang & Li, 2021). In contrast, bilinguals are expected to develop a shared or converged representational system due to their early exposure to and active use of both languages. Experiment 2 examines how participants process event similarity when access to language is restricted through concurrent verbal interference. The purpose of verbal interference is to reduce the online influence of language and disrupt the mutual feedback between labels and perceptual regularities by engaging the phonological loop in competing language-related activities (Baddeley, 2003; Gennari et al., 2002; Lupyan, 2012). We hypothesize that when language access is restricted, if cross-linguistic patterns persist despite verbal interference, this would indicate that language effects on cognition are stable and deeply ingrained, supporting the linguistic relativity hypothesis. Conversely, if language effects weaken under verbal interference, this would support the label-feedback hypothesis, which suggests that language influences cognition in a flexible, context-dependent manner.

Method

Participants

One hundred and twenty university students participated in the experiment. Thirty Cantonese L1 speakers (M age = 20.8 years, SD = 0.9) and thirty English L1 speakers (M age = 21.8 years, SD = 2.6) were recruited from local universities of Guangdong Province, China and London, UK, respectively. Sixty Cantonese-English bilinguals (M age = 19.7, SD = 1.4) came from local universities of Hong Kong, where both Cantonese and English are official languages. According to the language policies of ‘biliteracy and trilingualism’ in Hong Kong, children typically begin learning English at an average age of three (M age = 3.0 years, SD = 1.3). The learning of English continues throughout their school years and many schools use English as the medium of instruction for other subjects in academic settings. Participants were required to complete the Oxford Placement Test (Version 2,

Oxford University Press, 2004). Based on their scores (M score = 54.9 out of 60; SD = 3.0; C2 level (54–60)), bilinguals' English proficiency was at an advanced level. The bilinguals' Cantonese proficiency was assessed using the Language History Questionnaire (LHQ 2.0, Li et al., 2014). Participants were asked to self-evaluate their Cantonese proficiency based on a 7-point Likert-type scale where 7 is the maximal rating (M score = 6.72; SD = 0.37). All participants reported using Cantonese and English interchangeably in daily activities and communication contexts.

Materials

The experimental stimuli in the verbal encoding task consisted of 54 six-second-long dynamic video clips, including 36 test items and 18 filler items. The test items depicted a boy performing a self-induced action with various types of manner and path (i.e., a boy walking down a hill) against different ground settings (e.g., in the forest, along the river, up the snow mountain), while the control items depicted another type of events, namely causative events (i.e., an agent performing a specific action on an object). Specifically, the test items varied in manner types, including general (e.g., walk, run), specific movements (e.g., crawl, swim), and use of instruments (e.g., cycling, skateboarding)-combined with diverse paths: vertical (e.g., up, down), deictic (e.g., towards, away from), and boundary-crossing (e.g., into, out of) (Hickmann & Hendriks, 2010).

Eighteen sets of animated videos were used as test triads (N = 12) and fillers (N = 6) in the similarity judgement task. The test triads shared the same stimuli with the verbal task. There were three dynamic events in each triad: a target event (e.g., A boy WALKING INTO a room) and its two alternates with manner and path as the contrast of interest. For instance, manner-consistent alternates had the same manner as the target but a different path (e.g., A boy walking OUT OF a room), while path-consistent alternates shared the target's path but differed in manner (e.g., A boy JUMPING into a room). The agent and background settings remain constant across three events within each triad. To obscure the contrast of interest (manner vs. path) from participants, 6 sets of fillers were included, contrasting ground settings with manner or path in three trials each.

Procedure

Following Wang and Li (2019, 2021, 2022a), participants in the linguistic encoding task first watched the stimuli and then immediately described 'what happened' in each clip after viewing. The stimuli were presented in a pseudo-randomized order. Monolinguals used their L1 languages. Sixty Cantonese-English bilinguals were randomly assigned to either a Cantonese-speaking or an English-speaking condition (N = 30 for each), where they used either Cantonese or English to describe the target events. To maintain a monolingual mode throughout the experiment, instructions were given in the same language as that used for descriptions. The bilinguals in both contexts demonstrated comparable English proficiency, as indicated by their average Oxford Placement Test (OPT) scores: M = 54.90, SD = 2.46 in the Cantonese context and M = 54.39, SD = 3.50 in the English context.

To maximally boost language involvement, Experiment 1 required participants to describe events in the target language before proceeding to a nonlinguistic categorization task (Gennari et al., 2002; Montero-Melis et al., 2016). In this task, participants viewed three animated clips presented in a synchronized sequence. The target clip played first at the bottom of the screen and disappeared immediately after completion. Following a 500 ms black screen, the two manner- and path-match variants began playing simultaneously at the left and right top of the screen. The positioning of these variants was counterbalanced across stimuli in a fixed order. Participants were instructed to quickly determine which variant most closely resembled the target event by pressing

the designated keys (A or L), without the need to watch the alternate videos to completion. Upon pressing one of the relevant keys, the videos disappeared, followed by a 1000ms black screen before the next set of clips started to play.

Data coding and analysis

The linguistic data were coded by L1 speakers of Cantonese and English, and subsequently segmented into clauses (i.e., the unit that contains finite verbs) based on the coding guidelines of English (Hickmann & Hendriks, 2010) and Cantonese (Wang & Li, 2019). The coding mainly focused on (1) how frequently the target elements (i.e., manner and path) were expressed, and (2) where the target elements were expressed (i.e., verbs or satellites). Only test items were included in the analysis. Descriptions without a specific mentioning of motion elements were removed (e.g., The sky is blue). To establish the coding reliability, 15% of the data were randomly selected and double coded by a second coder. As indicated by the Kappa Index, a high inter-coder reliability was reached for both the Cantonese (Cohen's $\kappa = .95$) and English (Cohen's $\kappa = .98$) dataset.

The nonlinguistic data in the similarity task included categorization preferences and reaction times. Participants' reaction times to each triad was calculated from the onset of the alternate videos till the point when participants made their similarity judgements. Response data of extremely long or short values were replaced with 2 standard deviations (SDs) from the group mean. In Experiment 1, 3 out of 1,440 observations were missing due to technical issues. Altogether 54 out of 1,437 outliers (3.75%) were identified and replaced in the final dataset.

Results and discussion

The linguistic encoding task

A total of 4,316 target linguistic descriptions were involved in the final analysis. Subsequent logistic mixed-effects regression models were fitted to test how frequently manner and path were expressed as a function of language groups. All analyses were run in R (R Core Team, 2021) using the *lme4* package (Bates et al., 2015). In each model, we computed the encoding of manner or path (presence=1; absence=0) as the respective binary dependent variable.¹ The fixed effect was participant groups and the random effects included random intercepts for subjects and items.

For path expressions, computing language group as a fixed effect did not significantly improve the model fit compared with the null model², $\chi^2(3)=3.619$, $p=.306$, showing that language group was not a significant contributor. In fact, participants across different groups had a high proportion of path encoding (Cantonese monolinguals: $M=95.93\%$; $SD=4.99$; Bilinguals in Cantonese context: $M=94.72$; $SD=5.57\%$; Bilinguals in English context: $M=94.26\%$; $SD=4.61\%$; English monolinguals: $M=93.52$; $SD=6.12\%$), indicating that path is a core element (Talmy, 2000). However, the frequency of manner expressions exhibited language-specific patterns (Cantonese monolinguals: $M=79.04\%$; $SD=10.76\%$; Bilinguals in Cantonese context: $M=93.24$, $SD=9.64\%$; Bilinguals in English context: $M=97.50\%$, $SD=2.97\%$; English monolinguals: $M=96.57\%$, $SD=4.48\%$). Specifically, bilinguals in a Cantonese context expressed significantly more manner-of-motion compared with Cantonese monolinguals ($\beta_0=2.45$, $SE=0.43$, $z=5.66$, $p<.001$), but patterned with bilinguals in an English context ($\beta_0=1.13$, $SE=0.48$, $z=2.76$, $p=.2$) and English monolinguals ($\beta_0=0.96$, $SE=0.47$, $z=2.04$, $p=.08$).

In addition, the semantic distribution of manner- and path-of motion across each group adhered to language-specific patterns. English speakers and bilinguals using English predominantly used verb-satellite constructions to encode manner via the main verb (English: $M=92.13\%$; $SD=8.13\%$;

Bilinguals in English: $M=95.09\%$; $SD=5.19\%$), and path through the satellite (English: $M=85.37\%$; $SD=9.62\%$; Bilinguals in English: $M=89.26\%$; $SD=6.87\%$), as exemplified in (4).

In comparison, Cantonese monolinguals and bilinguals in a Cantonese context employed serial-verb-constructions and lexicalized manner (Cantonese: $M=56.85\%$; $SD=12.63\%$; Bilinguals in Cantonese: $M=71.85\%$; $SD=11.78\%$) and path (Cantonese: $M=55.65\%$; $SD=12.56\%$; Bilinguals in Cantonese: $M=45.24\%$; $SD=13.06\%$) in verb compounds, as shown in (5):

(4) A boy is jumping _[manner verb] away _[path satellite] from a river.

(5) 一男仔 跳 咗 落 去
 A boy jump ASP descend go
 'A boy jumped down (from the stairs) and left'.

Similarity judgement task

To further explore whether cross-linguistic differences in language impact one's thinking patterns, mixed-effects models were run to compare participants' categorical preferences and reaction times in making manner or path decisions. For the categorical preferences (Figure 1), we compared the mean proportion of manner-match preferences across different groups (Cantonese monolinguals: $M=46.94\%$, $SD=24.42\%$; Bilinguals in a Cantonese context: $M=61.39\%$, $SD=24.16\%$; Bilinguals in an English context: $M=63.75\%$, $SD=23.43$; English monolinguals: $M=69.44\%$, $SD=25.34\%$).

A logistic mixed-effects model³ was computed with participants' categorical choice as the binary dependent variable (1 = manner-match variant; 0 = path-match variant), and language group (four levels) as the fixed effect (dummy coded). The random effects included random intercepts for participants and items. Likelihood ratio tests were performed using *anova()* function to evaluate model fits. The results showed that adding language group in the model as a fixed effect significantly improved the model fit compared with the null model, $\chi^2(3)=33.71$, $p<.001$, indicating that language group was a main effect. We used forward coding to compare the grand mean of the log-likelihood of manner-match preference with the next group. It showed that bilinguals in a Cantonese context selected significantly more manner-match variants than Cantonese monolinguals ($\beta_0=-0.662$, $SE=0.161$, $z=4.115$, $p<.0001$), but patterned with bilinguals in an English context ($\beta_0=-0.165$, $SE=0.181$, $z=-0.920$, $p=.363$) and English monolinguals ($\beta_0=-0.514$, $SE=0.273$, $z=-1.887$, $p=.079$).

For the processing efficiency (Table 1), we used a forward stepwise selection for models and the most optimal linear mixed-effects model⁴ computed preference type (manner or path variants), language group and their interaction as fixed effects. The random effects included crossed random intercepts for participants and items. To satisfy the normality assumptions, the dependent variable RTs were log-transformed (Baayen et al., 2008).

The results (Table 2) showed that *preference type* was a significant predictor, indicating that participants responded faster to making manner-match decisions. Critically, a significant interaction was found between *preference type* and *language group*, indicating that participants in different groups had different RT patterns when responding to manner- and path-match variants.

To further examine the effect of *language group* on RTs, we built four separate mixed-effects models⁵ for within-group comparisons. For all models, the dependent variable was the log-transformed RTs. The fixed effect was preference type. The random effects were random intercepts for participants and items. Cantonese monolinguals had similar RTs when selecting manner- and

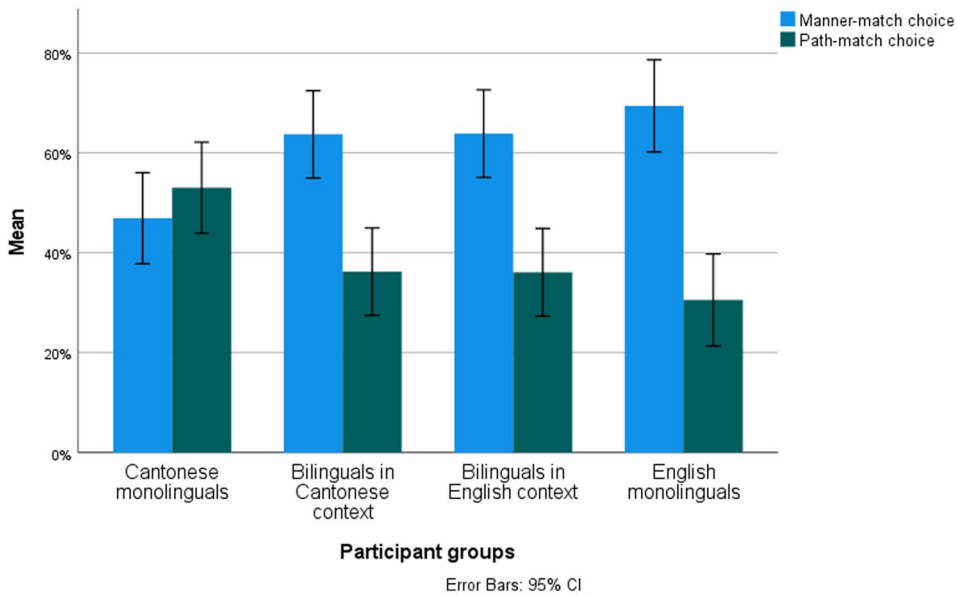


Figure 1. Mean proportion of manner- and path-match variants across different groups in verbal encoding condition.

Table 1. Reaction times to making manner- and path-match decisions.

Group	Mean RT to manner	Mean RT to path
Cantonese monolinguals	2058 (687)	2148 (646)
Bilinguals in a Cantonese context	1960 (646)	2126 (839)
Bilinguals in an English context	1842 (565)	2169 (531)
English monolinguals	1973 (653)	2239 (684)

path-match variants ($\beta_0 = -0.04$, $SE = 0.031$, $t = -.317$, $p = .189$). However, bilinguals in both language contexts (Cantonese context: $\beta_0 = -0.109$, $SE = 0.036$, $t = -3.045$, $p = .002$; English context: $\beta_0 = -0.202$, $SE = 0.033$, $t = -6.183$, $p < .001$), and English monolinguals ($\beta_0 = -0.167$, $SE = 0.037$, $t = -4.477$, $p < .001$) responded more quickly to manner-match variants than path-match variants. These cross-linguistic differences suggest that bilinguals followed an English-based way when processing manner- and path variants, irrespective of the language in operation.

Experiment 2: similarity judgements under verbal interference

Experiment 2 further explores whether cross-linguistic differences in the cognition of motion events persist in strictly nonverbal contexts. If the same patterns are observed, it suggests that participants may not rely on linguistic resources for their similarity judgements. Alternatively, this could indicate that linguistic resources permanently shape one's categorization behaviour, a warping that is impervious to verbal interference.

Table 2. Coefficients for the mean RTs to manner-and path-match choices in verbal encoding condition.

Fixed effects	Estimate	SE	t value	p
<i>Intercept</i>	7.594	0.035	214.362	< .001***
Preference type	−0.078	0.035	−2.253	0.024*
Language group	0.046	0.040	1.133	0.258
Bilinguals in English context	0.031	0.033	0.893	0.372
Cantonese monolinguals	0.088	0.053	1.625	0.099
English monolinguals	−0.112	0.049	−2.270	0.020*
Preference type * Language group	0.023	0.048	0.467	0.640
Manner*Bilinguals in English	−0.083	0.052	−1.609	0.108
Manner* Cantonese				
Manner* English				
<i>Random effects</i>	Variance		SD	
Participant (Intercept)	0.018		0.135	
Item (Intercept)	< .001		0.023	
Residual	0.093		0.305	

* $p < .05$. *** $p < .01$.

Method

Participants

One hundred and eight different university students participated in the experiment: twenty-seven Cantonese L1 speakers (M age=22.3 years, SD =1.8), twenty-five English L1 speakers (M age=20.7 years, SD =2.3) and fifty-six Cantonese-English bilinguals (M age=19.9 years, SD =3.1).

Materials and procedure

Experiment 2 utilized the same materials as Experiment 1, with the exception that participants were required to verbally repeat a sequence of two-digit random numbers in either Cantonese or English. This dual-task paradigm was designed to inhibit the subconscious use of language in cognitive processing (Trueswell & Papafragou, 2010). Monolinguals used their first language, while bilinguals were randomly assigned to one of two conditions: one half repeated the numbers in Cantonese, and the other half in English. Participants in both groups exhibited equal English proficiency, as measured by the OPT scores: M =53.98, SD =3.01 in the Cantonese interference language context and M =54.37, SD =2.78 in the English context.

Results and discussion

It was found that language-specific patterns in event categorization disappeared (Figure 2). Instead, participants in each language group showed an overall preference for path-match variants (Manner-match choices: Cantonese: M =43.52%, SD =21.35; Bilinguals with Cantonese interference: M =41.96%, SD =26.39%; Bilinguals with English interference: M =46.43%, SD =28.55%; English: M =37.32%, SD =27.16%). Further analysis⁶ confirmed that the participant group was not a significant contributor in the model, $\chi^2(3)$ =5.219, p =.156, indicating that after the verbal interference was introduced, participants' reliance on language was disrupted accordingly.

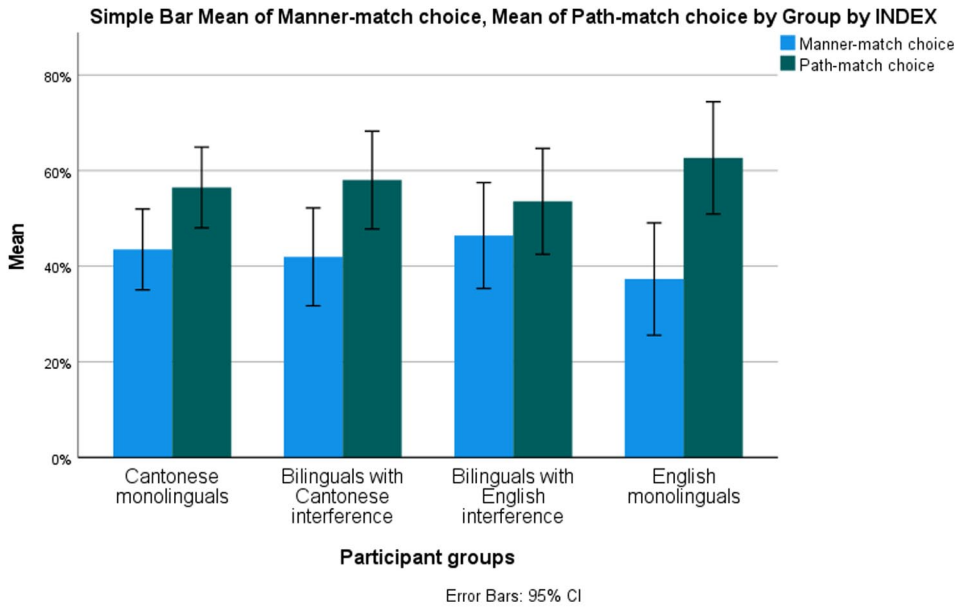


Figure 2. Mean proportion of manner- and path-match variants across different language groups with verbal interference.

Table 3. Mean RTs to manner- and path-match choices in verbal interference condition.

Group	Mean RT to manner	Mean RT to path
Cantonese monolinguals	2221 (676)	2342 (708)
Bilinguals with Cantonese interference	2112 (539)	2420 (703)
Bilinguals with English interference	2311 (451)	2519 (667)
English monolinguals	2214 (487)	2491 (795)

In terms of the RTs, altogether 17 out of 1,296 outliers (1.31%) were identified and replaced with ± 2 SDs from the grand mean in the final dataset. Table 3 presents participants' mean response times to manner- and path-match variants. A logistic mixed-effects model⁷ was built with the log-transformed RT as the dependent variable. The fixed effects included language group, preference types, and their interaction. The results showed a significant interaction between *preference type* and *language group* ($\beta_0 = 0.124$, $SE = 0.045$, $t = 2.725$, $p = .006$), indicating that participants in different groups had different RT patterns when responding to manner- and path-match variants.

Specifically, three separate mixed effect models were built to address the within-group differences in the processing efficiency of manner- and path-match variants. Results confirmed that Cantonese monolinguals showed comparable RTs with manner- and path-match variants ($\beta_0 = 0.013$, $SE = 0.033$, $t = 0.412$, $p = .68$), while bilinguals with Cantonese interference ($\beta_0 = -0.137$, $SE = 0.032$, $t = -4.266$, $p < .001$) and with English interference ($\beta_0 = -0.083$, $SE = 0.027$, $t = -3.066$, $p < .001$) patterned with English monolinguals ($\beta_0 = -0.124$, $SE = 0.035$, $t = -3.491$, $p < .001$) in responding significantly faster to manner-match variants, indicating that RTs are more resistant to verbal interference than categorical preferences.

General discussion

The current study aims to investigate how monolinguals and bilinguals of Cantonese and English gauge and process event similarity when their access to language ranges from maximal to minimal.

Experiment 1 revealed that English speakers significantly emphasized manner-of-motion and predominantly adopted the ‘manner verb + path satellite’ syntactic framing strategy. In contrast, Cantonese speakers frequently utilized both ‘manner-path compounds’ and ‘path-only’ constructions, reflecting the extensive use of path verbs in Cantonese. These findings support previous research on the lexicalization patterns of motion events in English and Cantonese (Wang & Li, 2019, 2021, 2022a; Yiu, 2014) and corroborate Slobin’s (2006) tripartite motion event typology.

At the conceptual level, English speakers tended to rely more on manner-of-motion and responded faster when making manner choices given its prominent verb status. Conversely, Cantonese speakers paid roughly equal attention to manner and path components and had comparable response time with manner and path decisions. These observations align with Slobin’s (2006) theory of manner salience, which posits that the cognitive prominence of manner is closely linked to its codability in lexicalization. Since English lexically encodes manner through finite verbs, the high codability of manner contributes to its greater cognitive salience and enhances its accessibility during processing. This ease of access might lead to a ‘sequential processing strategy’ in category formation (Ji, 2019), thereby shortening the response times for manner choices among English speakers. Cantonese, as a serial verb language, encodes manner and path equally in verb compounds. This allows Cantonese speakers to process and retrieve manner and path information ‘in parallel’ when language use is permitted during tasks (Wang & Li, 2021). Such findings support the label-feedback hypothesis (Lupyan, 2012), suggesting that the activation of verbal labels underpins the top-down influence of language on cognition. Specifically, eliciting linguistic descriptions before stimuli can enhance the interaction between labels and categories, promoting language-influenced cognitive behaviour on an ad hoc basis (Lupyan et al., 2020; Montero-Melis et al., 2016).

Data from Cantonese-English bilinguals demonstrate that, independent of the test language, bilinguals exhibited English-based processing in both tasks. They more frequently encoded manner when describing motion and showed a preference for manner-match variants in categorization tasks. In addition, their RTs were quicker for manner-match than for path-match variants, indicating an English-biased pattern. Notably, the bilinguals in this study were early bilinguals with high proficiency in both Cantonese and English, which likely encouraged the development of a unified representational system (Filipović & Hawkins, 2018; Kersten et al., 2010). Given that Cantonese can interchangeably use ‘manner + path’ and ‘path only’ constructions, while English consistently prefers ‘manner + path’, merging these systems into a single lexicalization pattern that aligns with both languages could reduce cognitive load. This approach exploits typological similarities to streamline bilingual language processing. This observation is consistent with findings that different bilingualism types influence language’s impact on cognition (Filipović, 2020). Specifically, early bilinguals tend to utilize a shared neural substrate for processing information, in contrast to late bilinguals, who rely on distinct neural substrates (Weber-Fox & Neville, 1999). This functional distinction may shield early bilinguals from the effects of short-term verbal mediation (Kersten et al., 2010; Lai et al., 2014; Wang & Li, 2019).

To further investigate if the cross-linguistic differences observed persist in a strictly nonverbal context, Experiment 2 implemented a dual-task paradigm with verbal interference, revealing an overall preference for path-match alternatives. This may suggest that the path of motion is a fundamental, universal component in the construction of motion events, which could account for the consistent emphasis on path across different languages (Slobin, 2006). Conceptually, it appears

that path may be prioritized over manner universally, supporting the Path Salience hypothesis (Talmy, 2000). When verbal interference blocks or suppresses participants' access to language, they cannot rely on previous labelling experiences as a strategy for problem-solving. Consequently, a universal conceptual foundation might override the effects of linguistic relativity, neutralizing language-specific observations (Ji & Hohenstein, 2018; Kersten et al., 2010). According to the label-feedback hypothesis, the interaction between labels and categories is facilitated by a phonological loop, essential for the verbal rehearsal of visual stimuli and category formation (Baddeley, 2003). However, when this loop is engaged in other verbally mediated tasks, it interrupts the feedback between stimuli and their labels, diminishing the influence of language on cognition (Lupyan, 2012; Lupyan et al., 2020). This indicates that categorical judgements, which are verbally mediated, can be affected by engaging in a verbal interference task (Wolff & Holmes, 2011, p. 256). The findings suggest that language dynamically interacts with cognition, reinforcing the idea that verbal labels can influence perceptual and categorization processes in real time.

Unlike categorical judgements, participants' processing efficiency still exhibited robust language-specific patterns. This discrepancy might be due to the nature of the measurements: categorical judgements are higher-level, explicit processes that are post-perceptual and typically occur at a later stage. In contrast, lower-level processes, such as reaction times, are automatic, subconscious, and generally happen in earlier stages (Athanasopoulos et al., 2015; Thierry et al., 2009). Reaction time is a nuanced metric that captures participants' automatic, nonreflective, and implicit responses to perceptual tasks (Tokowicz & MacWhinney, 2005), making it a valuable tool for detecting language-specific effects on simple, subconscious, and lower-level cognitive processes (Flecken et al., 2015; Ji & Hohenstein, 2018; Thierry et al., 2009; Winawer et al., 2007). As noted earlier, the typological differences between Cantonese and English may not significantly influence participants' categorical decisions in scenarios where linguistic resources are unavailable for category formation. However, examining reaction times can uncover subtle distinctions not immediately apparent in categorical data alone (Ji & Hohenstein, 2017, 2018). The findings align with prior research demonstrating that language effects emerge under specific conditions rather than applying uniformly across all cognitive tasks (Athanasopoulos et al., 2015; Bylund & Athanasopoulos, 2017).

Interestingly, when it comes to early Cantonese-English bilinguals, we found no significant variations in cognitive performance related to the language of interference, whether it was Cantonese or English. These bilinguals displayed similar processing efficiency to English monolinguals, indicating a unified, English-influenced cognitive approach regardless of the language in use. Although our findings differed from Athanasopoulos et al. (2015), it is worth highlighting that the bilinguals in the current study are early bilinguals with regular exposure to both Cantonese and English from a young age. They likely develop a shared neural basis for information processing, differing from late bilinguals in functional language organization and cognitive processing outcomes (Lai et al., 2014; Weber-Fox & Neville, 1999). Thus, it is important to consider the timing of bilingual exposure when examining language's influence on cognitive function (Kersten et al., 2010).

Conclusions

To conclude, this research advances our understanding of the interplay between language and thought, particularly regarding the influence of language on the conceptualization of motion. It confirms that variations in motion verb usage across languages shape mental representations of motion events, aligning with theories that verbal labels activate associated features through top-down processes (Bylund & Athanasopoulos, 2017; Lupyan, 2012). The study also sheds light on

differences between explicit and implicit event cognition processes, highlighting the impact of stimulus characteristics, linguistic structures, and task requirements on these discrepancies. Furthermore, these results support the nuanced perspective of linguistic relativity, suggesting that language effects on thought are complex and multidimensional, rather than a simple all-or-nothing phenomenon (Bylund & Athanasopoulos, 2017). To further explore the top-down modulation of language on different cognitive processes, future studies can use diverse behavioural and neuroscientific methods with diverse bi- and multilingual populations and language pairs. These would broaden linguistic relativity research and capture a more nuanced picture of language's impact on cognition at different processing stages.

ORCID iDs

Yi Wang  <https://orcid.org/0000-0003-4175-7467>

Li Wei  <https://orcid.org/0000-0002-2015-7262>

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The article was prepared in the framework of a research grant funded by the Ministry of Science and Higher Education of the Russian Federation (Grant ID: 075-15-2020-928) and the ESRC Postdoctoral Fellowship (ES/V012274/1).

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

1. `model1 <- glmer(Manner ~ Language group + (1|Subject) + (1|Item), family=binomial, data=data);`
`model2 <- glmer(Path ~ Language group + (1|Subject) + (1|Item), family=binomial, data=data)`
2. `model 3 <- glmer(Path ~ (1|Subject) + (1|Item), family=binomial, data=data)`
3. `model 4 <- glmer(Manner choice ~ Language group + (1|Subject) + (1|Item), family=binomial, data=data)`
4. `model 5 <- lmer(log(RT) ~factor(preference type) * Language group + (1|Subject) + (1|Item), data=data)`
5. `model 6 <- lmer(log(RT) ~factor(preference type) + (1|Subject) + (1|Item), data=data)`
6. `model7 <- glmer(Manner choice ~Language group + (1|Subject) + (1|Item), family=binomial, data=data)`
7. `model8 <- lmer(log(RT) ~factor(preference type)*Language group + (1|Subject) + (1|Item), data=data)`

References

- Ameel, E., Malt, B. C., Storms, G., & Van Assche, F. (2009). Semantic convergence in the bilingual lexicon. *Journal of Memory and Language*, 60(2), 270–290.
- Athanasopoulos, P., & Bylund, E. (2013). Does grammatical aspect affect motion event cognition? A cross-linguistic comparison of English and Swedish speakers. *Cognitive Science*, 37(2), 286–309.
- Athanasopoulos, P., Bylund, E., Montero-Melis, G., Damjanovic, L., Scharfner, A., Kibbe, A., & Thierry, G. (2015). Two languages, two minds: Flexible cognitive processing driven by language of operation. *Psychological Science*, 26(4), 518–526.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, 59(4), 390–412.

- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829–839.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.
- Bylund, E., & Athanasopoulos, P. (2017). The Whorfian time warp: Representing duration through the language hourglass. *Journal of Experimental Psychology: General*, 146(7), 911–916.
- Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition*, 106(2), 579–593.
- Filipović, L. (2019). *Bilingualism in action: Theory and practice*. Cambridge University Press.
- Filipović, L. (2020). Bilingual memory advantage: Bilinguals use a common linguistic pattern as an aid to recall memory. *International Journal of Bilingualism*, 24(3), 542–555.
- Filipović, L., & Hawkins, J. (2018). The complex adaptive system principles model for bilingualism: Language interactions within and across bilingual minds. *International Journal of Bilingualism*, 23(6), 1223–1248.
- Flecken, M., Athanasopoulos, P., Kuipers, J. R., & Thierry, G. (2015). On the road to somewhere: Brain potentials reflect language effects on motion event perception. *Cognition*, 141, 41–51.
- Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, 83(1), 49–79.
- Goldstone, R. L., & Hendrickson, A. T. (2010). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(1), 69–78.
- Hickmann, M., & Hendriks, H. (2010). Typological constraints on the acquisition of spatial language in French and English. *Cognitive Linguistics*, 21(2), 189–215.
- Ji, Y. (2019). Cognitive representation of spontaneous motion in a second language: An exploration of Chinese learners of English. *Frontiers in Psychology*, 10, 2706–2728.
- Ji, Y., & Hohenstein, J. (2018). English and Chinese children's motion event similarity judgments. *Cognitive Linguistics*, 29(1), 45–76.
- Kersten, A. W., Meissner, C. A., Lechuga, J., Schwartz, B. L., Albrechtsen, J. S., & Iglesias, A. (2010). English speakers attend more strongly than Spanish speakers to manner of motion when classifying novel objects and events. *Journal of Experimental Psychology: General*, 139(4), 638–653.
- Lai, V. T., Rodriguez, G. G., & Narasimhan, B. (2014). Thinking-for-speaking in early and late bilinguals. *Bilingualism: Language and Cognition*, 17(1), 139–152.
- Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language History Questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, 17(3), 673–680.
- Lucy, J. A. (2016). Recent advances in the study of linguistic relativity in historical context: A critical assessment. *Language Learning*, 66(3), 487–515.
- Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology*, 3, 54–67.
- Lupyan, G., Rahman, R. A., Boroditsky, L., & Clark, A. (2020). Effects of language on visual perception. *Trends in Cognitive Sciences*, 24(11), 930–944.
- Matthews, S., & Yip, V. (2011). *Cantonese: A comprehensive grammar*. Routledge.
- Montero-Melis, G., Jaeger, T. F., & Bylund, E. (2016). Thinking is modulated by recent linguistic experience: Second language priming affects perceived event similarity. *Language Learning*, 66(3), 636–665.
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108(1), 155–184.
- Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, 21(3), 793–797.
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Sato, S., Casaponsa, A., & Athanasopoulos, P. (2020). Flexing gender perception: Brain potentials reveal the cognitive permeability of grammatical information. *Cognitive Science*, 44(9), e12884.
- Slobin, D. I. (1996). From 'thought and language' to 'thinking for speaking'. In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge University Press.

- Slobin, D. I. (2006). What makes manner of motion salient? Explorations in linguistic typology, discourse, and cognition. In M. Hickmann & S. Robert (Eds.), *Space in languages: Linguistic systems and cognitive categories* (pp. 59–81). John Benjamins.
- Talmy, L. (2000). *Toward a cognitive semantics*. MIT Press.
- Thierry, G. (2016). Neurolinguistic relativity: How language flexes human perception and cognition. *Language Learning*, 66(3), 690–713.
- Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J. R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences*, 106(11), 4567–4570.
- Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in Second Language Acquisition*, 27(2), 173–204.
- Trueswell, J. C., & Papafragou, A. (2010). Perceiving and remembering event cross-linguistically: Evidence from dual-task paradigms. *Journal of Memory and Language*, 63(1), 64–82.
- Wang, Y., & Li, W. (2019). Cognitive restructuring in the bilingual mind: Motion event construal in early Cantonese-English bilinguals. *Language and Cognition*, 11(4), 527–554.
- Wang, Y., & Li, W. (2021). Cognitive restructuring in the multilingual mind: Language-specific effects on processing efficiency of caused motion events in Cantonese-English-Japanese speakers. *Bilingualism: Language and Cognition*, 24(4), 1–16.
- Wang, Y., & Li, W. (2022a). Multilingual learning and cognitive restructuring: The role of audiovisual media exposure in Cantonese-English-Japanese multilinguals' motion event cognition. *International Journal of Bilingualism*, 27(3), 331–348.
- Wang, Y., & Li, W. (2022b). *Thinking and speaking in a second language. Elements in second language acquisition*. Cambridge University Press.
- Weber-Fox, C. M., & Neville, H. J. (1999). Functional neural subsystems are differentially affected by delays in second language immersion: ERP and behavioral evidence in bilinguals. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 23–38). Lawrence Erlbaum Associates, Inc.
- Whorf, B. L. (1956). *Language, thought, and reality*. MIT Press.
- Winawer, J., Nathan, W., Michael, C. F., Lisa, W., Alex, R. W., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, 104(19), 7780.
- Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(3), 253–265.
- Yiu, C. (2014). Directional verbs in modern Cantonese: A typological perspective. In C. Yiu (Ed.), *The typology of motion events: An empirical study of Chinese dialects* (pp. 33–135). De Gruyter Mouton.

Author biographies

Dr. **Yi Wang** is a Lecturer of English Language and Linguistics at Cardiff University. She received her Ph.D. from University College London. Her research expertise lies in language and cognition, motion event processing, and experimental psychology.

Professor **Li Wei** is Director and Dean of the UCL Institute of Education at University College London. He works on different aspects of bilingualism and multilingualism. He is currently the Editor of the International Journal of Bilingual Education and Bilingualism and Applied Linguistics Review.