

Vtce{ Ni

The linguistic Effects of Motion Event Typology

Abstract: This study tests the hypothesis of linguistic relativity along two lines of research: a) how L2 learners of Chinese and English, respectively, syntactically package semantic components for caused motion (cause, manner, path) in an experimental situation in which they are asked to describe video clips showing caused motion events to an imagery addressee, and b) how monolingual native speakers of Chinese and English judge the similarity between caused motion scenes while viewing them. Our results regarding a) show that Chinese learners of English acclimate to the target pattern of organizing particularly dense caused motion information very rapidly, and English learners of Chinese also arrived at an inter-language showing considerable resemblance to the target system rather than traces of the L1 influence. Our findings regarding b) reveal that despite striking differences between Chinese and English in L1 motion descriptions, native speakers show an identical tendency to prefer the path–match alternate over the manner–match alternate. Overall, these observations suggest that language–specific constraints can be largely shaken off when encoding caused motion in a non–native language, and linguistic and non-linguistic representations of caused motion may be dissociable from each other.

Key words: L2 acquisition of caused motion expressions; linguistic relativity; motion event typology; similarity judgment task

1 Introduction

In the last three decades the domain of space has increasingly attracted the attention of researchers in varied disciplines of language and cognition. Two factors, among other things, contribute to this phenomenon. First of all, despite space being universal because spatial concepts regarding entities and relations among them are constrained by our biological endowment, the visual and the haptic–kinesthetic system in particular, the linguistic representation of one’s spatial experience does not closely mirror the contours of nonlinguistic spatial understanding. Rather, an increasing number of recent studies reveal that languages differ strikingly in how they describe space (e.g. the motion event typology as proposed by Talmy 1985 2000). These specific properties of space make it an ideal testing ground for the relation between language and thought, for instance, the Whorfian hypothesis of linguistic relativity. Further, the particular ways in which languages differ in the linguistic encoding of spatial events can be psychologically tested. Language–independent, objective means are available to measure discrimination for spatial conceptualization: examples include gestures, categorization tasks, recognition memory and similarity judgments. In this context, the present study aims to explore the effect of language typology on representation of space in two types of tasks: a) a linguist task (Study A) in which second language (L2) learners of Chinese and English, respectively, describe the semantic dimensions of caused motion events (e.g. A boy pushed [manner and cause of motion] a treasure bag up [path of motion] the pyramid), and b) a non–linguistic task (Study B) in which monolingual native speakers of Chinese and English judge the similarity between motion scenes while viewing short video clips illustrating caused motion events.

As proposed by Talmy (1985 2000), world languages mainly fall into two major categories with respect to motion event descriptions, depending on in which grammatical element the most fundamental semantic dimension of motion, viz., path of motion, is systematically encoded: in verb roots or in verb–supporting elements (i.e. satellites). Most Germanic languages (e.g. English, German) are classified as satellite–framed in which path is characteristically expressed in particles (e.g. A boy kicked the football up the hill), whereas most Romance languages (e.g. French, Spanish) are considered to be verb–framed in encoding path of motion in the marked grammatical category of verbs (e.g. ‘A boy make ascend the football uphill by kicking’). As regards the two languages under discussion, English is unanimously categorized as satellite–framed. The exact typological status of Chinese, on the other hand, was once a

disputable topic (i.e. satellite-framed or not). However, an increasing number of recent studies as well as the revised version of old claims seem to suggest that Chinese is better considered as a third category of languages standing midway along the satellite- vs. verb-framing continuum (Ji et al. 2011a, 2011b), a language with parallel framing systems in motion descriptions, viz., both satellite- and verb-framed (Talmy 2009), or an ‘equipollently-framed’ language in which multiple semantic components of motion can be simultaneously packaged via elements of equal grammatical status and formal significance, for instance, compound verbs (Slobin 2004; e.g. *ti1-shang4* ‘kick-ascend’).

2 Caused Motion Expressions in Adults’ L2 Acquisition

Previous investigations reveal that caused motion expressions in English are characteristically encoded in a ‘manner and cause verb + path satellite’ combination (example 1), whereas in Chinese two specific grammatical constructions are normally recruited, along with a Resultative Verb Compound (RVC) to express the specific type of caused motion events in the current study, namely, BA and ZHE constructions, as illustrated in example 2:

- (1) He dragged the treasure bag into the pyramid.
 (2) a. *Ta1 ba3 cai2bao3 dai4 tuo1-jin4 jin1zi4ta3*
 he BAF1F treasure bag drag-enter pyramid
 ‘He dragged the treasure bag into the pyramid’
 b. *Ta1 tuo1 zhe cai2bao3 dai4 jin4 / zou3-jin4 le*
 he drag ZHE treasure bag enter / walk-enter ASP
jin1zi4ta3
 pyramid
 ‘Dragging the treasure bag, he went into / walked into the pyramid’

Compared to the vast bulk of literature on L1 acquisition of motion expressions, very few studies have systematically investigated the acquisition of motion descriptions in an L2 context, particularly in relation to Chinese and English. The key issue explored centers around whether and how the typological properties of L1 constrain the rhythm and the particular way various semantic components of motion are reorganized and mentally construed when encoding motion events in a non-native language.

As regards the acquisition of English motion descriptions by Chinese adult learners, relevant studies (e.g. Zeng 2011) revealed that although L2 learners seemed to generally acquire the characteristic way of encoding motion events in English, viz., the ‘verb + satellite’ combination, their descriptions still differed from that of English native speakers and showed clear influence from L1 in two aspects in particular: a) they used motion verbs that were much more limited in number and much less specific in nature as compared to native speakers of English, and b) they tended to provide descriptions of physical settings of motion events more frequently than did English native speakers, a discourse feature that learners inherited from their L1 Chinese and carried over into their L2 English.

With respect to the acquisition of Chinese motion expressions by English adult learners, previous investigations suggested that the acquisition of the BA construction did not seem to pose particular challenges to L2 learners. Du’s (2006) study, for instance, showed that English learners of Chinese acquired a fairly good knowledge of the complexity constraint on the post-ba verb, that is, in most cases where they produced a BA sentence, they used, as is required, either a verb compound or a single verb with the perfective aspectual marker *le*. Our survey of literature on the acquisition of the ZHE construction showed that relevant studies are virtually absent. However, we believe that the acquisition of this construction may pose particular difficulties to L2 learners of Chinese due to the following semantic and syntactic constraints associated with this construction: a) the construction involves syntactic subordination despite that the two events concerned are presented as occurring simultaneously; and b) the first verb (V1) is syntactically subordinate to the second verb (V2) although it is semantically more salient (see Ji and Hohenstein 2014b for details). To illustrate, in (2b) above the event of ‘dragging the treasure bag’ and that of ‘walking into the pyramid’ co-extend along the temporal axis. However, the former is syntactically subordinate to the latter although it constitutes the semantic focus of the sentence.

3 Representation of Spatial Events at the Cognitive Level

The impact of language-specific factors on the acquisition of spatial language and the organization of spatial information is actually found in diverse sub-domains of space. Some scholars hold that the influence of language is limited to the level of linguistic expression only (e.g. Gennari et al. 2002, Papafragou et al. 2002). In this sense, spatial concepts are universal across languages and diversity in the linguistic representation of a motion event merely reflects different instantiations of a common conceptual framework. In contrast, other researchers argue that language-specific properties can even shape our spatial cognition at a deeper level (see, for instance, Bowerman 1999, Bowerman and Choi 2001, Hohenstein 2005). That is, conceptual representations of motion are not identical across languages and language-specific categories shape the way we perceive and conceptualize motion events (as summarized by Ji 2014). Slobin's 'thinking for speaking' proposal (1996) is representative of the latter view and normally considered as a weaker version of the Whorfian hypothesis. It is claimed that language-specific factors influence how we conceptualize motion events when those events are linguistically expressed. 'Thinking for speaking' thus involves picking those characteristics of objects and events that (a) fit some conceptualization of the event, and (b) are readily encodable in the language.

Studies investigating the influence of specific languages on spatial cognition have produced discordant results. For instance, Papafragou et al. (2002) compared the performance of English and Greek (satellite- vs. verb-framed) children and adults in memory and categorization of motion events. Their findings indicated no language effect despite that speakers of the two languages differed greatly in verbal descriptions of these events. In contrast, Hohenstein (2005) used a preferential looking paradigm to examine Spanish- and English-speaking children's responses to visual motion event stimuli. She found that in a match-to-sample task, participants of different languages (verb-framed Spanish with 'path salience' vs. satellite-framed English with 'manner salience') behaved differently towards video stimuli in ways that could be predicted by their respective languages: 7-year-old English children fixated on videos matching the manner (rather than path) of a target video more often than Spanish speaking 7-year-olds. However, the performance of the two groups of 3.5-year-olds did not differ significantly in preferential looking. Such findings revealed a clear effect of language typology on motion conceptualization in showing that children's cognition is similar prior to becoming accustomed to using spatial language in ways typical to their native language (i.e. at the age of 3.5 years) but then demonstrates differences after such habitual use (i.e. at the age of 7 years).

4 Research Questions and Hypotheses

Within the research context reviewed in section 3, the current study aims to address linguistic relativity along two lines of research: a) in the field of L2 acquisition: whether and to what extent L2 learners can shake off the typological pattern of their L1 and fully acquire the standard way of motion expressions in an L2, particularly in term of syntactic packaging of dense motion information beyond the clause level, and b) in the field of the relationship between language and thought: whether and to what extent native speakers of partially typologically different languages (i.e. satellite-framed English and equipollently-framed Chinese) judge the similarity between motion scenes on different criteria while viewing video clips illustrating caused motion.

If a strong Whorfian effect exists, we then expect, regarding a), that language-specific differences are not superficial and learning a different language implies learning a new way of conceptualizing motion. Thus typological properties of the source language will influence the learning process and L2 learners will have difficulties in fully adapting to the target language's patterns of motion. As regards b), following the hypothesis of linguistic relativity, language-specific properties affect a speaker's 'habitual behavior', that is, what speakers do most naturally by default in common situations. In Chinese, various aspects of motion are encoded in compound verbs with equal grammatical status and formal significance. When native speakers of Chinese attend to a motion event, they habitually attend simultaneously to both manner and path dimensions. In contrast, in English manner and path are presented in separate and syntactically unequal grammatical categories, as a result of which English speakers do not habitually perceive manner and path as equally salient and occurring simultaneously, so that they tend to focus on the element encoded in the marked grammatical category of verbs only, viz., only manner of motion. If such a linguistic contrast has an influence on a

speaker's mental representation of motion, we then expect English speakers to judge the similarity between motion scenes by employing a 'manner similarity' criterion whereas Chinese speakers will adopt '

5 Methodologies

5.1 Participants

Study A

There was a total of 96 adult participants in this study who fell into three categories: the first category included three groups of Chinese learners of English at three proficiency levels, all being Chinese university students passing College English Test Band 4 (CET-4; low proficiency), CET Band 6 (intermediate level) and Test for English Majors Band 8 (TEM-8; advanced level), respectively. The second category comprised three groups of English learners of Chinese at three proficiency levels, all being overseas students learning Chinese as a foreign language in China who passed annual placement tests of their respective universities at different mark levels. The third category was made up of two groups of monolingual native speakers of Chinese and English, all being university students in Beijing (China) and California (US), respectively. There were 12 participants in each of the 8 groups (6 males and 6 females). More detailed information about participants is set out in Table 1.

Table 1: Groups of participants in Study A

Group ID	Age range	Mean age	L1	L2	Proficiency level	Score range in relevant tests	Mean score
L2CH_Low	19–61	25.4	EN	CN	Low	30–44	38.25
L2CH_Int	19–26	22	EN	CN	Intermediate	45–89	69.5
L2CH_Adv	18–50	23.7	EN	CN	Advanced	90–118	102
CNS	17–18	18	CN	N/A	Native	N/A	N/A
ENS	20–30	23	EN	N/A	Native	N/A	N/A
L2EN_Adv	23–25	23.6	CN	EN	Advanced	60–100	77.8
L2EN_Int	19–21	21.7	CN	EN	Intermediate	426–604	546.7
L2EN_Low	19–21	19.1	CN	EN	Low	426–604	563.8

Notes on abbreviations:

EN: English

CN: Chinese

L2CH_Low: English learners of Chinese (low proficiency)

L2CH_Int: English learners of Chinese (intermediate level)

L2CH_Adv: English learners of Chinese (advanced level)

CNS: monolingual native speakers of Chinese

ENS: monolingual native speakers of English

L2EN_Adv: Chinese learners of English (advanced level)

L2EN_Int: Chinese learners of English (intermediate level)

Study B

Participants in Study B included two groups of monolingual native speakers of Chinese and English (32 participants in each group: 16 males and 16 females). They were all university students in Beijing (China) and London (UK), respectively (CNS: mean age = 19.3; ENS: mean age = 26).

5.2 Materials

Study A

The motion stimuli in this study included 16 short video clips (5 sec in duration each; see (A1)–(A16) in Appendix 1

for descriptions) illustrating a child performing a specific action that causes the displacement of an object during which time the child accompanies the object throughout the course of movement (e.g. Bonny dragged a toy car across the icy lake). The design of the stimuli followed the model for caused motion developed by Hickmann et al. (2009). The 16 clips presented two types of manner of motion: either pushing/dragging or kicking/throwing, and four types of path information: verticality (up, down), boundary-crossing (into, across), deixis (towards, away from) and course parallel to ground of motion (along, around). See (a) in Appendix 2 for an illustration.

Study B

The stimuli in Study B consisted of 16 triads of caused motion events: 16 targets and their 32 alternates. Each clip appeared for 5 sec with a 0.5 sec of black screen between a target and its two alternates and a 1 sec of black screen between triads.

Within each triad, the target video clip depicted a child performing a caused motion event (as illustrated in Study A). In the same-manner alternate, the path of motion was changed (e.g. Bonny dragged a toy car AROUND the icy lake), whereas in the same-path alternate, the manner of motion was altered (e.g. Bonny PUSHED a toy car across the icy lake). See (b) and (c) respectively in Appendix 2 for illustrations. All stimuli were arranged into two randomized orders that were counterbalanced across participants within each language group. Also, the presentation position of manner-match vs. path-match video clips (left or right of the screen) was counterbalanced across stimuli in a given order.

5.3 Procedures

Study A

Following Hickmann et al. (2009), participants were invited to describe to a remote listener what happened in these caused motion stimuli. Interviews were conducted individually in a quiet place at respective universities. A pre-test item was provided at the beginning of each session to direct the participant's attention to various motion components they were expected to report during the test (e.g. cause, manner and path of motion).

Study B

Participants were invited to view video clips in front of a MacBook Pro and asked to indicate their judgment of the similarity between motion scenes by pressing given keys on the keyboard. The 16 triads of stimuli were played to participants via the stimulus presentation software 'SuperLab 4.5', which automatically generated, at the end of each testing session, a file containing participants' choices and the time they used in reacting to each stimulus (in millisecond). All stimuli were played in a synchronized series with the target videos playing first in the center of a screen with a black background, followed by two simultaneous alternate videos placed side-by-side on the same screen. While viewing the stimuli, participants received the following audio instruction: 'This is X. Which is most like X?'. They were also required to repeatedly count backwards from 100 to 1 till the end of each triad. The purpose of this number shadowing condition was to prevent participants from subconsciously verbalizing motion scenes while watching them. A pre-test item was provided at the beginning of each session to direct the attention of participants to the fact that the matching judgment was expected to make on the basis of similarity of actions.

5.4 Coding

Study A

Following Hickmann et al. (2009), each stimulus was analyzed as presenting a rich set of caused motion information, these being a) path [P], b) cause of motion [C], and three types of manner information: manner of the agent [Ma], manner of cause [Mc] and manner of the object [Mo], as illustrated in example (3), below. To suit the purpose of investigating the syntactic packaging of multiple motion components, only responses containing simultaneously path, cause, and at least one type of manner information were selected for analyses.

(3) Bonny rolled [C+ Mc] a barrel of beer around [P] the kitchen table. [Target utterance for analysis]

In accordance with Allen et al. (2000), four patterns of wrapping up dense semantic motion information at the

syntactic level were distinguished:

- (4) a. ‘Compact’: English simplex clauses and Chinese BA constructions
Bonny rolled a barrel of beer around the kitchen table.
- b. ‘Semi-compact’: English complex sentences and Chinese ZHE constructions
Bonny went around the kitchen table rolling a barrel of beer.
- c. ‘Loose’: coordinated clauses and clauses in concatenation in English and Chinese
Bonny walked around the kitchen table and rolled a barrel of beer.
- d. ‘Others’: any other syntactic patterns used by the participant to encode caused motion events (e.g. those involving dependent clauses)
Bonny is rolling a barrel of beer that went around the kitchen table.

All target responses were identified as adopting one of the above-illustrated syntactic strategies.

Study B

Data for study B were analyzed using both categorical measure (i.e. participants’ preferences) and continuous measure (i.e. reaction time). According to the specific keys participants had pressed during the testing session, their judgments were, first of all, classified as preferring either the manner-match or the path-match. The reaction time (RT) for a given stimulus was calculated from the presentation of alternate videos in a triad to their completion with the theoretically longest RT of 6,000 milliseconds (ms). The RT data were first cleaned by excluding physically impossible short values (button pressed within 5 ms of stimulus onset). As for extremely long values, a prior screening for outliers was conducted by removing all observations that were at a distance of more than two standard deviations from the mean of the distribution.

6 Results and Discussion

Study A

Our results showed that each participant group produced utterances encoding path, cause and at least one type of manner for caused motion in a predominating proportion over a group total of responses of 192 (L2CH_Low: 56%, L2CH_Int: 81%, L2CH_Adv: 88%, CNS, ENS and L2EN_Adv: 100%, L2EN_Int and L2EN_Low: 99%). Figure 1 depicts how L2 learners of English and Chinese, respectively, syntactically organized particularly dense semantic information for caused motion as compared to monolingual speakers of English and Chinese.

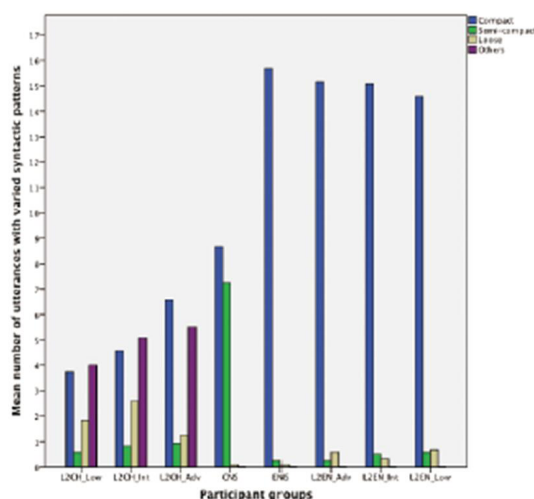


Figure 1: Mean number of utterances with varied syntactic patterns across groups of monolingual native speakers vs. L2 learners

Two observations from this figure merit further discussions. Firstly, the Chinese learners of English, across proficiencies, seemed to have acquired the characteristic way of encoding caused motion in English, that is, expressing cause + manner in the main verb while expressing path in verb particles. Such a propensity is found to be highly systematic across both learners and items. Thus, the Chinese learners of English resembled English native speakers in organizing multiple semantic components for caused motion in a highly ‘Compact’ syntactic pattern. Examples (5) and (6), using descriptions of the same item, illustrated this tendency.

- (5) Bonny pushed the treasure bag up the pyramid. (ENS_10A)F
- (6) a. He pushes the bag up the pyramid to the top. (L2EN_Adv_06A)
- b. Bonny is pushing up the treasure bag up the pyramid. (L2EN_Int_12B)
- c. Bonny is pushing a bag of dollars up to the pyramid. (L2EN_Low_03B)

Secondly, Chinese native speakers employed ‘Compact’ (BA) and ‘Semi-compact’ patterns (ZHE) in comparable proportions to encode caused motion at the syntactic level (example 7), as is revealed in a paired samples t-test ($t(11) = 0.875$, ns). This result is in line with previous findings regarding the ‘equipollent’ status of Chinese in motion event typology (e.g. Chen and Guo 2009, Ji et al. 2011b, Ji and Hohenstein 2014a, Slobin 2004). To give an example, Slobin (2004) reported that in narrating the ‘frog’ story, Chinese adults produced syntactically simplex utterances ‘flying out of the hole’ and complex sentences involving gerunds ‘exiting the hole flying’ equally frequently.

- (7) a. ‘Compact’: BA construction
Bonny ba3 cai2bao3 dai4 tuo1-jin4 le jin1zi4ta3
Bonny BA treasure bag pull-enter ASP pyramid
‘Bonny pulled the treasure bag into the pyramid’ (CNS_01A)
- b. ‘Semi-compact’: ZHE construction
Bonny tuo1 zhe cai2bao3 dai4 jin4ru4 le jin1zi4ta3
Bonny pull ZHE treasure bag enter ASP pyramid
‘Bonny, pulling the treasure bag, went into the pyramid’ (CNS_06A)

The performance of English learners of Chinese differed from that of Chinese native speakers in several aspects. Although their use of the BA construction increased in frequency with proficiency, their option for the ZHE construction was occasional with a mean frequency below 10% across proficiency levels. Most importantly, these Chinese L2 learners seemed to have adopted a syntactic pattern of information organization that can only be categorized as ‘Others’ (i.e. example 8). A closer examination revealed that such responses were characterized by: a) the use of two verbs in concatenation. The first verb was not marked for any aspect whereas the second verb was normally marked for the perfective aspect (i.e. ending with *le*), and b) a single intonation pattern of ‘rise-fall’ rather than two intonation patterns with two rising tones, which signaled that the construction under discussion was a single grammatical unit rather than two coordinated clauses exemplifying a ‘Loose’ pattern of syntactic organization.

- (8) Bonny la1 wan2ju4 che1 huan2 le
Bonny pull toy car surround/around ASP
bing1 hu2 yi4 zhou1 (L2CH_Adv_04B)
ice lake one circle
‘Bonny, pulling the toy car, went around the iced lake in one circle’

Ji and Hohenstein (2014b: 112) put forward three theoretically possible interpretations of this specific syntactic pattern:

- a) The English learners of Chinese mirrored the typical pattern for motion descriptions in their L1 onto their L2, that is, they were treating V1 in constructions as demonstrated above (e.g. ‘pull’) as the main verb while V2 as a particle (e.g. ‘around’).
- b) The L2 learners produced a hybrid syntactic pattern that integrated typological features of both L1 and L2. That is, V1 was used as a verb, reflecting the satellite-framing properties of the source language (English). Meanwhile, V2

was also regarded as a verb (e.g. ‘surround’), showing the verb–framing aspect of the target language (Chinese).

c) The English learners of Chinese chose to produce a ‘reduced’ form of ZHE construction in which the durative aspectual marker suffixing V1 (i.e. *zhe*) was omitted.

Ji and Hohenstein (2014b) argued that of the three interpretations, c) seemed most likely. The problem with a) is that if learners were indeed using V2 as a particle, they should not have suffixed it with the perfective aspectual marker *le*, in most cases, to signal the ending of an event. Further, since Chinese lacks morphological devices to mark the exact grammatical status of a lexical item, it is difficult to conclude that V1 in b) is used as a full verb, particularly when it is not followed by any aspectual markers. It is more likely that the English learners of Chinese meant to use the ZHE construction, however, due to the heavy semantic and syntactic load in producing an entirely grammatically correct ZHE sentence, they gave a ‘weakened’ form of the target construction instead. By doing so they produced a grammatically not-so-accurate utterance but achieved the communicative goal of expressing multiple semantic components for caused motion. In other words, efforts may have been made on the part of the learners to strike a balance between ‘linguistic effort’ and ‘communicative effect’. Seen in this way, the syntactic pattern of information distribution as demonstrated in example (8) bears more resemblance to the target language (Chinese) rather than showing significant influence from the source language (English).

Study B

As previously stated, we analyzed the judgment data using two types of measures: a) categorical preferences (i.e. manner– or path–match) and b) reaction time in millisecond. Following the Whorfian hypothesis, we predicted that English participants would be more manner–oriented while Chinese speakers would be equally manner– and path–oriented. In addition, given that English directs its speakers’ attention to manner only in default situations while Chinese encourages its speakers to pay an equal amount of attention to manner and path, we further hypothesized that the reaction time in making path–similarity judgments would be significantly shorter in Chinese than in English.

Paired samples *t*-tests, conducted regarding Chinese and English, respectively, revealed that within each language group, monolingual speakers chose the path–match video significantly more frequently than the manner–match video (Chinese: $t(31) = -3.370$, $p = 0.002$; English: $t(31) = -3.832$, $p = 0.001$). Further, differences in the frequency of the path–match judgment were trivial between the group of Chinese and English participants, as is confirmed in an independent samples *t*-test (Chinese: Mean = 0.62, SD = 0.20, English: Mean = 0.64, SD = 0.21; $t(62) = -0.352$, ns). These results suggested a shared tendency of path–match judgments between different language groups. Further investigations were conducted to see how systematic this tendency was across individual test items. Figure 2, below, presents the mean frequency of path preferences across 16 items. The responses to individual items were not uniform. Some caused motion events yielded more path preferences (e.g. item 3) while others eliciting more manner preferences (e.g. item 12). This indicated that apart from a common tendency of preferring the path–match video, there are certain aspects of the designed items that direct the attention of speakers to either path or manner constantly, irrespective of language group. To give an example, item 3 (pulling the treasure bag up the pyramid) was significantly more path–match inducing than item 12 (crawling up the cave). Relevant chance analyses showed that the responses to item 3 fell above chance levels (i.e. more path–matches) across groups, $t(63) = 17.015$, $p = .000$, although the responses to item 12 fell below chance levels (i.e. more manner–matches) across groups, $t(63) = -14.346$, $p = .000$. A repeated measure analysis of variance (ANOVA) with language group as the between–subjects factor (Chinese, English) and items (16) as a within–subjects factor revealed a significant effect of item only, $F(15, 930) = 20.604$, $p = .000$, suggesting that there was a pattern, irrespective of language group, in which a particular test item was viewed as more salient in path or in manner. In general, the results demonstrate that monolingual adult speakers of Chinese and English seemed to perceive and judge the similarity between caused motion events on the same criterion, viz., path–similarity. Such findings did not map onto the differences in how native speakers of Chinese and English linguistically construct these types of motion events (example [5] vs. [7a] and [7b]; see Ji et al. 2011a, 2011b for more details).

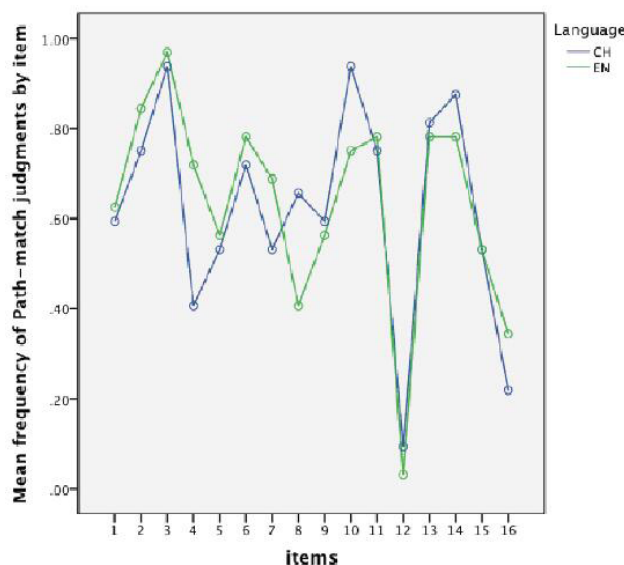


Figure 2: Mean frequency of path-match judgments by language and item

A different perspective was taken on these results by looking at the RT used in making path-match judgments between two groups of participants. An independent samples *t*-test showed that Chinese participants were significantly quicker (Mean = 2090.60, SD = 471.14) in making path-match judgments than their English counterparts (Mean = 2544.73, SD = 584.41), $t(62) = -3.422$, $p = .001$. Such findings may be interpreted in different ways. For example, it might be in line with the difference in how the English / Chinese language directs its speakers' habitual attention to manner and / or path by default in common situations. It is also likely that such observations are generally culture-specific: Chinese speakers tended to make their choices at the earliest point of time possible, whereas their English counterparts seemed to adopt a 'wait-and-see' strategy and tended to make their judgments near the end of the stimuli.

7 General Discussion and Conclusion

In the current study, we aimed to test the hypothesis of linguistic relativity by looking at whether language-specific influences can function at both the linguistic level in an L2 learning context and the cognitive level in a match-to-sample task. Results of Study A showed that the Chinese learners of English across proficiencies seemed to acquire the characteristic pattern of caused motion expressions in English fairly quickly (i.e. 'manner and cause' verb + path particle). Although the English learners of Chinese did not seem to fully acquire all target patterns for caused motion information packaging (particularly the ZHE construction) even at the advanced stage, their interlanguage exhibited substantial resemblance to the target language rather than showing significant influences from their native language, especially when examined from a particular perspective (i.e. the 'weakened' form of ZHE construction). Overall, these findings indicated that the typological pattern of L1 can be generally shaken off when acquiring the standard pattern of information distribution in the target language, thus suggesting that learning to encode caused motion in a non-native language may not necessarily involve a new way of conceptualizing events.

As regards Study B, our results revealed that adult participants, irrespective of language, tended to be more path-oriented than manner-oriented, despite that Chinese and English differ strikingly in how to encode caused motion linguistically. Although the analysis of the RT suggested that Chinese participants were significantly quicker in making path-match judgments than did English speakers, such differences can be equally attributed to 'general culture' as to potential language-specific influence.

Generally, our results provide little support for the Whorfian hypothesis of linguistic relativity. As far as L2 acquisition is concerned, Slobin, in his 'thinking for speaking' hypothesis (a weaker version of linguistic relativity),

argued that each native language ‘has trained its speakers to pay different kinds of attention to events and experiences when talking about them; this training is carried out in childhood and is exceptionally resistant to restructuring in adult second language acquisition’ (Slobin 1996: 89). Our findings, however, suggest that in a particular context where the source and the target language share partial typological similarities, the reconstruction of one’s thinking pattern may not be as daunting as was assumed. For instance, the ‘Compact’ pattern of information packaging prevails in English whereas the ‘Compact’ and the ‘Semi-compact’ pattern occur equally frequently in Chinese. When Chinese learners of English acquire the target pattern, what they need to do may merely involve activating a subset of their ‘thinking for speaking’ pattern and then mirror it onto the L2 system. For English learners of Chinese, despite that one of the target patterns (ZHE) is not characteristically recruited for caused motion expressions in their L1, the online reorganization of particularly dense motion information in an L2 did not seem to pose unconquerable acquisitional difficulties to them, as is attested from the considerable similarity between their interlanguage and the L2.

Our results in the nonlinguistic similarity judgment task are not consistent with some findings in previous studies as well. For example, Hohenstein (2005) found that in a match-to-sample task, English- and Spanish-speaking adults judged the similarity between motion scenes differently, and in a way that could be predicted by the specific typological contrast between the two languages involved in motion event descriptions. Such a clear effect of language typology on conceptualization is absent from the current investigation. These two studies differ in many aspects, two of which merit particular attention. First of all, the present study employed a ‘shadowing’ condition in which participants were prevented from subconsciously verbalizing motion events while viewing the video clips (i.e. a truly ‘nonlinguistic’ task), whereas in Hohenstein’s (2005) study, no ‘shadowing’ was provided and the potential effect of language typology on the similarity judgment cannot be entirely excluded. Further, Hohenstein (2005) used the preferential looking paradigm and examined the amount of time a participant fixated to the path-match minus his / her fixation to the manner-match. This continuous measure, as compared to our categorical judgment (i.e. preferences), is more likely to reveal the degree of differences between language groups.

To conclude, the results from both linguistic and non-linguistic experiments in the current study did not provide strong evidence for the Whorfan hypothesis. It seems that language-specific properties constrain the L1 linguistic representation of motion only. When complex caused motion events are represented in an L2 learning context or beyond the linguistic level at the cognitive realm, the language-specific influence is either insignificant or almost invisible. This suggests that the striking differences in L1 motion descriptions across languages tend to be superficial, and linguistic and non-linguistic representations of caused motion may be dissociable from each other. Overall, our findings provide fresh insights into the relation between language and thought in general. Future research along the same lines can investigate the effect of language typology in a more diverse L2 context (e.g. L2 learners from satellite-, verb- and equipollently-framed languages), and employ more subtle means such as the continuous variable of fixation time to analyze the non-linguistic (e.g. preferential looking) data.

Acknowledgements

I would like to thank the British Academy Postdoctoral Fellowship, which made an award to me in 2010, for its generous financial support of the current study (pf100022). I am also indebted to the Shenzhen University Social Sciences Foundation for supporting me, in various ways, in finalizing this study (project code: 801).

My special thanks go to Dr. Jill Hohenstein at King’s College London, who read different versions of this manuscript and provided many insightful observations.

References

- Allen, S., A. Ozyurek & S. Kita. (2007). Language-specific and universal influences in children’s syntactic packaging of Manner and Path: A comparison of English, Japanese, and Turkish. *Cognition*, 102 (1), 16–48.
- Bowerman, M. (1999). Learning how to structure space for language: A crosslinguistic perspective. In P. M. Bloom & L. Peterson (Eds.), *Language and Space* (pp. 385–436). Cambridge, MA: MIT Press.

- Bowerman, M. & S. Choi. (2001). Shaping meanings for language: Universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman & S. Levinson (Eds.), *Language Acquisition and Conceptual Development* (pp. 475–511). Cambridge: Cambridge University Press.
- Chen, L. & J. Guo. (2009). Motion events in Chinese novels: Evidence for an equipollently-framed language. *Journal of Pragmatics*, 41, 1749–1766.
- Du, H. (2006). *The Acquisition of the Chinese ba-construction*. Muenchen: LIMCOM EUROPA.
- Gennari, S., S. Sloman, B. Malt & W. Fitch. (2002). Motion events in language and cognition. *Cognition*, 83, 49–79.
- Hickmann, M., H. Hendriks & C. Champaud. (2009). Typological constraints on motion in French and English child language. In J. Guo, E. Lieven & N. Budwig (Eds.), *Crosslinguistic Approaches to the Psychology of Language: Research in the Tradition of Dan Isaac Slobin* (pp. 209–224). New York: Lawrence Erlbaum Associates.
- Hohenstein, J. (2005). Language-related motion event similarities in English- and Spanish-speaking children. *Journal of Cognition and Development*, 6, 403–425.
- Ji, Y. (2014). *The Expression of Motion Events: Typological and Developmental Perspectives*. Beijing: China Social Sciences Press.
- Ji, Y., H. Hendriks & M. Hickmann. (2011a). How children express caused motion events in Chinese and English: Universal and language-specific influences. *Lingua*, 12, 1796–1819.
- Ji, Y., H. Hendriks & M. Hickmann. (2011b). The expression of caused motion events in Chinese and in English: Some typological issues. *Linguistics*, 49 (5), 1041–1076.
- Ji, Y. & J. Hohenstein. (2014a). The expression of caused motion by adult Chinese learners of English. *Language and Cognition*, 6 (4), 427–461.
- Ji, Y., & Hohenstein, J. (2014b). The syntactic packaging of caused motion components in a second language: English learners of Chinese. *Lingua*, 140, 100–116.
- Papafragou, A., C. Massey & L. Gleitman (2002). Shake, rattle, ‘n’ roll: The representation of motion in language and cognition. *Cognition*, 84, 189–219.
- Slobin, D. I. (1996). From ‘thought to language’ to ‘thinking to speaking’. In J. Gumperz & S. Levinson (Eds.), *Rethinking Linguistic Relativity* (pp. 70–96). Cambridge: Cambridge University Press.
- Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömquist & L. Verhoeven (Eds.), *Relating Events in Narrative Vol. 2: Typological and Contextual Perspectives* (pp. 219–257). Mahwah, NJ: Lawrence Erlbaum Associates.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical form. In T. Shopen (Ed.), *Language Typology and Syntactic Description Volume 3* (pp. 36–149). Cambridge: Cambridge University Press.
- Talmy, L. (2000). *Toward a Cognitive Semantics, Vol. 2: Typology and Process in Concept Structuring*. Cambridge, MA: MIT Press.
- Talmy, L. (2009). Main verb properties and equipollent framing. In J. Guo, E. Lieven & N. Budwig (Eds.), *Crosslinguistic Approaches to the Psychology of Language: Research in the Tradition of Dan Isaac Slobin* (pp. 389–402). New York: Psychology Press.
- Zeng, Y. H. (2011). *Acquisition of English Motion Event Expressions by Chinese EFL Learners* (Doctoral dissertation). Hunan: Hunan Normal University.