# Cognitive restructuring in the multilingual mind: motion event construal in Cantonese-English-Japanese multilingual speakers

Yi Wang

Thesis submitted to University College London
for the degree of Doctor of Philosophy

March, 2020

# Declaration

I, Yi Wang confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

## Abstract

Languages differ typologically in motion event encoding (Talmy, 2000). Furthermore, the cross-linguistic variations in expressions tend to modulate cognition in a dynamic and task-dependent manner (Slobin, 1996a; Wolff & Holmes, 2011). Although evidence shows that language-specific structures affect spatial cognition with L1 and L2 speakers, little is known about how multilinguals process spatial information during the course of L3 acquisition. The present study addresses the question whether, and to what extent, the acquisition of an L2-English (satellite-framed) in childhood and an L3-Japanese (verb-framed) in adulthood restructure L1-Cantonese-based (equipollent-framed) lexicalization and categorization patterns in processing spatial concepts when the target language is actively involved in the decision-making process.

Participants (N=150) were Cantonese-English bilinguals, Cantonese-English-Japanese multilinguals, and monolingual controls (N=30 each). A cartoon-based test was specially designed for the study, with a verbal encoding and a triads-matching task. In verbalization, monolinguals were instructed and narrated 'what happened' in each stimulus in their native languages. Bi- and multilinguals were assigned to a monolingual and a bilingual context by manipulating immediate language use in their oral descriptions. Then participants were given a subsequent similarity judgment task where they needed to decide as soon as possible which alternate (manner- or path-oriented) was more similar to the target event. Their reaction time of decision-making was recorded.

Results from monolinguals confirmed language-specific effect of semantic-conceptual structures on event conceptualization. Results from bi- and multilinguals demonstrated an ongoing restructuring from L1-based patterns towards the L2- or L3-based patterns in both event lexicalization and categorization regardless of the language context. And the degree of restructuring is modulated by the amount of language contact with the L2 and L3. The findings suggest that learning a language means internalizing a new way of thinking and provides positive evidence for cognitive restructuring of L1-based patterns within the frameworks of thinking-for-speaking and associative learning.

## Impact Statement

The research reported in this thesis is of basic research nature and aims to extend the interplay between language and cognition to the domain of multilingualism. It examines the potential effects of language learning on perceptual and attentional processing of motion events with multilinguals across different modalities (i.e. verbal encoding, co-verbal processing efficiency and non-verbal categorization).

It is a multidisciplinary research that incorporates different approaches from linguistics, experimental psychology, as well as second language and additional language acquisition within the broad discipline of language education. Situated within a broader frame of language and thought research, this study aims to answer the fundamental question of whether learning a new language changes the way we think by combining innovative psycholinguistic experiments and techniques (reaction time) with basic cognitive processes (perception and categorization) across different modalities in the context of dynamic multilingualism. This project has theoretical, methodological and pedagogical implications for the language-and-thought debate and additional language learning by linking the conceptual development in bi- and multilingual speakers to second/additional language acquisition. It can benefit both academic researchers and larger social communities in the following ways.

First, this study adopts a multimodal approach into language-and-thought research, which benefits cognitive scientists, psychologists, linguists, and L2 researchers by engaging linguistic relativity research with the context of additional language learning. It also takes various individual factors into consideration when addressing the cognitive mechanisms underlying the language effect on cognition.

Second, this study links the conceptual development of bi- and multilingual speakers to second/additional language teaching and learning and sheds light on what aspect of thinking for speaking may pose challenges for learning during different stages of additional language acquisition. It can benefit a range of different audiences such as language learners, parents, teaching professionals, educators, curriculum designers, and policy makers in terms of how to use multimodal pedagogical tools and constructive teaching materials to maximize the learning potentials and the cognitive benefits of language learning. With a multi-dimensional examination of different aspects of multilingual behaviour in the domain of spatial cognition, this study will have pedagogical implications on various education practices, such as teacher training, teaching materials, language teaching syllabuses and language curriculum design. This may also benefit and empower parents to make their decisions in terms of bringing up their children multi-lingually based on the latest research findings.

Last but not least, this study also has practical implications for everyday human life as it bridges gaps between cross-linguistic and cross-cultural differences in thinking and acting with more than two languages. This facilitates successful inter-cultural communication and social interaction for speakers of different linguistic and cultural backgrounds in real-life behaviour.

# Acknowledgements

This thesis would not have been possible without the guidance, support, and encouragement of so many people over the past three years. I would like to take this opportunity to express my sincere thanks and gratitude to my supervisors, colleagues, participants, friends and family.

First and foremost, I would like to express my deepest and greatest gratitude to my principle supervisor, Professor Li Wei, for his unwavering support and careful guidance at every stage of my PhD journey. His vision, expertise and profound knowledge in the field of language learning, bilingualism and cognitive science have inspired me to read widely and take a multi-disciplinary approach in my doctoral research. His encouragement and inspiration greatly motivate me to follow his steps in becoming a scholar devoted to academic research, teaching and mentoring students.

I am also very grateful for my subsidiary supervisor, Professor Andrea Révész, for her tremendous help and insightful comments on how to conduct statistical analyses for my research with the most up-to-date statistical methods.

I also wish to extend my sincere appreciations to Professor Stephen Matthews, my MA supervisor, for sparking my interests in cross-linguistic research and encouraging me to embark on my PhD journey. I also would like to show my sincere thanks to Professor Jeanine Treffers-Daller, Professor Marianne Gullberg, and Professor Scott Jarvis for kindly sharing their insightful comments and suggestion on my research on several occasions.

My sincere appreciations go to all members of the London Second Language Acquisition Research Forum (L-SLARF) for holding and organizing annual conferences and seminars for doctoral students. I have received lots of constructive and insightful comments and feedback on my current research.

My special thanks go to all my participants across different parts of the world, for taking an active part in my study. I also would like to express my sincere thanks to the editorial board of Language Learning, for their financial support. The Language Learning Dissertation Grant they offered helps me carry out the current research smoothly.

My heartfelt thanks also go to all my friends and colleagues both in London and in China. Their generous support, help and companionship along the way make my PhD a pleasant and memorable experience.

Last but not least, I wish to show my greatest thanks for my beloved family. I am in great debt to my parents for their unconditional love and support all these years. It is their love, care and belief in me that make who I am today.

## List of Publications and Presentations based on the thesis

- 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. (in revision). 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.*

- Wang, Y., & Li, W. (under review). Cognitive restructuring in the multilingual mind: the effects of multiple language learning on motion event construal in Cantonese-English-Japanese speakers. *Language Learning.*

- Wang, Y. (2020). Language-specific effects on processing efficiency of caused motion events in Cantonese-English bilinguals learning an L3. Paper presented at AAAL, 28-31, March 2020, Denver, USA.

- Wang, Y. (2019). Cognitive restructuring in the multilingual mind: the role of long-term L2 exposure and short-term L3 exposure in expressing and conceptualizing caused motion. Paper presented at EuroSLA 29, 28-31, August 2019, Lund, Sweden.

- Wang, Y. (2019). Cognitive restructuring in the multilingual mind: motion event construal in Cantonese-English-Japanese multilinguals. Paper presented at AAAL, 9-12, March 2019, Atlanta, USA.

- Wang, Y. (2018). Motion event acquisition: does learning additional languages change the way we perceive the world? Paper presented at Second Language Research Forum (SLARF), 26-28 October 2018, Montreal, Canada.

## TABLE OF CONTENTS

## A List of Tables

## A List of Figures

## Chapter 1. Introduction

For centuries, thinkers have wondered about whether the language we speak influences the way we think. This age-old question has recently received a lot of renewed interest as a number of new research paradigms have evolved that allow addressing the connection between language and thought empirically. Experimental evidence shows that language can exert temporary and immediate, or habitual and durable effects, on various cognitive processes, such as perception, categorization, reasoning and recognition memory, in a flexible and context-dependent manner (Slobin, 1996b; Whorf, 1956; Wolff & Holmes, 2011). For example, cross-linguistic differences in colour vocabularies tend to cause differences in colour perception and categorization, indicating that language effects are deep and profound in the sense of affecting even basic visual perception (Athanasopoulos, 2009; Lupyan, 2012; Winawer et al., 2007).

On the one hand, language effects have been detected in various cognitive domains such as time (Boroditsky, Fuhrman, & McCormick, 2011; Casasanto & Boroditsky, 2008; Casasanto, 2010), colour (Athanasopoulos, 2009; Athanasopoulos, Damjanovic, Krajciova, & Sasaki, 2011; Winawer et al., 2007), objects (Cook, Bassetti, Kasai, Sasaki, & Takahashi, 2006; Pavlenko & Malt, 2011) and motion (Ji & Hohenstein, 2018; Park, 2019). On the other hand, these effects are context-bound and visible under certain conditions. For example, the linguistic relativity effects are consistently observed when language is actively used during online thinking (Filipovic, 2018; Montero-Melis, Jaeger, & Bylund, 2016; Trueswell & Papafragou, 2010), or when language is used as a strategy to solve a subsequent cognitive task (Lai, Rodriguez, & Narasimhan, 2014; Lupyan, 2012). However, such effects disappear when the access to language is blocked by task manipulation (Gennari, Sloman, Malt, & Fitch, 2002; Montero-Melis & Bylund, 2017).

These mixed findings for the language-specific effects on cognition have motivated researchers to examine in more depth when and under what conditions such effects are most likely to occur (see Bylund & Athanasopoulos, 2014b, for a detailed review). For instance, language effects on cognition are most likely to appear when language-specific labels are explicitly involved at the very moment of decision-making. This process, termed thinking-for-speaking (Slobin, 1996a), emphasizes the language effects on online thinking when speakers are actively engaged in language-driven activities, such as language comprehension or production. In addition, there are cases where language is used as a strategy to solve a subsequent cognitive task, especially when the task lacks a correct or objective answer and requires a higher level of cognitive processing (Finkbeiner, Nicol, Greth, & Nakamura, 2002). Such a 'thinking with language' effect, which can be manipulated by various experimental set-ups, concerns the spontaneous recruitment of linguistic resources as a meddler or augmenter to aid working memory and facilitate answer formulation (Lupyan, 2012; Wolff & Holmes, 2011). In addition, in a more recent view, the label-feedback hypothesis suggests that the effects of language on cognition are dynamic and occur in an ad hoc fashion (Lupyan, 2012). In this view, recent linguistic experience can activate related non-linguistic representations which speakers can draw upon for a subsequent cognitive task. And the mutual feedback between linguistic and non-linguistic representations can be manipulated by short-term language mediation, such as linguistic priming, language training and verbal encoding, immediately prior to or during mental processing (Montero-Melis & Bylund, 2017). In fact, a growing body of research has demonstrated that the effects of language on cognition draw on many factors, such as the nature of the experimental stimuli (simplex or complex), the involvement of language (explicit, implicit or with verbal interference), and experimental manipulation (linguistic priming, language in operation). Thus, instead of asking whether or not language determines thought, current studies have shifted their focus to considering which language-specific categories affect which cognitive domains under what conditions (Bylund & Athanasopoulos, 2014b).

Research on the effects of language learning on cognition starts with monolinguals of typologically contrastive languages, although recent studies have started to extend the scope by including bilingual speakers and L2 learners of various types (Athanasopoulos, Damjanovic, Burnand, & Bylund, 2015b; Bassetti & Cook, 2011; Brown & Gullberg, 2013; Flecken, Carroll, Weimar, & Von Stutterheim, 2015). Bilingualism research on language and cognition mainly focuses on the dynamic relationship between the process of additional language learning and the changing of the entire cognitive state of bilingual speakers. Empirical evidence shows that a successful acquisition of an L2 needs not only the internalization of novel linguistic frames, but also associated conceptual distinctions. This may give rise to the restructuring of original conceptual categories acquired through the L1 (Jarvis & Pavlenko, 2008). This process, termed conceptual or cognitive restructuring, refers to the conceptual changes that bilinguals undergo during the acquisition of a new language. It is a gradual process and occurs in bilinguals' verbal and non-verbal behaviours (Jarvis, 2007; Pavlenko, 2011). The conceptual changes bilinguals have are gradient and exhibit various forms, such as co-existence of the L1- and L2-based concepts (Hohenstein, Eisenberg, & Naigles, 2006; Sachs & Coley, 2006) , convergence (Brown & Gullberg, 2013; Cook et al., 2006), shift to L2-based concepts (Athanasopoulos, Damjanovic, Burnand, & Bylund, 2015a; Park & Ziegler, 2014) and attrition of L1-based concepts (Bylund, 2009b; Bylund & Jarvis, 2011). The connections between cognitive restructuring and linguistic relativity research are based upon the observation that conceptual representations within the bilingual mind are multimodal, dynamic and highly inter-correlated across modalities (Bylund & Athanasopoulos, 2014b; Casasanto, 2008). Studies to date have demonstrated that the degree of cognitive restructuring may be modulated by various long-term learning effects, such as L2 proficiency (Ji, 2017; Park, 2019) , age of L2 acquisition (Boroditsky, 2001; Lai et al., 2014), and the amount of language exposure, as well as the frequency of language use (Bylund & Athanasopoulos, 2014a, 2015, Park, 2019).

More recently, another line of research suggests that the conceptualization patterns that bilinguals have are context-bound and susceptible to immediate experimental manipulation, such as linguistic priming (Lai et al., 2014; Montero-Melis et al., 2016), biased instructions (Brown & Gullberg, 2008; Kersten et al., 2010) and language context (Athanasopoulos, Bylund, et al., 2015; Stocker & Berthele, 2019). For example, recent linguistic exposure can trigger bilingual switches between language-specific representations in perceptual judgement and recognition, depending on the language they are using (Athanasopoulos, Bylund, et al., 2015). That is, bilinguals in an L1-instructed context display an L1-based conceptualization pattern, but an L2-based conceptualization pattern when the instructions are given in the L2.

Despite the ongoing interest into the language-and-thought research, several issues remain unresolved. First, studies regarding the language-and-thought debate mainly focus on bilingual speakers or different types of L2 learners through cross-linguistic comparisons, while research on speakers of more than two languages remains limited. With the exceptions of Bylund, Athanasopoulos, and Oostendorp (2013) and Bylund and Athanasopoulos (2014a), from a grammatical perspective, very little has been done to examine how speakers of more than two languages (multilingual speakers) conceptualize motion events from a lexical perspective of manner versus path in the broad context of multilingualism and additional language learning. As multilingualism is a common linguistic phenomenon worldwide (Aronin & Singleton, 2012) and multilingual speakers display unique linguistic and cognitive features (Cook & Li, 2016), extending research into language and thought to the domain of multilingualism will allow consideration of how different languages within a multilingual mind affect thought together.

Second, most studies along these lines are conducted with late bilingual speakers or adult L2 learners with typologically contrastive languages (Athanasopoulos, Bylund, et al., 2015; Montero-Melis et al., 2016; Stocker & Berthele, 2019). However, little is known about how bi- and multilinguals with partial overlap linguistic systems tend to behave: whether they have a single integrated way of thinking-for-speaking, or if they can switch between distinct sets of thinking patterns depending on which language they are using.

Third, the overall picture of the effects of language on cognition suggests that bilinguals' conceptual representations are dynamic and multimodal in the sense that they can be influenced by both long-term factors and short-term linguistic mediation. However, it remains unclear how long- and short-term variations interact with each other in modulating bi- and multilinguals' cognitive patterns together.

In order to address these research gaps, the present thesis aims to extend language-and-thought research to the context of multilingualism and additional language learning and to take a first step in exploring how Cantonese-English-Japanese multilinguals with three typologically different languages encode and gauge event similarity in the motion domain under different conditions. More specifically, the current study aims to address the questions of whether, and to what extent, the acquisition of an L2-English (satellite-framed) in early childhood and an L3-Japanese (verb-framed) in adulthood restructures the lexicalization and conceptualization patterns acquired through L1-Cantonese (equipollently-framed). In addition, it also explores how long-term effects of language learning and short-term effects of language manipulation interact with each other in modulating the process of cognitive restructuring within the bi- and multilingual mind. The study combines two research paradigms: a linguistic encoding paradigm that allows us to manipulate participants' short-term language activation in language production and a triads-matching paradigm which is higher-level cognitive and has been widely used To assess participants' perceptions of motion event similarity and their processing efficiency in the decision-making process (Athanasopoulos, Damjanovic, et al., 2015b; Ji & Hohenstein, 2018; Park, 2019; Vanek, 2019).

The thesis is organized as follows. There are altogether 10 chapters. Chapter 1 serves as the introduction to the whole thesis. Chapter 2 presents the principle of linguistic relativity as the theoretical backup for the current study, together with basic assumptions on the dynamic interplay between language and cognition. Chapter 3 pushes the boundaries of language-and-thought research to the context of bilingualism with an overview of the framework of cognitive restructuring that underlies the dynamics of language learning regarding changing the cognitive status in the bilingual mind. Chapters 4 and 5 narrow the scope of cognitive semantics to the domain of motion, with detailed reviews and descriptions of motion event typology across different target languages and empirical evidence for the process of cognitive restructuring in bilingual speakers and L2 language learners. At the end of this chapter, research gaps are identified and the specific research questions are proposed. A detailed methodology is presented in Chapter 6, which includes the overall design of the study, the instruments used for the data collection and experimental procedures. At the end of this chapter, a detailed coding outline is provided for both linguistic and non-linguistic data, together with why the mixed-effects methods are used for the data analysis. Chapters 7-9 conduct a comprehensive analysis of the experimental results of the present study. Chapter 7 presents the results for motion event lexicalization and categorization with monolingual speakers, while Chapters 8 and 9 place their focuses on bilingual and multilingual speakers. Finally, in Chapter 10, the main research findings are summarized, along with a detailed discussion and a thoughtful conclusion. This chapter ends with a careful consideration of the theoretical and methodological implications, and possible limitations of the current thesis, as well as recommendations and directions for future language and cognition research.

# Chapter 2. The effect of language on cognition: an overview of the Linguistic Relativity Hypothesis

Does the language we speak change how we think? This age-old question has generated extensive debate among linguists, philosophers, anthropologists and psycholinguists (see Bylund & Athanasopoulos, 2014b, for a recent review). The core issue, known as Linguistic Relativity (Whorf, 1956), shows that cross-linguistic differences in semantic encoding affect one's habitual thinking, even when language is not actively involved in cognitive processing. However, other studies show that the language effects on cognition are context-bound and visible under certain conditions. For example, the effects are most likely to occur when language is used as a strategy to solve a task (Gennari et al., 2002; Lai et al., 2014), but interrupted when a verbal interference is introduced (Montero-Melis & Bylund, 2017; Trueswell & Papafragou, 2010) .

Given the controversies in the language-and-thought debate, this section will first approach the basic tenet of the linguistic relativity from its historical context and then move on to its development with contemporary formulations and applications. Then, it will focus on the cognitive mechanisms underlying the interplay between language and cognition by introducing the up-to-date theoretical and methodological underpinnings. Finally, it will conduct an in-depth discussion of how different contemporary approaches to the language-and-thought debate integrate with each other and provide new insights into how languages may affect thought under different conditions.

## 2.1. Historical background of the Linguistic Relativity hypothesis

The Linguistic Relativity (LR) Hypothesis (also termed 'Whorfian effects' or 'linguistic relativity effects'), which suggests that language-specific categories shape one's thinking patterns, has a quite long history. Originating from the provocative writings and observations of French and German scholars in the 18th century, it was then advanced by many prominent philosophers and linguists, such as Wilhelm von Humboldt (1767-1835), Franz Boas (1858-1942), Edward Spair (1884-1939) and Benjamin Lee Whorf (1897-1941) (Jarvis, 2016; Pavlenko, 2014). The philosophical underpinnings of LR argue that each language represents the spirits and characters of a nation and speakers of different languages hold different viewpoints (von Humboldt, 1963). However, this review sparked an intensive debates on the notion of 'spirit' and 'character' and was soon abandoned due to a lack of proper means to adequately test this assumption (Casasanto, 2008; Lucy, 1992a; Lucy, 2016).

A century later, the development of psychology and cognitive science gave rise to a reformulation of the linguistic relativity hypothesis. It was revisited and gained more prominence via the work of two American scholars. According to Edward Sapir and Benjamin Whorf, the basic idea of LR is that structural differences across languages result in speakers' cognitive diversity, as illustrated in the following quotation:

'We dissect nature along lines laid down by our native languages. The categories and types we isolate from the world of phenomena we do not find there because they stare every observer in the face; on the contrary, the world is presented in a kaleidoscopic flux of impressions which has to be organized by our minds—and this means largely by the linguistic systems in our minds. We cut nature up, organize it into concepts, and ascribe significances as we do, largely because we are parties to an agreement to organize it in this way—an agreement that holds throughout our speech community and is codified in the patterns of our language. […] We are thus introduced to a new principle of relativity, which holds that all observers are not led by the same physical evidence to the same picture of the universe, unless their linguistic backgrounds are similar, or can in some way be calibrated.' (Whorf, 1956, pp. 213–214).

As mentioned by Whorf (1956), the LR hypothesis puts emphasis on the effects of 'language' on 'habitual thought' with a more precise definition of each notion, that is, 'language' here refers to language-specific properties in grammatical structures and 'habitual thought' is operationalized as 'conceptual categories' that modulate speakers' habitual or routinized ways of conceptualizing and perceiving the reality. The basic tenet of the LR hypothesis is that if the language we speak constraints us to attend to the external world in certain ways, speakers of different languages will develop distinct thinking patterns based on their language-specific categories.

The LR hypothesis is often viewed and interpreted as two versions: a strong version of linguistic determinism (i.e. whether language determines cognition), and a weak version of the linguistic relativity hypothesis (i.e. whether language influences cognition). The former version of linguistic determinism highlights the role that language plays in cognition and claims that language structures determine basic patterns of thought. In this view, the effect of language is so powerful that it can even override speakers' basic conceptual and perceptual abilities. Language determinism is difficult to hold ground as no evidence has ever been found to support this claim. In contrast, empirical evidence from cognitive science suggests that the interaction between language and thought allows speakers to construe the reality in accordance with language-specific units, indicating that language can influence certain aspects of thought via conventionalized form-meaning pairings (Evans, Bergen, & Zinken, 2007; Evans & Green, 2006; Lakoff, 1990; Wolff & Holmes, 2011).

Over the years, the LR hypothesis has sparked lots of debates and controversies. Criticism of the LR partly stems from the lack of empirical evidence to support this hypothesis and is partly due to the misinterpretation of Whorf's original statement (see, Pavlenko, 2016, for an overview). The most fundamental attack is from the Universal Grammar (UG) (Chomsky, 1988), which claims that the general cognition of mankind is universal and not subject to language-specific properties. However, several ground-breaking studies in experimental psychology and the cognitive sciences have successfully challenged the universal dominance of human cognition and view language as an essential and indispensable part of human cognition (Boroditsky, 2001; Casasanto, 2008; Lucy, 1992a; Regier & Kay, 2009; Roberson, 2005). Another criticism concerns with the oversimplified question of the interaction between language and cognition, especially with regard to the strong version of language determinism. However, a more recent review of Sapir and Whorf's work (Everett, 2013; Lucy, 1992b; Pavlenko, 2016; Wolff & Holmes, 2011) suggests that interpreting the linguistic relativity hypothesis as a simple 'yes-or-no dichotomy' is misleading and inadequate, because neither the strong-weak distinction nor the doctrine of linguistic determinism was ever put forward by Whorf (Pavlenko, 2011; Pederson, 2007). In fact, the dichotomy is a later invention introduced by those who attempted to reformulate Sapir's and Whorf's descriptive ideas but lost their original arguments in translation (Pavlenko, 2016). Thus, language-and-thought research should not end with an 'all-or-none' approach due to the misunderstanding of Whorf's original ideas (Lucy, 1992a; Lucy, 2016).

With the development of multi-disciplinary research in the 20th century, LR received renewed attention after the dominance of the universal-based approach. The revival of LR can be attributed to: 1) the integration of experimental psychology with cognitive semantics (Talmy, 2000) and cognitive grammar (Langacker, 1987, 1991, 2008); 2) the establishment of the critical role that language plays in one's cognitive development (Lucy, 1992a) and 3) the diversity in the cognitive domains under investigation (see Wolff & Holmes, 2011 for an overview). A new wave of interest has arisen to re-examine the interplay between languages and thought with a more diverse and rigorous methodology. Researchers have begun to place the experimental approach at the centre of research and to emphasize the necessity of taking not only verbal but also on-verbal evidence into consideration when addressing the language-and-thought debate. This new surge of interest has moved the study of language-and-thought research forward with more testable hypotheses and theoretical approaches.

### 2.2. Contemporary approaches to language-and-thought research

Contemporary approaches to the LR hypothesis (also known as neo-Whorfian approaches) are located within the framework of Cognitive Linguistics (Evan, Bergen, & Zinken, 2007; Lakoff, 1990) and characterized by three prominent features. First, the contemporary LR hypothesis holds the view that language-specific properties, cognitive operations and external experience are highly interactive and the experience of language learning is governed by general cognitive principles (Evans, Bergen, & ZinKen, 2007; Evans & Green, 2006; Lakoff, 1990). That is, the basic mental operations of human cognition are manipulated by experience of the external world. Second, languages across the world differ syntactically and semantically and cross-linguistic differences in grammatical constructions of form-meaning mappings influence how speakers perceive and interpret the external world (Langacker, 2008; Talmy, 2000). For example, regarding the grammatical aspect, speakers of languages with obligatory aspect marking (i.e. the progressive) tend to mention endpoints less often (e.g. a car is driving along the road) and focus more on the ongoing phrase, whereas speakers of languages that lack aspect markers of ‘ongoingness’ tend to show a linguistic bias towards action goals or endpoints (e.g. a car is approaching the village) and adopt a holistic perspective in perceiving the same event. Third, assuming that language is an essential part of human cognition, a change in the grammatical or syntactic structures or forms entails a change in meaning which, in turn, has immediate consequences for different cognitive processes, such as perception, reasoning, and categorization (Langacker, 1987, 1991, 2008).

Grounded in Cognitive Linguistics, contemporary approaches to linguistic relativity research are shifting from observation to experimentation and involve a wide range of psychological experiments with both verbal and non-verbal evidence. By directly utilizing a wide range of psychological (Guillaume, Panos, Alison, Benjamin, & Jan-Rouke, 2009; Levinson, 2001; Papafragou, Hulbert, & Trueswell, 2008), behavioural (Lucy & Gaskins, 2001; Regier & Kay, 2009; Roberson & Davidoff, 2000) and neuroscientific techniques (Athanasopoulos, Dering, Wiggett, Kuipers, & Thierry, 2010; Flecken, Athanasopoulos, Kuipers, & Thierry, 2015), the contemporary approach explores the effect of language on thought in a wide array of conceptual domains, such as colour (Athanasopoulos, 2009; Winawer et al., 2007), object and substances (Lucy & Gaskins, 2003; Pavlenko & Malt, 2011), time (Boroditsky et al., 2011; Casasanto, 2008), space (Levinson, 2001, 2003) and motion (Athanasopoulos, Bylund, et al., 2015; Flecken, Carroll, et al., 2015).

This emerging synthesis start to acknowledge that linguistic and non-linguistic representations are fundamentally distinct and focus on how thought or cognition is operationalized (Lucy, 1992b, 2014). For example, thought can take different forms, such as perception, categorization and recognition memory. Thus, instead of asking whether or not language affects thought, research along these lines starts to narrow down the basic enquiry to which language-specific categories tend to affect which non-linguistic representations under which conditions (Bylund & Athanasopoulos, 2014b; Wolff & Holmes, 2011).

Contemporary approaches to the linguistic relativity hypothesis are characterized by three prominent features: 1) no more linguistic determinism; 2) more stringent methodological requirements and 3) expansion of empirical evidence with a wide array of cognitive domains.

### 2.2.1. No more linguistic determinism

Contemporary approaches distance themselves from linguistic determinism and do not regard language and thought as the same thing. Evidence from cognitive science suggests that the relationship between thought and the world is much closer than that between thought and language (Levinson, 2001; Wolff & Holmes, 2011; Wolff & Malt, 2010), especially for nominal concepts such as colour, time, gender, objects and motion. The relatively loose interaction between thought and language stems from the fact that cross-linguistic differences in linguistic labelling are much more diverse than differences observed in thinking patterns. Some evidence shows that cross-linguistic differences in linguistic encoding are not always manifested in non-linguistic behaviour (Malt, Sloman, Gennari, Shi, & Wang, 1999; Munnich, Landau, & Dosher, 2001). For example, Munnich et al. (2001) examined English, Japanese and Korean speakers' memory of spatial locations. Results suggested that although the three languages differed cross-linguistically in naming patterns for spatial expressions, speakers' memory of spatial locations exhibited a more convergent pattern. Given the mixed evidence, this suggests that a shifting away from a simple all-or-nothing question can benefit current research by acknowledging that language and thought are not exactly the same. But this does not imply that language cannot have an impact on thought. Instead, given the differences between language and thought, we can expect some interaction between them and the interaction within each system can exert some impact on the other.

### 2.2.2. More stringent methodological requirements

Contemporary approaches to language-and-thought research have shifted their focus from anthropology to psychology with linguistic analysis and cognitive assessment. Simultaneously, this emerging synthesis starts to examine the interface between language and thought with novel perspectives and innovative methodological advances (Lucy, 1992b, 2016; Pavlenko, 2011).

Contemporary approaches to linguistic relativity research are characterized by two prominent features: 1) transformation from description to experimentation and 2) involvement of a wide range of non-verbal tasks for the measurement of non-verbal behaviour in addition to purely linguistic data only (Lucy, 2016; Pavlenko, 2016). The distinctions between verbal and non-verbal behaviours serve as a prerequisite for modern approaches to the linguistic relativity research, as the involvement of non-verbal measurement can largely avoid the methodological flaw of language-thought circularity (Athanasopoulos & Bylund, 2013b; Lupyan, 2012). Theoretically, verbal behaviours refer to behavioural data concerning language comprehension or production, with narrative tasks being operationalized as one of the most commonly used elicitation techniques (i.e. picture description or retelling of the stories). While non-verbal behaviour, or thought, is operationalized as a wide range of mental processes, such as attention, reasoning, recognition memory, similarity judgements and categorical perception (Bylund & Athanasopoulos, 2014b, 2015b; Lucy, 1992a, 2014).

The psychological engagement with linguistic relativity research is built upon two fundamental assumptions. First, similarity is the basis of categorization (Nosofsky, 1986). Second, categorization is an indispensable part of human cognition (Harnad, 1987). Following this rationale, non-verbal measures in linguistic relativity research are developed along a continuum of a wide range of cognitive tasks incorporating categorization as an inherent part. Non-verbal tasks can be further divided into two main types: higher-level cognitive tasks such as triads-matching, recognition memory and high-level reasoning, and lower-level cognitive tasks such as pre-linguistic perception, reaction time and event-related potentials (ERPs). To be more specific, higher-level processing refers to mental processes that involve post-perceptual processing, which often takes place at a later stage of processing, while lower-level processing is unconscious and automatic and usually takes place in various cognitive processes (Athanasopoulos et al., 2010; Thierry, Athanasopoulos, Wiggett, Dering, & Kuipers, 2009). For example, the similarity judgement task within the triads-matching paradigm, which is used by the current thesis, serves as one of the most successful measurements to probe speakers' cognitive processes, as perception of similarity is a complicated mental process based not only on the objective properties of the objects compared, but also on their subjective conceptual representations in speakers' mind (Bylund & Athanasopoulos, 2014b; Athanasopoulos et al., 2015). For example, using a triads-matching task, Lucy & Gaskins (2001, 2003) examined the effect of grammatical marking of number on object categorization. In each trial, participants were presented with a target picture (e.g. a wooden glass) and its two alternatives: the same material triad (e.g. a wooden bowl) and the same shape triad (e.g. a mental glass). Then participants needed to decide which of these two alternatives were more like the target triad based on their own perceived similarity. This design could avoid research from running into the language-thought circularity, as mentioned above. Empirical evidence shows that language effects on thought tend to be strongest when the task lacks an objective answer, or when speakers are uncertain about the perceptual stimulus (Flecken, Athanasopoulos, et al., 2015; Montero-Melis, Jaeger, & Bylund, 2016).

In addition, it has been argued that conclusions on whether language affects thought cannot be drawn from only one type of test, but rather from the incorporation of different types of tests and measurements (Casasanto, 2008; Flecken, Carroll, et al., 2015; Lucy, 2014). Current research is starting to introduce a wide range of psycholinguistic techniques, such as eye-tracking, event-related potentials (ERPs) and reaction time under different conditions in order to make the results more generalizable and interpretable. According to Tokowicz and MacWhinney (2005), reaction time is a subtle type of measurement which directly reflects participants' automatic, non-reflective and implicit responses in cognitive processes. This type of measurement has been widely applied in various cognitive domains to capture language-specific effects on simple, sub-conscious and perceptual decisions (Flecken, Athanasopoulos, et al., 2015; Guillaume et al., 2009; Winawer et al., 2007).

Given the methodological advances in linguistic relativity research, the study combines two research paradigms: a linguistic encoding paradigm that allows manipulating participants' short-term language activation in language production, and a triads-matching paradigm which is higher-level cognitive and has been widely used to assess participants' perceptions of event similarity and processing efficiency (as measured by reaction time) during the decision-making process.

### 2.2.3. Expansion of empirical studies in different conceptual domains

Contemporary approaches to linguistic relativity research have included a wider array of conceptual domains, such as colour (Athanasopoulos, 2009; Athanasopoulos et al., 2011; Regier & Kay, 2009; Thierry et al., 2009; Winawer et al., 2007), time (Boroditsky, 2001; Boroditsky et al., 2011; Casasanto, 2008; Casasanto & Boroditsky, 2008), gender (Kurinski & Sera, 2011), objects and substances (Cook et al., 2006; Lucy & Gaskins, 2001, 2003; Pavlenko & Malt, 2011), space (Levinson, 2001, 2003; Von Stutterheim, Bouhaous, & Carroll, 2017) and motion (Athanasopoulos & Albright, 2016; Flecken, Carroll, et al., 2015; Gennari et al., 2002; Han & Cadierno, 2010; Ji & Hohenstein, 2018; Kersten et al., 2010). Expanding linguistic relativity research to different cognitive domains allows current research to 1) generate more empirical evidence for the language-and-thought controversy and 2) investigate whether relativist theories are applicable to all cognitive domains or they are on the contrary, domain-specific.

Most of the earlier empirical evidence on language-and-thought research started with the interplay between basic colour vocabularies and colour perception (Berlin & Kay, 1969; Brown & Lenneberg, 1954). On the one hand, numerous studies show that cross-linguistic differences in colour naming have an impact on colour recognition, discrimination and categorization (Roberson, Davies, & Davidoff, 2000; Thierry et al., 2009; Winawer et al., 2007). However, other studies report no such language effects and demonstrate a universal trend in colour cognition (Berlin & Kay, 1969; Heider & Olivier, 1972). For example, Winawer et al. (2007) examined how basic colour items affect processing efficiency in colour recognition with English and Russian speakers. Russian has two basic colour items for the colour 'blue' (light blue vs. dark blue) whereas English only has one item. This cross-linguistic difference in colour naming affected participants' processing speed in colour recognition, as Russian speakers matched two shades of the same blue much more quickly than their English counterparts (200ms faster). Similar results were reported by Thierry et al. (2009), who found that Greek speakers (with two basic items for the colour 'blue') differed from their English (with only one basic colour item for 'blue') counterparts in pre-linguistic colour perception. By measuring the electric activity in the brains of speakers of different languages, results showed that Greek-speaking brains differed in neuronal activities from English-speaking brains. If cross-linguistic differences in the naming patterns affect colour perception, we could expect that speakers of Dani, a language with only two colour items: light colour and dark colour, to differ in every aspect of colour recognition compared with their English counterparts in non-linguistic colour categorization. However, this is not the case (Heider & Olivier, 1972). In fact, colour perception is largely constrained by the physical properties of different colours. Thus, cross-linguistic differences in colour naming may exert some influence on colours with subtle differences (i.e. light blue vs. dark blue), but not for the perception and categorization of the focal ones (Regier & Kay, 2009).

The mixed findings for colour naming suggest that linguistic relativity research on language effects on cognition is far from an 'all-or-nothing' issue and highlight the importance of including more cognitive domains in such investigation. In addition, other studies argue that colour naming is not a grammatical category with morphological properties as it lacks the morphological marking to highlight language-specific features in colour structures and labelling (Lucy, 1996, 2016). Research to date shows that language-on-thought effects arise from structural differences rather than the impact of general cognition (Han & Cadierno, 2010; Lucy, 2014; Slobin, 2006; Talmy, 2000). For this reason, contemporary approaches start to include a wide range of cognitive domains, especially ones with prominent structural differences, such as number, gender, and aspect marking. Among these conceptual domains, motion events have become a popular testing ground for the interplay between language and cognition as world languages differ drastically in how they select and package information about motion. A detailed discussion of the diversity in motion event encoding and its potential influence on event categorization will be conducted in Chapters 4 and 5.

### 2.3. The interplay between language and cognition: to what extent does language mediate cognition

Given the complexity of cognitive effects, this section focuses on how language affects thought under different circumstances. Current research shows that different degrees of language involvement affect non-verbal cognition. Empirical evidence shows that the effects of language are most likely to appear when the access to language is facilitated at the very moment of decision-making (Bylund & Athanasopoulos, 2014b, Slobin, 1987, 1996b; Wolff & Holmes, 2011). Such a 'thinking-for-speaking' effect, postulated by Slobin (1996b), emphasizes the online utilization of language during cognitive processing (Slobin, 1987, 1996b; Wolff & Holmes, 2011). In addition to using languages online, there are cases where speakers need to depend on recent linguistic experience to complete a subsequent task. This 'thinking-after-language' effect, a term coined by Wolff and Holmes (2011), focuses on the use of language as a strategy or means to facilitate working memory. In both cases, language is used as a meddler or augmenter to facilitate perception and categorization during cognitive processing (Lupyan, 2012; Wolff & Holmes, 2011).

This section focuses on three fine-grained hypotheses, on the interplay between language and cognition, which are widely applied in current language-and-thought research. These widely adopted frameworks are the thinking-for-speaking hypothesis (or sometimes termed 'thinking-with-language' effects), speech-planning process (Levelt, 1989; Von Stutterheim & Nuse, 2003) and the label-feedback hypothesis (Lupyan, 2012; Lupyan & Clark, 2015; Lupyan, Rakison, & McClelland, 2007). These three hypotheses share common ground in that they all attempt to outline when and under what conditions that language tends to mediate cognition.

### 2.3.1. The thinking-for-speaking hypothesis

Language may influence one specific type of thinking, namely the online thinking, which occurs during language use. This process, termed *thinking-for-speaking* (TFS) (Slobin, 1996a, 2003, 2006), proposes that effects of language on cognition only occur under circumstances when speakers are involved in language-driven activities such as language production or comprehension. In other words, it refers to a 'special form of thought that is mobilised for communication' (Slobin, 1996:76). Specifically, as Slobin explains (Slobin, 1996a, 2000), when speakers are engaged in the process of speech planning or production, they need to attend to the properties that 1) fit some conceptualization of the event, and 2) are readily encodable in language (Slobin 1987, p.435). Such influence is expected to produce differences in thought across languages because the processing of information requires speakers to fit their thinking patterns into language-specific references or frames in order to achieve the purpose of successful communication. For example, when talking about a past event, speakers of English, but not Chinese, must pay attention to when the event took place and encode the grammatical feature [+tense] in the main verb. Similar cases can be found in Turkish since, when describing a past event, Turkish speakers must specify whether the event was witnessed or not.

The rationale of the TFS framework is that speakers are engaged in language use from time to time and, therefore, research on language and cognition remains incomplete if cognitive processes related to speech production are not fully taken into consideration (Athanasopoulos & Bylund, 2013b). Current literature suggests that TFS effects are documented across different time frames and periods during language production (Slobin, 2003). The first frame or period is called the 'anticipatory' phase, which concerns the effects of language on cognition immediately before language production. During this pre-linguistic phase of speech preparation, speakers attend to language-specific dimensions and units that are directly relevant to the linguistic encoding. The second phase refers to speaking time proper, when linguistic symbols are activated in tandem with the non-linguistic representations. Thus, this process is also known as 'thinking with language' effects, or 'using language as a strategy', as appears in most of the current literature today. The third frame is situated immediately after language use or speech production and relates language-specific encoding to subsequent information processing or non-linguistic representations. Thus, such effects are also called 'consequential effects' or 'thinking-after-language' effects (see Wolff & Holmes, 2011, for a detailed review).

The crucial difference between the hypotheses of linguistic relativity and thinking-for-speaking lies in the former emphasizing the language effects on general cognition no matter whether language is involved or not, while the latter focuses on the online effects of language on cognition only during the process of speaking or language production. From a theoretical perspective, it is important to distinguish between these two approaches from the following perspectives. First, linguistic relativity puts emphasis on non-verbal behaviours, whereas thinking-for-speaking focuses on verbal performance. Thus, interpreting the thinking-for-speaking hypothesis as either a weak or new version of linguistic relativity conflicts with Slobin's original statement and definition (Athanasopoulos & Bylund, 2013b; Bylund & Athanasopoulos, 2014b; Lucy, 2016). Second, as argued by Lucy (1997) and others, these two accounts represent different levels of mental representations. For instance, the cross-linguistic difference in information selection and structuring as captured by TFS is related to a linguistic level of representations which are language-derived and context-dependent, while LR concerns a non-linguistic level of representations that are pervasive and durable. For example, empirical evidence demonstrates that cross-linguistic differences in motion event cognition are most likely to appear with explicit linguistic encoding (Hendriks & Hickmann, 2015; Slobin, 1996b; Soroli, Sahraoui, & Sacchett, 2012). However, such influence disappears when the access to language is blocked by task manipulation (Papafragou, Massey, & Gleitman, 2002; Trueswell & Papafragou, 2010). Thus, these two hypotheses have different constructs and should be regarded as distinct concepts and theoretical frameworks (Athanasopoulos & Bylund, 2013b; Bylund & Athanasopoulos, 2014b).

Although linguistic evidence shows that speakers of different languages select and organize information differently in their oral descriptions (Brown & Gullberg, 2008, 2011; Daller et al., 2011; Hohenstein et al., 2006), it has recently been questioned whether language-specific thinking-for-speaking with linguistic data or language production only is sufficient to prove that people with different linguistic patterns have different thinking patterns. As pointed out by both Athanasopoulos and Albright (2016) and Casasanto (2016), using linguistic data alone may run the risk of circular reasoning, such as : why do speakers of different languages think differently? The reason is simply because they speak differently. In order to better argue whether language can influence thinking or general cognitive patterns, it is crucial to involve different types of measurements across different modalities, such as co-verbal (i.e. gestures, eye movements and reaction time) and non-verbal measures (i.e. perception, reasoning and categorization), rather than the linguistic measures alone. This triangulation of methods is adopted by the current study as it combines different tasks in different modalities: linguistic data of motion event elicitation, non-linguistic data of motion event categorization, and co-verbal data of reaction time. Detailed information on the experimental instruments will be reviewed in Chapter 6, Methodology.

To investigate whether thinking, rather than speaking, is language-specific, contemporary approaches within the framework of TFS are beginning to combine verbal measurements with co-verbal ones such as gestures (A. Brown, 2015; Brown & Gullberg, 2008; Gullberg, 2011), eye movements (Papafragou et al., 2008; Soroli, Hickmann, & Hendriks, 2019; Trueswell & Papafragou, 2010), reaction time (Boroditsky, 2001; Von Stutterheim & Nuse, 2003) and event-related potentials (ERPs) (Boutonnet, Athanasopoulos, & Thierry, 2012; Flecken, Athanasopoulos, et al., 2015). For example, Boroditsky (2001) examined the reaction time of English and Chinese speakers when processing the truth values of temporal relations. In English, temporal relations are typically described by horizontal spatial metaphors (e.g. Christmas is ahead of us), whereas in Chinese, both vertical metaphors, like *xia* 'down' in xia *ge yue* (next month), and horizontal metaphors are used to indicate time and spatial relations. Participants were primed with different sentences of horizontal and vertical spatial expressions. After linguistic priming, participants needed to decide the truth value of a statement (e.g. March comes earlies than April) as fast as possible, with their processing efficiency being evaluated. Results suggested that English speakers were significantly faster in verifying vertical relations than horizontal ones, whereas Chinese speakers were the opposite. In light of linguistic relativity, these results suggest that language-specific mappings of spatial metaphors (i.e. vertical or horizontal) affect speakers' sensitivity and implicit processing efficiency of spatial relations, that is, speakers of different languages do not share the same reasoning patterns and use different processing strategies in decision-making, as demonstrated by the RTs.

Similar findings were reported by Papafragou et al. (2008). By applying an eye-tracking paradigm, Greek and English speakers' attention allocation was examined during the process of preparing for speech when watching different motion scenes. Results indicated that while watching different target videos of motion events (i.e. a boy is skating across an icy lake), speakers of each language allocated their visual attention to the components that were typically encoded and highlighted by their language. That is, speakers of English (a satellite-framed language where the manner of motion is encoded via the main verb) prioritized their attention on the region of the manner of motion, whereas speakers of Greek (a verb-framed language where the path of motion is encoded via the main verb) looked at the region of the path in the first instance during speech planning. However, such language effects disappeared when participants watched the same video under a free encoding condition or when a verbal interference was introduced. Via the integration of both the verbal and non-verbal measurements, the results painted a more accurate picture that cross-linguistic differences in cognition were mostly likely to appear when participants were engaged in activities of language production or comprehension, which lends support to the framework of 'thinking for speaking' or' thinking with language'.

### 2.3.2. The process of speech planning

Another framework closely related to the TFS hypothesis is the speech production model (Levelt, 1989, 1996). This account seeks to demonstrate how speakers select and structure relevant information in accordance with the typological constraints of each language. The speech production model has four basic components: conceptualization, formulation, articulation and self-monitoring (Levelt, 1989).

The first element, conceptualization, refers to ‘a process of forming temporary mental representations of complex situations and events’ and is used to account for language-specific patterns of thinking during language production (Von Stutterheim & Nuse, 2003, p256). The process of conceptualization comprises three levels of representation: the conceptual level (conceptualizer), the lemma level (formulator) and the phonological level (articulator). The process of lexicalization takes place at the conceptual level in the conceptualizer, where a pre-verbal message is transformed into temporary conceptual structures and maps onto different linguistic forms (the formulator) before articulation (the articulator). In the conceptualizer, there are four specific processes of speech planning (Habel & Tappe, 1999): segmentation, selection, structuring and linearization. In segmentation, certain components need to be extracted from the knowledge base and segmented into smaller units depending on the complexity of the situation. Then in selection, speakers choose the units that fit the situation and their corresponding linguistic components that can be used for verbalization. These two processes involve the selection of ‘what to say’ and are labelled as *macroplanning* (Levelt, 1989). The next step is structuring, where speakers need to structure the chosen components in accordance with linguistic frames or references. In this process, speakers need to select a specific viewpoint from which the event can be reported. The final step is the linearization of words and grammatical components to fit the one-dimensional medium of language. The latter two processes are concerned with linguistic formulation and are clarified as *microplanning* (Levelt, 2000), as illustrated in Figure 1.

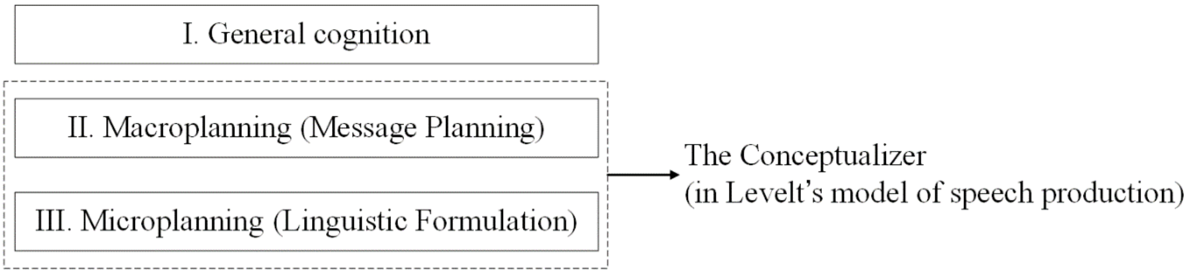



Figure 1. Three levels of conceptual representations

The speech production model shares some commonalities with the TFS framework from the following perspectives. First, both hypotheses assume that the process of speech planning is shaped by language-specific categories made available in the speaker's language system (Von Stutterheim & Nuse, 2003). Thus, language-specific categories may influence event representation and conceptualization in the speaker's mind. Second, both accounts assume that language only affects thought during language use. Thus, language-specific patterns of conceptualization are most likely to appear in the context of communication when a speaker segments, selects and structures different cognitive components in verbal encoding (Flecken, Von Stutterheim, & Carroll, 2014).

Previous studies provide compelling evidence for cross-linguistic differences of event conceptualization in segmentation, selection and temporal representations. For example, with regard to event segmentation, Von Stutterheim and Nuse (2003) investigated how speakers of Arabic, English (an *aspect language* with focus on event 'ongoingness') and German (*non-aspect language* with a focus on event completion) segmented events by using a film retelling task. Participants were required to retell 'what is happening' in a silent film while watching it. Results showed that German speakers encoded significantly fewer events compared with speakers of Arabic and English, especially for ‘smaller’ and unbounded events (i.e. looking around). Such information was more likely to be omitted by German speakers. Similar findings are reported in event selection, where many studies have examined how speakers of different languages depict goal-oriented motion during online descriptions (Carroll, Von Stutterheim, & Nuese, 2011; Carroll, Weimar, Flecken, Lambert, & Von Stutterheim, 2012a; Flecken, 2011). Results reported in these studies suggest that when describing and perceiving the same event, such as ‘A man is driving along the road towards a village’, speakers of languages without obligatory aspect markings are more likely to encode event endpoints in their oral descriptions (e.g. A man is approaching a village), whereas speaker of languages with aspect marking generally omit this information and focus more on mentioning the ongoing phrase of the same event (e.g. A man is driving along the road). In sum, these findings lay solid ground for where and how the effects of language on cognition are most likely to take place.

### 2.3.3. The label-feedback hypothesis

Another model closely related to the TFS framework is the label-feedback hypothesis (Lupyan, 2012). This account seeks to explore the dynamic relationship between language activation and linguistically-modulated conceptual representations under different conditions. For example, it aims to answer the question of why recent linguistic experience and immediate language use can affect speakers’ categorical perception, and how language can produce transient modulation during higher-level processing, such as categorization and object perception. It also proposes that such language effects on cognition are regulated in two directions: they can be either increased by recent verbal training, linguistic priming and perceptual learning, or decreased through verbal interference (Lupyan, 2012).

This hypothesis is built upon three basic tenets, that 1) categorization is one of the basic human capacities, 2) language may augment categorization, and 3) language plays a fundamental role in categorization and object perception (Athanasopoulos & Albright, 2016; Lupyan, 2012; Vanek, 2019). For example, in colour perception, the visual perception of colour (blue) activates language-specific labels/representations of 'blue' in a bottom-up fashion, while at the same time, verbal labels/representations automatically co-activate associated perpetual features in a top-down manner via a phonological loop. According to the account of short-term memory (Baddeley, 1992, 2003), a phonological loop (where language-specific labelling is assumed to be encoded) provides space for the verbal rehearsal of the visual stimuli and facilitates its transition from short- to long-term memory. Thus, the phonological loop can be flexibly up- or down-regulated by recent linguistic experience. For instance, training or priming participants with language-specific labels before presenting visual stimuli can facilitate the co-activation of linguistic and non-linguistic representations, guiding participants to categorize events or objects in accordance with the linguistic categories they are primed with. In contrast, when participants' access to language is blocked by task manipulation (i.e. under verbal interference), the mutual feedback between linguistic and non-linguistic representations is interrupted and, as a result, the cross-linguistic influence of language on cognition is greatly diminished.

Empirical evidence demonstrates that the immediate use of language-specific labels may guide participants' thinking patterns in a subsequent cognitive task, and such language-specific effects are context-bound and task-dependent (Athanasopoulos, Bylund, et al., 2015; Lupyan & Clark, 2015; Lupyan et al., 2007; Montero-Melis et al., 2016; Vanek, 2019; Von Stutterheim et al., 2017; Winawer et al., 2007). For example, Winawer et al. (2007) examined whether an obligatory colour boundary affects processing efficiency in colour discrimination. In colour perception, there is an obligatory distinction between lighter blues ('goluboy') and darker blues ('siniy') in Russian, but English has no such distinction. By using a triads-matching task, it was shown that Russian speakers reacted more quickly in discriminating between-category colours rather than within-categories colours. However, such effects disappeared under the condition where verbal interference was introduced. Similar findings were reported by Montero-Melis et al. (2016) in the domain of motion where recent linguistic priming and overt verbal-training could up-regulate the influence of language-specific labels on cognitive processing. That is, participants with manner-primed prompts were more likely to categorize motion based on manner-match criteria than those with path-primed prompts. It is suggested that conceptual representations are flexible and can be modulated by recent linguistic priming by establishing temporary visual-linguistic associations in an ad hoc fashion. This finding is further supported by Vanek (2019), who examined how L2 learners categorised motion events by manipulating different degrees of language involvement in a perceptual learning paradigm. Results suggested that language learning had a gradient impact on cognition: categorization with explicit encoding demonstrated the strongest effects, followed by free encoding, whereas categorization with verbal interference had the least effect. The results suggest that the effects of linguistic labels are context-induced and flexible, indicating an online and immediate effect of language on perceptual representations and categorization.

In sum, the current chapter summarizes three contemporary approaches to linguistic relativity research with supportive empirical evidence from various cognitive domains. The overall findings from existing research suggest that language can exert a strong impact on thought but has different roles to play, depending on when and where the access to language is facilitated: when thinking occurs before or during speaking (thinking-for-speaking and speech-planning), or when thinking and language compete with each other in tandem (language as a meddler in the 'label-feedback' hypothesis), or thinking occurs after language where language is used as a strategy or meddler to solve a subsequent cognitive task (thinking-with-language or thinking-after-language). This will shed light on the theoretical and methodological innovations in current language-and-though research.

# Chapter 3. Language and thought in bilingual speakers: the effect of language learning on cognitive restructuring in the bilingual mind

As most of the world's population speak more than one language (Cook, 2003; Cook et al., 2006), combining contemporary approaches to linguistic relativity research with the domain of bilingualism has become a new trend to tap into the interplay between language and cognition in bilingual speakers, L2 speakers or foreign language learners. Adopting a bilingual perspective is regarded as a natural extension of linguistic relativity research: if speakers of different languages have different thinking patterns, then the question is: how is the world represented by bilinguals who have more than one contrastive language system? This offers us a unique opportunity to examine the interplay between language and cognition from a new perspective as bilinguals are the ones who are directly experiencing linguistic relativity effects.

This section will combine the latest findings of linguistic relativity research with the multi-competence framework of bilingualism and focus on the dynamics between L2 learning and the changing cognitive status of bilingual speakers. Based on the framework of cognitive (or conceptual) restructuring and associative (or attentional) learning, this chapter further discusses the linguistic and extra-linguistic factors that underpin the recalibration of language-specific thinking patterns during L2 or additional language learning.

## 3.1. Expanding the language-and-thought debate to bilinguals and L2 learners

The fact that language can modulate cognition in various ways raises many intriguing questions: if speakers of different languages have different modes of thinking, how do bilinguals with two typologically different languages behave? Do they have two independent modes of thought and behave like monolinguals of each language? Or do they develop a convergent mode of thinking integrating typical features of all the languages they know? These questions are concerned with the learnability of L2-specific ways of thinking and the permeability of already established patterns associated with the L1. Thus, expanding the boundary of language-and-thought research to the domain of bilingualism and SLA will not only facilitate our understanding of the process of addition language learning, but also provide unique insights into the dynamics between language and cognition, which are not easily to be observed in monolingual speakers. Despite this being of great importance, the language-and-thought research on speakers of two or more languages remains quite limited (Bassetti & Cook, 2011; Bylund & Athanasopoulos, 2014b; Daller, Treffers-Daller, & Furman, 2011; Flecken, Gerwien, Carroll, & Von Stutterheim, 2015; Jarvis, 2016; Jarvis & Pavlenko, 2008).

The interface between language and cognition in speakers of two or more languages is closely related to 1) the extent to which bilinguals restructure their linguistic encoding and non-linguistic cognition patterns as a result of L2 acquisition; 2) the transfer phenomena (forward vs. backward transfer) that characterize the interaction between linguistic and non-linguistic representations; and 3) linguistic or extra-linguistic factors that modulate a learner's cognitive behaviour during L2 learning.

#### 3.1.1. Combing the framework of multi-competence with bilingualism

To systematically examine the interaction between language and cognition in bilinguals or L2 speakers, current research adopts different frameworks with a wide range of experimental methods. One of the most acknowledged theoretical framework is the frameworks of multi-competence (Cook, 1992, 1999, 2003; Cook & Li, 2016), which views bilinguals as independent learners with unique linguistic and cognitive features.

The successful application of the multi-competence framework into language-and-thought research can be attributed to the following reasons. First of all, in line with the linguistic relativity hypothesis, the multi-competence framework assumes that language and cognition are highly interactive (Cook, 2016). Thus, cross-linguistic differences will not only be manifested in the linguistic domain, but also in non-linguistic conceptual representations. In this way, this framework brings together two lines of cognitive research, namely, bilingual linguistic encoding and bilingual conceptual representations. Second, this framework does not view bilinguals as a simple sum of two monolinguals in one head, but instead, assumes them to be independent, multi-competent speakers of unique representations of the world (Pavlenko, 2016). This allows us to shift from regarding monolinguals or native speakers as the 'social norm' for investigation and provides a ready platform to explore the unique cognitive characteristics that bilinguals have. In addition, in line with the framework of 'thinking-for-speaking', the multi-competence framework also views conceptual representations in bilinguals as a dynamic process and situated along a continuum of integration and separation, where different language-specific concepts interact and merge with each other (Cook, 2002; 2003; 2016). For example, at one end of the continuum, language-specific concepts are separate from each other within the bilingual mind, whereas at the other end, language-specific concepts merge into new categories. The degree of separation and integration of different conceptual categories that bilinguals have depends on various individual and contextual factors, such as language proficiency, frequency of language use and the amount of language exposure, as well as short-term experimental manipulation, such as linguistic priming, biased language instructions and verbal interference. This account provides a theoretical background for the investigation of cross-linguistic influence (or language transfer) in bilinguals or L2 speakers and sheds light on how language learning affects bilinguals' both verbal and non-verbal performances from a new perspective.

Following the multi-competence framework which regards speakers of more than one language as an independent language user (Cook, 200, 2003) and adopting a user-based approach (Smith & Samuelson, 2006), bilingualism in the current study is defined as someone 'who regularly needs to understand or use more than one language at home and at school' (Frederickson & Cline, 2002, p. 246), while multilingualism refers to someone who knows and uses more than two languages in daily communication and interaction. Thus, bilingualism and multilingualism in the current thesis are conceptualized in terms of the total number of languages someone is exposed to, and the frequency of use with each language on a daily basis. Due to a high frequency of language use, bi- and multilingual speakers in the current thesis usually have achieved a pretty high level of language proficiencies in each of their languages. Following Bylund and Athanasopoulos (2015b), the current thesis did not qualitatively differentiate bilingual speakers from L2 learners as long as the participants' language learning trajectories meet the definition here.

### 3.1.2. Language learning and cognitive restructuring in the bilingual mind

The Conceptual Transfer Hypothesis (Jarvis, 2011; Jarvis & Pavlenko, 2008) within the cognitive domain suggests that a 'certain instance of cross-linguistic influence in a person's use of one language originates from the conceptual knowledge and patterns of thought that the person has acquired from another language' (Jarvis, 2007, p.44). This analytical framework assumes that speakers of different languages have distinct sets of conceptual categories and mental representations of the reality. In the case of bilingual speakers and L2 learners, these conceptual differences in one language may trigger transfer effects  (either forward or backward) across languages, which may affect the learning and use of an L2 or additional language (Jarvis, 2007; Jarvis & Pavlenko, 2008; Pavlenko, 2011).

Conceptual transfer is a dynamic and multi-directional process where L1 may affect and be affected by additional language learning. As suggested by the framework of multi-competence, conceptual categories in a bilingual mind are presented along an integration continuum where language-specific concepts may integrate and merge into unique presentations sharing features of different languages (Cook, 2002; 2003; 2016). Thus, conceptual transfer in bi-multilingual speakers exhibits different forms in terms of directionality, such as forward transfer from L1 to L2, reverse transfer from L2 to L1, and bidirectional transfer between L1 and the L2. Although great efforts have been put to examine how L1- based conceptual transfer affects L2 learning, effects from L2-mediated cognition patterns on the L1 remains to be a neglected area in the field of bi- or multilingual cognition (Bylund & Jarvis, 2011).

However, the mechanism of cross-linguistic influence in the conceptual domain is far more complicated than just focusing on its directionality. Empirical evidence shows that acquiring a new language can lead to conceptual changes in one's thinking. When speakers learn a new language, they not only need to learn novel linguistic references or frames, but also associated conceptual distinctions (Jarvis & Pavlenko, 2008; Pavlenko, 2011). This may give rise to the restructuring of existing conceptual categories associated with the L1 (Jarvis & Pavlenko, 2008; Pavlenko, 2011). This process, termed conceptual or cognitive restructuring, refers to conceptual changes that bilinguals undergo in the process of acquiring a new language. It is a gradual process and occurs in both bilinguals' both verbal and non-verbal behaviours (Jarvis & Pavlenko, 2008; Pavlenko, 2011, 2014).

The conceptual changes bilinguals have are gradient and exhibit various forms, such as the co-existence of L1- and L2-based concepts (Hohenstein et al., 2006; Sachs & Coley, 2006), the convergence of the L1- and L2-based concepts, (Brown & Gullberg, 2013; Cook et al., 2006), shift to L2-based concepts (Athanasopoulos, Damjanovic, et al., 2015b; Park & Ziegler, 2014) and the attrition of L1-based concepts (Bylund, 2009b; Bylund & Jarvis, 2011), with details of these discussed below. It should be noted that these seven types of conceptual changes are distributed along a continuum without an absolute fixed order. In other words, cognitive restructuring in the bilingual mind is an ongoing process which occurs either simultaneously or developmentally.

1) Influence of L1-based concepts
2) Internalization of L2-based concepts
3) Co-existence of L1- and L2-based concepts
4) Convergence of L1- and L2-based concepts
5) Shift to L2-based concepts
6) Influence of L2-based concepts on L1-based concepts
7) Attrition of L1-based concepts

The first stage of cognitive restructuring during L2 learning is the stage when no restructuring is evident and speakers continue to follow L1-based patterns for thinking, seeing and speaking. Current literature on L1-based influence has been well-documented in various cognitive domains, such as objects (Athanasopoulos, 2006; Cook et al., 2006), spatial concepts (Levinson, 2003) and motion events (Finkbeiner et al., 2002; Hendriks, Hickmann, & Demagny, 2008; Von Stutterheim & Nuse, 2003).

Then conceptual restructuring proper begins with the destabilization of L1-based concepts where speakers diverge from their original thinking patterns to accommodate for newly-established references or frames in the target language. This may bring about the incorporation and internalization of new words, phrases or conceptual categories for target-like performance from both linguistic and cognitive perspectives (Jarvis & Pavlenko, 2008). To date, evidence for the internalization of L2-based concepts comes from various cognitive domains such as spatial concepts (Park & Ziegler, 2014), motion events (Athanasopoulos & Kasai, 2008; Daller et al., 2011) and emotion (Panayiotou, 2004a, 2004b; Pavlenko, 2003). One typical example related to the internalization of new concepts comes from Pavlenko (2003), who examined Russian-English bilinguals' expressions of abstract concepts. Results indicated that bilinguals with an intermediate level of proficiency began to use expressions denoting the notion of 'privacy', which is absent in Russian but available in English, suggesting an internalization of new concepts as a result of L2-English learning. In addition, one typical question related to the process of internalization concerns the relationship between newly-internalized concepts and already-existing concepts. Further investigation is needed to examine whether the integration of new concepts may lead to an overall or a partial change in the cognitive state of the bilingual mind (Pavlenko, 2011, 2014).

Another process that is closely related to internalization is the co-existence and convergence of L1- and L2-based concepts. Co-existence refers to the maintenance of two or more respective sets of language-specific frames or conceptual categories in agreement with the typical constraints of each language. One example of immediate relevance is the study by Sachs and Coley (2006), who explored Russian-English bilinguals' conceptualizations of different emotional concepts (i.e. envy and jealousy).

It was reported that bilinguals' conceptual representations switched in accordance with different language contexts, that is, bilinguals demonstrated an English-way of conceptualizing envy when tested in an English context, but patterned with Russian monolinguals when tested in a Russian context. Similar findings have been reported in other conceptual domains, such as shape (Barner, Inagaki, & Li., 2009), space (Levinson, 2001, 2003) and motion event construal (Athanasopoulos, Bylund, et al., 2015; Bylund 2011). These results showed that bilinguals exhibited language-specific cognitive patterns from both L1 and L2, indicating an ongoing process of co-existence of different conceptual categories in the bilingual mind.

With an increase in the degree of being bilingual, the L1- and L2-based conceptual categories may converge with each other and lead to a unitary conceptual category integrating characteristics of both the L1 and L2 (Jarvis, 2011). This process, termed convergence, has two different manifestations. The first is that speakers may exhibit conceptualization patterns incorporating both L1 and L2 features and display an 'in-between performance' (Pavlenko, 2011c, p.247). One typical example was reported by Park (2019) in motion event encoding. Results suggested that Korean-English bilinguals performed differently from Korean and English monolinguals in the frequency of manner encoding when describing voluntary motion: bilinguals used more manner verbs than Korean monolinguals, but fewer manner verbs than English monolinguals, indicating an 'in-between' performance between Korean and English monolingual speakers. Meanwhile, convergence can also be manifested as neither L1- nor L2-based, also called divergence. Bassetti and Cook (2011) suggest that convergence does not necessarily only take the form of the 'in-between' performance. Concepts diverging from either language can also be viewed as instances of convergence. For example, Park and Ziegler (2014) examined the acquisition of spatial concepts in Korean-English bilinguals with varying degrees of English proficiency. Results showed that bilinguals' categorical preferences of space (put in vs. put on) resembled neither L1- nor L2-based patterns. Further details demonstrated that they had formed a unique category of spatial cognition distinct from either of the monolingual concepts.

The interaction between different conceptual categories within the bilingual mind may lead to L2-based conceptual changes; this process, termed a shift to L2-based concepts, refers to a process whereby, in order to reduce the cognitive load of maintaining two separate conceptual systems, bi- or multilingual speakers are prone to shift from their L1-based concepts, but reassembling, albeit not necessary fully, to L2-based concepts during the process of language learning. Empirical evidence suggests that this L2-based restructuring can be manifested in speakers' verbal and non-verbal behaviours in various conceptual domains, such as colour (Athanasopoulos, 2009; Athanasopoulos et al., 2011), objects and substances (Athanasopoulos, 2006; Cook et al., 2006; Malt & Sloman, 2003) and motion event encoding and gesturing (A. Brown, 2015; Cadierno, 2010; Carroll, Weimar, Flecken, Lambert, & Von Stutterheim, 2012b). For example, Athanasopoulos and Kasai (2008) reported a cognitive shift of English-Japanese bilinguals from L1- to L2-based cognitive patterns in object categorization. By using a triads-matching task, results indicated that advanced bilinguals patterned with Japanese monolinguals in grouping objects in terms of shape, while speakers of intermediate levels of proficiency maintained an L1-based preference for categorizing objects based on colour. The results suggested that L2 proficiency was a strong predictor in modulating the process of bilinguals' cognitive restructuring towards an L2-based pattern of conceptualization.

In addition, prolonged exposure and increased proficiency in the L2 or additional language may lead to another type of conceptual change, that is, restructuring of L1-based concepts under the influence of an L2, which involves ‘a partial modification of already existing language-mediated conceptual categories’ (Jarvis & Pavlenko, 2008, p. 160). The experience of language learning, especially for languages with typologically different properties, requires learners to adjust concepts or categories already established in the L1 to form new conceptualization patterns associated with the L2. Unlike the aforementioned conceptual changes, the restructuring of L2-based concepts is an instance of backward influence in terms of the directionality (reverse transfer). It is an under-explored area with limited evidence from different cognitive domains. To date, empirical evidence for L2 influence on the L1 is from object naming (Pavlenko & Malt, 2011), spatial concepts (Park & Ziegler, 2014), motion event encoding (Hohenstein et al., 2006; Papafragou et al., 2008), motion event gesturing (Brown & Gullberg, 2008; 2010) and motion event cognition (Bylund, 2009a; Bylund and Athanasopoulos, 2014a, Bylund et al., 2013; Bylund & Jarvis, 2011; Wang & Li, 2019). As this process is closely related to the current thesis, more examples will be given in Chapter 5 with regard to the specific domain of motion events.

The final stage of cognitive restructuring is related to the attrition of L1-based concepts as a result of L2 or additional language learning. Pavlenko (2011) reported that Russian-English bilinguals living in the US no longer replied using the Russian emotion category ‘perezhivat’ (suffer things through) when expressing sorrow in an elicited speech compared with Russian monolinguals. Similar findings were reported by Athanasopoulos (2009), that is, Greek-English bilinguals demonstrated a weakening of obligatory conceptual contrasts between different shades of blues (light vs. dark) in colour perception. These findings demonstrated an ongoing process of attrition in both linguistic and non-linguistic domains for bilingual speakers. However, future studies need to adopt a longitudinal approach to capture actual processes of cognitive shift from target-like performances to non-target like performances in bilingual speakers or L2 learners (Pavlenko, 2011).

In sum, the multifaceted picture shows that bilinguals' conceptual representations are dynamic and multimodal across different modalities (verbal, co-verbal and non-verbal), depending on where the participants locate along the continuum of bilingualism. However, it remains unclear why and how the learning of another language can affect speakers' mental presentations of different concepts, as well as the various mechanisms underpinning the process of cognitive restructuring in bilingual speakers or L2 learners.

### 3.2. Mechanisms beyond cognitive restructuring: an account of associative learning

In order to address how linguistic knowledge may give rise to the relativistic effects in bilingual speakers or L2 learners, a usage-based approach to language learning, namely the associative and attentional learning account, has been successfully applied (Bylund & Athanasopoulos, 2014b; Casasanto, 2008; Ji & Hohenstein, 2018; Kersten et al., 2010). The connections between cognitive restructuring and the associative learning account are built upon the observations that conceptual representations are multimodal, dynamic and highly inter-correlated across different modalities (Athanasopoulos, Damjanovic, et al., 2015b; Casasanto, 2008). The psychological account of associative and attentional learning (Samuelson, 2002; Smith, 2010; Smith & Samuelson, 2006) concerns the process of learning as a process of forming different instances of form-meaning associations. Take the domain of motion as an example, the presence or absence of grammatical marking (i.e. progressive marking) in different languages tends to modulate the degrees to which speakers express and allocate their attention to event trajectories or endpoints (Bylund & Athanasopoulos, 2014a; Flecken, Carroll, et al., 2015; Von Stutterheim, Andermann, Carroll, Flecken, & Schmiedtova, 2012). For instance, speakers of an aspect language (i.e. languages with grammatical means to present aspectual contrasts) are less prone to mention endpoints in event lexicalization and focus more on the ongoingness of the target event, while speakers of a non-aspect language (i.e. languages without a grammatical means to encode aspect) tend to express event endpoints more frequently and adopt a holistic perspective in event categorization and perception. These findings suggest that languages may direct the speaker's attention to the more codable elements that are made salient by language.

Neuro-scientific evidence also shows that linguistic and non-linguistic representations in the bilingual mind are highly dynamic and inter-connected. For example, when speakers process colour words such as blue or green, there is an unconscious, automatic and concurrent activation of vision and language expected in an area of the brain (Regier & Kay, 2009; Siok et al., 2009; Lupyan, 2012). Similarly, when participants are processing motion words such as walk, kick or run, a similar sensorimotor cortex will be activated by language-specific symbols (Pulvermüller, Shtyrov, & Ilmoniemi, 2005). Empirical evidence shows that language-specific representations are built up, or emerge, in an up-regulation fashion, due to continuous exposure to numerous form-meaning pairings (Lupyan, 2012; Smith & Samuelson, 2006). Then linguistic relativity effects arise, as a result of the co-occurrence of frequently occurred associations between linguistic and non-linguistic patterns. Thus, continuous exposure to novel events throughout one's lifetime will facilitate learners to form new form-meaning associations based on statistical regularities of co-occurring associations in different contexts. From the perspective of L2 learning, the main concern is how the internalization of L2-specific associations interacts with existing L1-based associations, which may result in constant restructuring of conceptual categories in the bilingual mind (Athanasopoulos, Damjanovic, et al., 2015b). In fact, cross-linguistic differences in conceptual representations are affected by the degree of exposure to language-specific associations, that is, the more conventionalized an association becomes, the more efficiently it is retrieved from memory and unitized for the purpose of categorization (Langacker, 2000, 2008).

From an associative learning perspective, the process of conceptual changes during language learning can be interpreted from three perspectives. First, the associative learning account suggests that linguistic and non-linguistic associations are context-bound. For example, in object naming, different grammatical cues tend to guide the speaker's attention to different aspects of the same object (cake or cakes). Similar findings are reported for the role that linguistic priming plays in motion event construal (Athanasopoulos & Albright, 2016; Montero-Melis et al., 2016). For example, Athanasopoulos and Albright (2016) explored the relationship between grammatical aspect and event categorization with monolinguals of English [+aspect] by using a perceptual learning approach. Results suggested that English monolinguals could be trained to categorize motion in a novel way when presented with grammatical cues of non-aspect marking. These findings suggest that experimental contexts play an important role in modulating the effects of language on conceptual representations. Second, the associative and attentional learning account views the similarities in categorical preferences as statistical, but far from absolute and sufficient (Smith & Samuelson, 2006). In fact, when learning a new language, learners need to figure out the statistical regularity of a certain association and the conditions where it is most likely to appear. However, this tendency is never absolute. Thus, cross-linguistic difference from one language to another is a matter of relative degree, instead of being on a definite basis. Last but not least, as speakers are continuously exposed to new experiences and situations throughout their whole lifetime, already-established associations are open to ongoing changes and restructuring in accordance with learners' language-learning trajectories, such as age of language acquisition, language proficiency and language proficiency, as well as the language-specific domains under investigation.

As mentioned above, cross-linguistic differences in conceptualization demonstrate a liner and gradual process. This process is context-bound and highly open to individual differences. One key factor is related to the frequency of exposure to a specific form-meaning association. Empirical evidence shows that conceptual representations are subject to constant changes as a result of sufficient language exposure and frequent language use (Athanasopoulos et al., 2011; Bylund & Athanasopoulos, 2014a; Bylund et al., 2013; Park, 2019). In addition, the degree of cognitive restructuring of the form-meaning associations is modulated by various extra-linguistic factors, such as age of L2 acquisition (Athanasopoulos, 2009; Boroditsky, 2001; Lai et al., 2014), L2 proficiency (Athanasopoulos, Damjanovic, et al., 2015b; Ji, 2017; Park & Ziegler, 2014), language context (Filipović, 2011; Montero-Melis et al., 2016) and length of immersion in an L2-speaking community (Cook et al., 2006; Daller et al., 2011; Park, 2019), which will be discussed in more detail in the next section.

### 3.3. Predictors of cognitive restructuring in the bilingual mind

As supported by the associative and attentional learning account, conceptual or mental representations of language-specific categories within the bilingual mind are multimodal and highly interactive, and subject to changes across different contexts and modalities. Thus, the degree of cognitive restructuring during language learning is closely related to the degree of variation in learners' language learning history, as well as short-term experimental manipulation. This section will focus on four factors that may modulate the degree of cognitive restructuring in the bilingual mind; they are long-term biographical variables, such as language proficiency, language use with different languages and age of acquisition, and short-term language mediation of language context (or language mode, as suggested by Bylund and Athanasopoulos (2015a). These variables will be well controlled and closely examined when designing my own study.

### 3.3.1. Language proficiency

Language proficiency usually refers to the general proficiency bilinguals have with each of their languages (Bylund & Athanasopoulos, 2014a; Pavlenko, 2011). Empirical evidence shows that language proficiency plays an important role in modulating different degrees of cognitive restructuring in various cognitive domains, such as colour (Athanasopoulos, 2006; Athanasopoulos 2007; Malt & Sloman, 2003; Park & Ziegler, 2014), motion (Athanasopoulos, Damjanovic, et al., 2015b; A. Brown, 2015), and selective attention (Chen, Su, Lee, & Seaghdha, 2012). For example, Athanasopoulos (2006) examined how number was cognitively represented in Japanese-English bilinguals with varying degrees of L2 proficiency by measuring their sensitivity to countable and non-countable nouns. Results suggested that speakers of high L2 proficiency patterned with English monolinguals [+grammatical plural number marking], in showing more sensitivity to number differences between countable and non-countable objects, whereas speakers of intermediate L2 proficiency patterned with Japanese counterparts [-plural number marking], showing no such sensitivity. Similar findings were reported by Athanasopoulos (2007), who found that increased number marking knowledge was closely correlated with the selection criteria speakers used for object categorization (substance vs. shape). These results are in line with the associative and attentional learning account that increased proficiency in the L2 will provide learners with statistically more instances of L2-based cues, which will strengthen the form-meaning associations typical in that language.

However, other studies have failed to report proficiency effects on verbal encoding and non-verbal cognition (Athanasopoulos, 2009; Bylund & Athanasopoulos, 2014a; Cook et al., 2006; Daller et al., 2011). For example, Cook et al. (2006) examined Japanese-English bilinguals' categorization preferences for objects, whether shape-based or material-based. Results showed that all bilinguals preferred to use material similarity as the classification criterion in simple object categorization regardless of their L2 proficiency.

These discrepancies in the overall results can be attributed to various factors, such as 1) how L2 proficiency is measured (i.e. by self-reported scores or language proficiency tests), 2) in what ways it is measured (i.e. domain-specific language proficiency or general language proficiency, as well as a possible proficiency threshold where the effect of language on cognition starts to appear. With regard to the third factor, current literature shows that on the one hand, a moderate level of language proficiency can exert an impact on the cognitive restructuring of one's thinking patterns (Athanasopoulos, 2006; Ji, 2017, 2018; Park, 2019). For example, Ji (2017) examined how Chinese learners of English at different proficiencies conceptualized and processed motion events. Results showed that learners with intermediate proficiency started to show a cognitive shift towards the L2-based patterns in talking and thinking about motion.

However, other studies failed to show the effect of language proficiency on cognitive restructuring, especially with highly advanced language learners (Athanasopoulos, 2009; Bylund & Athanasopoulos, 2015a; Cook et al., 2006). For instance, Bylund and Athanasopoulos (2015a) examined the degree to which foreign language learners with high proficiency restructured their categorization preferences of immediate temporal frame towards the target language. Results suggested that no effects were found for the factor language proficiency, indicating that the effect of proficiency on cognitive restructuring was not a linear process. In other words, there may be possible intervals where the effects of proficiency are the most prominent. But when the proficiency level past this certain point, the potential effects may level out.

Given the mixed findings on the impact of language proficiency on cognitive restructuring, it is important to take into consideration different types of measurements used in different studies, as well as a possible proficiency threshold when interpreting the results.

### 3.3.2. Language use

Another factor that is closely related to the degree of cognitive restructuring is language use. Language use is defined as the total amount of time or contact that bilingual speakers or L2 learners have with each of their languages. In a narrower sense, it sometimes refers to the frequency of use of the L2, especially within the context of second or foreign language acquisition. Empirical evidence shows that language contact that different languages bilinguals have plays an important role in modulating different degrees of conceptual changes (Bylund & Athanasopoulos, 2014a, 2015c; Bylund et al., 2013; Daller et al., 2011; Flecken, 2011; Park, 2019). For example, Bylund and Athanasopoulos (2014a) examined how L1-isiXhosa (a language without obligatory aspect marking) multilinguals categorized voluntary motion under the influence of acquiring two aspect languages. Results suggested that the more frequently the aspect languages were used, the more likely speakers were to shift away from the L1-based associations and to pattern with speakers of aspect languages (L2/L3) in event categorization. Similar results were reported by Flecken (2011) in the examination of motion events with early bilinguals of L1-Dutch [+progressive aspect] and L2-German [-progressive aspect] by using an eye-tracking paradigm. By examining participants' verbal encoding and co-verbal attention allocation, results showed that when conceptualizing the same event (e.g. A woman is walking along the road towards a village), early Dutch-German bilinguals focused more on completion of the event and looked longer at event goals (e.g. A woman walks to a village), whereas Dutch monolinguals looked longer at the movement and focused more on the action itself (e.g. A woman is walking). The results also showed a close link between the frequency of use of specific grammatical features and attention allocation to specific aspects of motion, indicating a backward influence from L2 on L1 in event categorization and attention allocation. These findings are in line with associative learning in that the readjustment of originally established categories is related to the cumulative experience that speakers have with each of their languages (Bylund & Athanasopoulos, 2014a). Thus, the frequent use of newly established language-specific categories will strengthen corresponding language-specific form-meaning associations, whereas more infrequent use of the target language tends to weaken these associations.

To date, the factor of language use is often operationalized through self-reported scores on a daily or weekly basis via a language background survey. However, it is necessary to bear in mind that language contact may sometimes indirectly correlate with other predictors, such as general language proficiency, and the onset of L2 acquisition, as well as the length of immersion in a foreign country (Jarvis, 2007). Thus, in order to disentangle the effect of language contact from various other factors, further studies need to conduct more systematic analyses to assess the internal relationships across different factors when addressing their functions in the process of cognitive restructuring (Bylund & Athanasopoulos, 2015b).

In addition, most studies operationalize language contact as the amount of language use with either the L1 or L2, assuming that speakers use their other language in the rest of the time. This way of operationalization only works for bilingual speakers, i.e. with only two languages (Park & Ziegler, 2014). However, for multilingual speakers of more than two languages in a multilingual context, it is essential to keep track of all the languages they know as they may have more than one L1 and multiple L2s due to complicated learning backgrounds and profiles (Bylund & Athanasopoulos, 2015b). From this perspective, the current study keeps a record of the language use in all the languages that bi- and multilingual speakers have.

### 3.3.3. Age of language acquisition

The age of L2 acquisition refers to the age of onset of when bilingualism began. Empirical evidence suggests that, on the one hand, age effects have been well documented in various domains such as colour (Pavlenko, 2012), object naming (Ameel , Storms , Malt , & Sloman, 2005; Malt & Sloman, 2003), and motion events (Bylund, 2009a; Kersten et al., 2010; Lai et al., 2014). However, other studies have failed to document such age effects on bilingual cognition (Athanasopoulos, 2009; Bylund et al., 2013). For example, Pavlenko and Malt (2011) reported that Russian-English bilinguals who started learning English between the ages of 1 and 6 years were most susceptible to L2 influence in the object naming. Bylund (2009a) also showed that Spanish-Swedish bilinguals who started Swedish before the age of 12 showed the most observable patterns of L2 influence in motion event encoding. The results seem to suggest that L2 influence on L1 is more observable in early bilinguals who started L2 learning before puberty. However, the internalization of new concepts is also found in late bilinguals and adult learners (Pavlenko, 2003; Pavlenko & Malt, 2011). These findings suggest that speakers seem to maintain some degree of brain plasticity throughout their whole lives and the process of cognitive restructuring should not be limited to early bilinguals and L2 learners.

In addition, the age factor sometimes correlates with other predictors, such as language context and language proficiency (Athanasopoulos, 2009; Kersten et al., 2010; Lai et al., 2014). For example, Boroditsky (2001) examined how Mandarin-English bilinguals conceptualized the notion of time. Results suggested that bilinguals who started L2 acquisition late were more likely to categorize temporal relations based on vertical cues such as ‘shang’ up and ‘xia’ down (typical of Chinese patterns of categorization). However, bilinguals with earlier L2 acquisition displayed L2-based ways of conceptualizing time as ‘before’ and ‘after’. One possible explanation is that the age of acquisition may correlate with language proficiency, as it is generally the case that the earlier speakers start learning a language, the more frequently they use that language, and the more likely they are able to achieve higher levels of proficiency in that language (Athanasopoulos, 2009; Pavlenko, 2011). Thus, further studies need to adopt a more systematic approach and rigorous statistical methods in examining whether this age factor serves as a mediating factor and is associated with other variables (i.e. language proficiency or language use) in modulating the process of cognitive restructuring within the bi- and multilingual mind.

### 3.3.4. Language context

Language context, sometimes termed bilingual mode, is another predictor for different degrees of conceptual changes that bilinguals may have (Grosjean, 1998). Unlike the previously mentioned long-term factors that are closely related to bilinguals' language learning trajectory, this factor is closely related to different degrees of language activation and inhibition during cognitive processing and is regarded as a short-term effect and a function of language mediation. Thus, this factor is closely related to recent L2 exposure and is often operationalized by different experimental manipulation, such as linguistic priming (Lai et al., 2014; Montero-Melis et al., 2016), biased instruction (Athanasopoulos 2007; Boroditsky, Ham, & Ramscar, 2002; Brown & Gullberg, 2008; Kersten et al., 2010) and language mode (Athanasopoulos, Bylund, et al., 2015; Stocker & Berthele, 2019).

For example, Boroditsky et al. (2002) examined how bilinguals categorized action events. Results suggested that Indonesian-English bilinguals switched between language-specific categorization patterns as a function of language in operation. Participants receiving instructions in English behaved more like English monolinguals compared with those who received language instructions in Indonesian. Similar findings were reported in the domain of motion events (Athanasopoulos, Bylund, et al., 2015; Kersten et al., 2010; Lai et al., 2014; Montero-Melis et al., 2016; Vanek, 2019). For example, Athanasopoulos, Bylund, et al. (2015) reported that English-German bilinguals switched between end-point and trajectory preferences in motion event categorization as a function of biased instructions. That is, bilinguals receiving task instructions in English tended to pattern with English monolinguals in preferring event trajectories, while those receiving German instructions patterned with their German counterparts in preferring event endpoints.

However, other studies have not reported similar results as a function of short-term language mediation (Athanasopoulos, 2007; Filipović, 2011; Lai et al., 2014; Wang & Li, 2019). For example, Athanasopoulos (2007) examined how Japanese-English bilinguals categorized objects in different language contexts. Results showed that no significant differences were found for bilinguals with either L1-or L2-based instructions. That is, both groups demonstrated 'in-between' cognitive patterns irrespective of different language contexts. Similar findings have reported in the domain of motion. For instance, Lai et al. (2014) examined how English-Spanish bilinguals categorized voluntary motion in two language contexts: an English- and German-based context of linguistic priming. Results suggested that no differences were reported for bilinguals in the English- and German-priming context. They displayed L2-based categorization patterns regardless of which language was involved in their decision-making process.

These mixed findings suggest that bilinguals' mental representations are highly multi-modal and context-dependent. Under the view of associative and attentional learning, the effects of language context on cognitive processing reflect a 'contextually cued shift in attention' (Bylund & Athanasopoulos, 2015a). However, this tendency is statistical rather than absolute. And therefore, whether bilinguals can switch between different conceptualization patterns of different languages largely depends on whether they have established distinct sets of conceptual representations while learning an L2 or additional language (Bylund & Athanasopoulos, 2014b). Thus, it goes back to the original question posed at the beginning of this chapter, which concerns the mental representations of different conceptual categories in the bi- and multilingual mind. That is, how do speakers of two typologically different languages tend to behave? Do they have two independent modes of thought and behave like monolinguals of each language? Or do they develop a convergent mode of thinking that integrates the typical typological features of all the languages they know? Based on this controversy, the current study aims to explore how Cantonese-English-Japanese speakers of partial overlap language systems tend to behave in motion event lexicalization and conceptualization under different language contexts (i.e. a monolingual context where the L1 is the only activated language and a bilingual context where both the L1 and L2 are activated). Detailed information on motion event lexicalization across three target languages will be reviewed in the next chapter.

To sum up, these mixed findings suggest that cognitive restructuring in the bilingual mind is complex and dynamic. Thus, it is of great importance to take various factors into consideration when accounting for the effects of language on cognition in the context of bilingualism or multilingualism.

# Chapter 4. Motion event encoding in Cantonese, English and Japanese: linguistic devices and their effects on cognition

As explained in Chapter 1, the present study focuses on the processing of motion events with several reasons for it. First, motion events play a critical role in people's conceptualization of the world. Not only our perception of motion constitutes the basic and omnipresent part of our daily life (i.e. navigation, communication for spatial and non-spatial meanings); our conceptual understanding of motion also serves as a fundamental and indispensable part in human cognition (Jackendoff, 1990; Levinson, 2003). In addition, the theory of Cognitive Semantics (Tamly, 1985; 2000) proposes that languages across the world differ typologically in how motion events are typically expressed (i.e. via different grammatical forms and syntactical structures). Thus, extending motion event cognition to the domain of multilingualism will shed light on how human cognition functions in processing complex and dynamic linguistic systems such as motion events.

This section will focus on language-specific patterns of motion event encoding in three target languages (i.e. Cantonese, English and Japanese) from a linguistic and cognitive perspective. By comparing and contrasting language-specific features of motion event encoding in three target languages, it aims to address the questions of whether, and how these language-specific ways of expressing motion correlate with how motion is conceptually represented in cognition by introducing the hypothesis of Manner Salience. At the end of this section, predications will be made concerning how language-specific properties in lexicalization tend to influence categorization in the target language of Cantonese, English and Japanese.

## 4.1. Talmy's bipartite typology on motion event encoding: from a typological perspective

Among different cognitive domains, the domain of motion serves as a suitable testing ground to explore the interplay between language and cognition as world languages exhibit great diversities in how motion is typically expressed. A motion event refers to 'a situation containing movement of an entity or maintenance of an entity at a stationary location' (Talmy, 1985, p.60). A motion event is composed of four internal semantic elements: Figure, Ground, Path and Motion.

- Figure: a salient moving object
- Ground: a reference frame or object related to which the figure moves
- Path: path or trajectory of motion which a figure moves
- Motion: the presence of motion of the figure object

| | | | |
|---|---|---|---|
| (1) A boy | came | into | the classroom. |
| [Figure] | [Motion] | [Path] | [Ground] |

In addition, a motion event also contains two co-events or secondary components and namely they are manner of motion (i.e. the subsidiary action of a figure which includes rate, posture, and motor patterns) and cause of motion (i.e. qualitatively different agent causing the figure to move), as illustrated in Example (2) and (3).

| | | | |
|---|---|---|---|
| (2) A boy | walked | into | the classroom. |
| | [Motion + Manner] | | |

| | | | | |
|---|---|---|---|---|
| (3) A boy | rolled | a ball | into | the classroom. |
| | [Motion + Cause] | | | |

Given the great variety of semantic elements in motion encoding, languages differ in a number of ways in how motion events are typically expressed, both semantically and syntactically. From a semantic point of view, manner and path of motion have different specifications and dimensions. For example, manner of motion is a cover term which includes different sub-dimensions, such as motor pattern (run, hop, jump), rate of action (walk, run), attitudes of figures (limping), and manners with instruments (cycle, skating). Similarly, path of motion can be further divided into several sub-types in terms of the event trajectory: trajectory path such as up and down, and boundary-crossing path such as into and out of, or in terms of explicit semantic distinctions, such as vertical path (up/down), and horizontal path (towards/away from).

Due to a wide range of semantic components in motion event descriptions, it is difficult for speakers to encode all elements within a single utterance. Thus, speakers may prefer to select and encode certain elements, while omitting the others based on the perceived salience of each element. Then, speakers need to choose proper linguistic devices to encode these selected elements (Talmy, 2000). This process, termed lexicalization, refers to the process 'when a particular meaning component is found to be in regular association with a particular morpheme' (Talmy, 1985, p.59). Under this view, when speakers lexicalize a specific motion event, they need not only to decide which element to encode, but also where to encode it. For example, Talmy (2000) proposes that path, rather than manner, is the core element in a motion event as removing path will have greater semantic implications than removing manner of motion (also known as the Path Scheme). In fact, world languages provide various linguistic means to encode path of motion, but this is not the case for manner (Slobin, 1996b). This path-manner asymmetry suggests that in lexicalization, speakers of different languages may differ from each other in the frequency of manner encoding, but not in the encoding of path. This can be attributed to the reason that path is pragmatically obligatory in denoting a motion event.

When identifying the most characteristic expressions in motion events across various languages, there are three basic criteria: '1) colloquial in style, rather than literary; 2) frequent in occurrence, rather than occasional and 3) pervasive, rather than limited- that is, a wide range of semantic notions are expressed in this type' (Talmy, 2000, p.27). Based on the variabilities in lexical distributions and syntactical structures of motion, Talmy (1985, 1992, 2000) divided world languages into two distinct categories: satellite-framed languages (S-languages) and verb-framed languages (V-languages), depending on the semantic distribution of path. In the satellite-framed languages, manner is expressed via the main verb, while path is expressed outside of the verb in a satellite (i.e. into). The S-languages include a wide range of different languages families such as Germanic languages (German, English, and Dutch), Slavic languages (Polish and Russian), as well as Finno-Ugric languages. The most typical construction in S-languages is Manner verb + Path satellite, as shown in example (4)

Examples in English: Manner verb+ Path satellite

(4) A boy ran [Motion+Manner] across [Path] the street.

On the contrary, the verb-framed languages typically encode path in the main verb, leaving manner not expressed (by default) or via peripheral devices (positional phrases or gerundive constituents). The V-languages contain Semitic and Romance languages, Turkic languages, as well as Japanese and Korean. Examples for V-languages are illustrated in (5) and (6).

(5) Example in French: Path verb + Manner adjunct
Il traverse [Path] la rue.

'He crossed the street.'

(6) Example in French: Path verb only
Il traverse [Path] la rue en courant [Motion + manner]

'He crosses the street running'.

The S- and V-language dichotomy has been fruitful in analysing many Indo-European languages. However, it fails to address other language-specific properties. First, this bipartite classification fails to accommodate some newly-investigated languages. For example, Brown (2004) reported that the Mayan languages, such as Tzeltal, motion is often encoded in parallel patterns (i.e. the parallel encoding of path in either verb or satellite). This parallel conflation makes it difficult to apply the binary typology. Second, some Indo-European languages, such as Greek and French, exhibit great intra-type cross-linguistic differences with properties of both S- and V-languages (Hickmann, Hendriks, Harr, & Bonnet, 2018; Pavlenko & Volynsky, 2015b). For example, French allows speakers to express path of motion via different linguistic forms: both the main verb and the satellite. This double framing strategy makes it difficult to classify French as a pure verb-framed language. In fact, as Berman and Slobin pointed out, 'these topological characteristics often reflect tendencies rather than absolute differences between languages' (Berman & Slobin, 1994, p.118). In addition, one genuine challenge for Tamly's typology is that this rule is not applicable to serial-verb languages, such as Chinese, Tai and other Sino-Tibetan languages where path and manner are encoded in compound forms of equal grammatical status based on the 'one-verb-per-clause' constraint (Beavers, Levin, & Shiao, 2010). Thus, this binary typology fails to accommodate one typical type of language where 'a single clause representing the same general event can contain two or more verbs with shared nominal arguments and further these verbs are usually of equal status; neither is a modifier of the other' (Zlatev & David, 2004, p.4).

The path-manner asymmetry as well as intra-typological variances within languages have led researchers to try other ways of motion event classification. For example, Slobin introduced a third type known as equipollent-framed languages, where 'both

Manner and Path are expressed by equipollent elements, that is, elements that are equal in formal linguistic terms, and appear to be equal in force or significance' (Slobin, 2004, p. 226). Different from the original typology proposed by Tamly (1985, 2000), Slobin proposes that cross-linguistic difference in motion expressions is better to be interpreted as a cline of manner salience, rather than an absolute difference (see Table 1 for a detailed illustration).

According to Slobin (2003,2006), an equipollent-framed language refers to a type of language that meets at least one of the following criteria: 1) path and manner receive equal amount of salience in terms of their grammatical or morphological status or 2) it incorporates typological features of both verb-framed and satellite-framed languages. For example, Mandarin Chinese is regarded as an equipollent-framed language (Chen & Guo, 2009; Ji, Hendriks, & Hickmann, 2011a, 2011b, 2011c; Slobin, 2004, 2006; Zlatev & Yangklang, 2004). In Chinese, the most typical way of motion description is to encode manner and path via a Resultative Verb Compound (RVC). A RVC consists of up to three components, with the first verb encoding manner or cause of motion, the second denoting path, and the third (optional) as deixis, as shown in Example (7). Most importantly, the second (V2) and third components (V3) in a RVC form can stand alone as full verbs, as shown in (8) and (9).

(7) Yi4 zhi1 mao1 cong2 shu4 shang4 pao3-xia4-lai2 (V1-V2-V3).

One CL cat from tree on run-descend-come

'A cat ran down from the tree towards us.'

| | | |
|---|---|---|
| pao3 | xia4 | lai2 |
| Manner/cause verb | path verb | deictic verb |
| Run | descend | come |

(8) Yi4 zhi1 mao1 xia5 le shu4.

One CL cat descend (V2) ASP tree

'One cat went down the tree.'

(9) Lai2 le yi4 zhi1 mao1.

Come (V3) ASP one CL cat

'Here comes a cat.'

Table 1. Tripartite typology of motion event construction adapted from Slobin (2006)

| **Language type** | **Semantic distributions** | **Typical construction types** | **Examples** |
|---|---|---|---|
| *Verb-framed* | Path in the main verb, with manner subordination | Path verb+ Manner subordination | Romance, Semitic, Turkic, Basque, Japanese, Korean |
| *Satellite-framed* | Path in non-verb element, manner in the verb | Manner verb + Path satellite | Germanic, Slavic, Finno-Ugric |
| *Equipollently-framed* | Path & Manner expressed by equivalent grammatical forms | serial verb: Manner verb+ Path verb<br>bipartite verb: [manner +path] $_{verb}$<br>generic verb: coverb $_{manner}$ + coverb $_{path}$ + verb $_{generic}$ | Niger-Congo, Sino-Tibetan, Tai-Kadai, Algonquian, Jaminjungan |

## 4.2. Slobin's tripartite typology of motion event encoding and the hypothesis of Manner salience

Given that path is an obligatory element in motion expressions (Tamly, 2000), Slobin (1996a, 1996b, 2003, 2006) further proposes that when classifying motion events, one should consider the following criteria regarding manner of motion: 1) the size and diversity of lexicons available for manner expressions, and 2) the overall frequency of manner selection within and across languages. These criteria serve as the basic tenet for the Manner Salience hypothesis.

With regard to the size of manner lexicons, the underlying psychological assumption is that learners with a richer set of lexicons are less likely to ignore the distinctions between different types of manner than those with a smaller set of lexicons. For example, English has around 23 words for the manner 'walk', while Turkish only has one single word. In addition, cross-linguistic comparisons via corpus analysis and translation studies further suggest that S-languages have a much larger vocabulary for manner expressions than V-languages (Slobin, 2003, 2005, 2006). Thus, a richer set of lexicons tends to make lexical innovations much more difficult and thus engenders more attention to this domain.

In addition, the degree of manner salience can be reflected by the overall frequency of manner selection. For example, speakers of V-languages encode path in the main verb, so manner of motion must be encoded somewhere else, either via an adverbial clause, or not expressed at all. In addition, as path of motion is regarded as the core element (Talmy, 2000), this manner-path asymmetry further makes the encoding of manner less regular and predictable. For instance, when describing the same event, speakers of V-languages can either say 'Tom entered the room' or 'Tom entered the room running', depending on whether the speaker wants to convey this information to the listener or not. Thus, as languages differ in how manner of motion is lexicalized, Slobin only proposes Manner salience, but not Path salience.

In addition, the omission of manner is more frequent in V-languages, especially when the target event denotes a notion of boundary-crossing (e.g. A boy walks across a road) (Aske, 1989; Slobin & Hoiting, 1994). However, this boundary-crossing constraint does not apply to other types of languages (i.e. S- and E-languages). In sum, cross-linguistic difference in motion event encoding is as a matter of manner salience, that is, satellite-framed languages have a higher degree of manner salience, while verb-framed languages have a lower degree of manner salience.

### 4.3. Cognitive consequences: the effect of typological differences in linguistic encoding on cognition

Regarding the interaction between language and cognition in the motion domain, a directly related question is: whether cross-linguistic differences in lexicalization have cognitive implications on how people conceptualize motion. Slobin's hypothesis of Manner Salience (Slobin, 2000, 2004, 2006) provides an excellent stating point. Given that world languages attach different degrees of salience to manner of motion, it is reasonable to hypothesize that if one language does not have to talk about feature X, then speakers will pay less attention to the feature X.

In line with the hypothesis of thinking-for-speaking (Slobin, 1996a) and cognitive semantics (Talmy, 2000; Langacker, 2008), the basic tenet of manner salience lies in that speakers' attention is automatically drawn to the element highlighted by grammar. In terms of processing, the frequent omission of manner in verb-framed languages tends to reduce its relevant cognitive salience in speakers' mental representations, at least during the process of thinking-for-speaking. In addition, encoding manner in subordinate forms (i.e. adverbial clauses or gerunds) requires more processing load, as this construction (i.e. non-finite verb form) is regarded as 'heavy' and 'redundant' in language comprehension and production (Slobin, 2000, 2004, 2006). Thus, the Manner Salience Hypothesis suggest that speakers of a lower frequency of manner encoding tend to pay less attention on manner in conceptualization.

A number of studies have confirmed that speakers of different languages organize and structure information differently based on the grammatical devices made available in their languages (Allen et al., 2007; Choi & Bowerman, 1991; Hickmann & Hendriks, 2006; Hohenstein et al., 2006; Slobin, 2003, 2004). And these cross-linguistic variations in linguistic encoding will have far-researching implications on event conceptualization. Empirical studies on motion event lexicalization and cognition will be reviewed in more detail in Chapter 5.

#### 4.3.1. Linguistic encoding of voluntary motion in English

Voluntary motion, or spontaneous motion, refers to an event where an object spontaneously initiates the movement without any particular causes, which eventually leads to its change of location. Examples (10) illustrates the most prototypical motion event construction in English. Being a typical S-language, English generally encodes manner in the verb root whereas path in a satellite (i.e. verb particles or prepositions). Although English has other ways to encode voluntary motion, such as conflating path in the main verb (i.e. ascend, descend, and exit), as illustrated in Example (11), these verbs are Latinate borrowings and represent a more formal register (Talmy, 2000). Thus, these expressions are seldom used in daily communication and not the most prototypical characteristics of English. A great number of empirical studies have confirmed the typological status of English as a typical S-language, with various linguistic and non-linguistic methods, such as linguistic elicitation, non-verbal categorization, co-verbal attention allocation and reaction time (Hendriks & Hickmann, 2015; Ji et al., 2011a; Ji & Hohenstein, 2017; 2018; Pavlenko & Volynsky, 2015a; Slobin, 2006; Talmy, 2000; 2009, to name a few). These methods will be reviewed in more detail in Chapter 5.

(10) **Manner verb+ Path satellite**

The boy walked [Manner] across [Path] the street.

(11) **Path verb +Manner subordination (optional)**

The boy <u>descended</u> [Path] the mountain walking.

### 4.3.2. Linguistic encoding of voluntary motion in Japanese

On the contrary, Japanese is a V-language (Slobin, 2004, 2006; Talmy, 1985, 2000), or head path coding language (Matsumoto, 2018). This means Japanese typically conflates path of motion in the verb root, leaving manner unexpressed (by default), as shown in example (12a), or via a subordinate or gerund form (the *-te* conjunctive marker), as illustrated in (12b).

(12) **a. Path verb, without mentioning Manner** (Yui, 2013).

| | | | |
|---|---|---|---|
| Kara | wa | sangai-ni | agatta [path verb] |
| S/he | TOP | third floor-DAT | ascend. PST |

'S/he ascended to the third floor.'

**b. Path verb + Manner in subordination**

| | | | | |
|---|---|---|---|---|
| Kara | wa | arui-te [manner gerund] | heya-ni | hait-ta [path verb] |
| S/he | TOP | walking | room-DAT | enter. PST |

'S/he walked into the room.'

Regarding the expression of manner of motion, Japanese encodes manner via various other means such as: (a) the complex verb predicate, (b) verb compound (V1-V2), and (c) mimetics (Brown & Gullberg, 2010; Matsumoto, 1996, 2018; Yoshinari, 2016). The complex predicate, also called as ‘conjunctive compound construction’ (Shibatani, 2003), refers to the $V_1$-te $V_2$ complex. In this construction, manner of motion is expressed in the $V_1$-te form as a co-verb, while path of motion $V_2$ serves as the verb root. For instance, the complex predicate *hasit-te noboru* [running-ascend] ‘run up’ consists of two parts: a manner gerund (non-head) *hasit-te* ‘running’, and a path verb (head) *noboru* ‘ascend’ (Matsumoto, 2018).

(13) Kara-wa heya-ni **hasit-te [manner gerund]** **hait-ta [path verb]**
S/he-TOP room-DAT run-CONJ enter-PST
‘S/he ran into the room.’

Alternatively, manner can also be conflated in a verb compound. The verb compound (V1-V2) in Japanese is composed by combining two verbal elements at the word level, with the first element (V1) as a satellite and second element (V2) served as the main component (Matsumoto, 1996). For example, in *kake-agar-u* [run-ascend-NPST] ‘run up’ and *suberi-de-ru* [slip-exit-NPST] ‘slip out’, the first elements (V1) in these two compounds *kake-* ‘running’ and *suberi-* ‘sliping’ denote the manner of motion, whereas the second elements (V2) *–agar-u* ‘ascend’ and *–de-ru* ‘exit’ represent the direction of motion. However, the verb combinations mentioned above are not productive in use and there are several restrictions on the compounding process. For example, ‘run up’ [*kake-agar-u*] serves as a fixed term in motion event descriptions, but not for ‘walk up’ [aruki-agar-u]. In addition, these compound verbs only combine manner with non-deictic verbs, but not with the deictic ones (Lamarre, 2008; Matsumoto, 1996, 2018).

(14) Kara-wa kaidan-o **kake-agat-ta**
S/he-TOP stairs-ACC run-ascend-PST
‘S/he ran up the stairs.’

In addition, manner in Japanese can also be encoded by mimetics, or sound-symbolic forms, which are used to express detailed and vivid manner information (Hamano, 1998), as illustrated in example (15) and (16). In example (15), *batan*- 'thud' is used to express a sudden falling off of a particular object, while *sutasuta*- in example (16) denotes the manner of walking and provides a vivid description of how the figure moves out of the scene. However, the size of the mimetics is very small.

(15) Hon-ga **batan-to** taore-ta (Toratani, 2016)
Book-TOP MIMETIC-P fall-PST
'The book fell with a thud.'

(16) Mari-wa **sutasuta-to** kaet-te-it-ta
Mari-TOP MIMETIC-P return-L-go-PST
'Mari strode off to her home.'

In sum, manner of motion in Japanese is most typically encoded as a non-head or external-head position in subordinate forms such as in (4b), non-head position of verb complexes as in (15) and co-verb form in (16). However, it can never serve as the head position of the verb complexes, as shown in example (17).

(17) *Kara-wa heya-ni **hait-ta [path gerund]** **hasit-ta [Manner verb]**
S/he-TOP room-DAT enter-CONJ run-PST
'S/he ran into the room.'

Although manner of motion in Japanese can be conflated in the simple main verb, the occurrence of manner verbs with prepositional or postpositional (PP) phrases are limited and can only be used under certain conditions (Inagaki, 2001, 2002; Matsumoto, 1996; Talmy, 1985). For example, manner verbs in Japanese are not allowed to appear together with goal PPs, as shown in example (18) and (19).

(18) *Kara-wa heya-ni **aruita** **[manner verb]**

S/he-NOM room-to walk-PST

'S/he walked to room.'

(19) *Kara-ga ie-no naka-ni **hassita** **[manner verb]** (Inagaki, 2001)

S/he-NOM house-of inside-at ran

'S/he ran into the house.'

In addition, as a V-language, the boundary-crossing constraint also applies to Japanese that the use of manner verbs is not licensed when the event denotes a categorical change of location (Aske, 1989; Slobin & Hoiting, 1994). That is, Japanese only licenses the use of path-to-verb mapping in motion expression in the boundary-crossing situation, as illustrated in example (20). However, this constraint is not applicable to English (S-languages) or Cantonese (E-languages).

(20) Kara-wa **arui-te** **[manner gerund]** heya-ni **hait-ta** **[path verb]**

S/he-NOM walking room-DAT enter. PST

'S/he walked into the room.'

Given the typological features of manner encoding, empirical evidence suggests that Japanese (V-language) encodes manner less frequently than English (S-language). For example, Matsumoto (2017) examined the Balanced Corpus of Contemporary Japanese (BCCWJ) and found that manner of motion was only mentioned 202 of 882 descriptions (23%), far much less than the overall frequency of manner encoding in the COBUILD corpus in English. Similar results were found by the experimental data that Japanese had a lower frequency of manner encoding, especially for the default manner of walking: only 35 cases of manner were mentioned out of 882 examples (Matsumoto, Akita, & Takahashi, 2017). The relatively lower frequency of manner encoding in Japanese can be attributed to 1) the lack of obligatory syntactic and semantic slot for manner 2) the requirement of additional morphology to express manner information and 3) the boundary-crossing constraint that restricts the use of certain manner verbs when there is a categorical change in the location.

With regard to path, corpus and experimental data confirms that Path is prototypically encoded in the main verb, such as agaru 'ascend', oriru 'descend' and deru 'exit'. However, it can also appear in other positions such as pre-and post-positional particles (i.e. kara 'from' and made 'to'), and deictics (i.e. iku 'go' and kuru, 'come'). Corpus data of Japanese (BCCWJ) shows that more than 50% of voluntary motion expressions encode path in simple path verbs and around 40% of the descriptions use deictics as the main verb of a sentence. However, manner verbs are used less than 10% in all motion descriptions (Matsumoto et al., 2017). Regarding different types of path (i.e. up, down and into), on the one hand, directional path such as UP/DOWN, and path with a boundary-crossing such as ACROSS/INTO are predominantly encoded in the main verb (Matsumoto, 2018), as shown in (21) and (22).

(21) Kara-wa kaidan-o **agat-ta** **[path verb]**
S/he-TOP stairs-ACC ascend-PST
'S/he ascended the stairs.'

(22) Kara-wa miti-o **watat-ta** **[path verb]**
S/he-TOP road-ACC cross-PST
'S/he crossed the road.'

On the other hand, path with directional goals, such as TO, are often encoded via source/goal prepositions as a case marker, whereas TO+IN is most often double encoded via both the main verb and prepositional phrases, as shown in (23) and (24).

(23) Kara-wa eki-**ni/kara** **[path satellite]** **hasit-te** **[manner gerund]** **it-ta** **[deictic verb]**

S/he-TOP station-GOAL run-CONJ go-PST

'S/he ran to/from the station.'

(24) Kara-wa heya-**ni** **[path satellite]** **hait-ta** **[path verb]**

S/he-TOP room-GOAL enter-PST

'S/he entered the room.'

#### 4.3.3. Linguistic encoding of voluntary motion in Cantonese

Cantonese, widely spoken in Hong Kong and Guangdong Province in China, is one regional variety of the Chinese languages (Yiu, 2013, 2014). Compared with Mandarin Chinese, little work has been done to systematically examine the typological status of Cantonese in the construction of motion events, especially with regard to the semantic and syntactic properties of path of motion.

According to the limited amount of available research, some scholars propose that like Mandarin Chinese, Cantonese is a serial-verb language (Matthews & Yip, 2011). And a serial-verb construction in Cantonese consists of two or more components. Each component is able to stand alone as an independent element (Matthews, 2006). Based on this rationale, Cantonese, like Mandarin Chinese, is an equipollent-framed language where manner and path of motion are equipped with equal grammatical significance or status. However, other scholars argue that elements within a serial verb construction may not have the equal grammatical status (Yiu, 2013, 2014; Yuan, 2011). For example, in the description of a specific scene of voluntary motion 'walking into the classroom', haang4-**zo2** jap6-haak2 si1 (**walk-ASP-into**-classroom), the aspect marker **zo2** 'Past', in Cantonese should be attached to the first verb **haang4** 'walk' but not the second one **jap6** 'enter'. However, in Mandarin Chinese, the aspect markers are usually attached to the verb compound (the RVC form), such as zou3jin4-

**le** jiao4shi, (**walk-enter-ASP**-classroom). Thus, it is difficult to prove that each component in a serial verb construction actually has an equal grammatical status as the aspect marker prototypically follows the element with higher level of grammatical salience. In fact, Cantonese does not totally pattern with Mandarin Chinese in that the integration of manner and path tighter in the latter than the former due to the wide application of the RVC form (Chen, 2005; Chen & Guo, 2009; Ji, 2009; Ke, 2003).

The other way of classifying Cantonese in motion typology is to regard the serial verb construction as a combination of satellite- and verb-framed properties (Yiu, 2013, 2014). This approach divides path in Cantonese into two sub-categories, directional verbs (path verbs) that can stand along as independent elements, and directional complements (path satellites) can must appear together with a main verb and denote its specific direction. In total, Cantonese has twelve single directional verbs: seong4 上 'to ascend', lok6 落 'to descend', ceot1 出 'to exit', jap6 入 'to enter', hoi1 開 'to depart', maai4 埋 'to approach', gwo3 過 'to pass', hei2 起 'to rise', dou3 到 'to arrive', faan1 翻 'to return', 嚟 lai4, 'to come' and heoi3 去 'to go'. And these single path verbs can combine with each other to form various other compound directional verbs.

For example, the path of motion, faan1 and jap6, can be expressed either as verb complements in example (25), or as independent verbs in example (26) and (27). Both constructions are frequently and pervasively used in the oral production (Yiu, 2013). Following this way, Cantonese is an equipollent-framed language, standing midway on the continuum of S- and V-languages (Lamarre, 2007; Yiu, 2013, 2014).

(25) Keoi5 paau2 zo2 faan1 jap6 seoi6 fong2.
S/he run ASP return enter bedroom
'S/he run back into the bedroom.'

(26) Keoi5 faan1 zo2 seoi6 fong2

S/he return ASP bedroom

'S/he returned to the bedroom.'

(27) Keoi5 jap6 zo2 seoi6 fong2

S/he enter ASP bedroom

'S/he entered the bedroom.'

From this perspective, the typological status of Cantonese is similar to Mandarin Chinese (Ji et al., 2011a, 2011b, 2011c) as these two languages have equipollent framing systems with satellite- and verb-framed properties. In addition, Talmy (2009) has suggested that Chinese is the only language that fits the case of equipollence. This can be attributed to diachronic transformations that Classical Chinese went through from a V-language to an S-language (Peyraube, 2006), and such typological transformations in Cantonese are not completed yet (Xu, 2006; Yiu, 2013). Thus, it has been argued that typological distinctions between S- and V-languages should not be viewed as an absolute dichotomy, but as a continuum with various degrees of manner and path salience (Slobin, 2004; Zlatev & Yangklang, 2004).

In fact, regarding Cantonese as an equipollently-framed language located along the satellite-framed and verb-framed continuum also supports Slobin's tripartite typology (Slobin, 2004, 2006) that an equipollent-framed language can be defined in two ways: it either refers to languages with multiple components encoded in equal grammatical forms or languages with various degrees of satellite- or verb-framed properties. In order to make the linguistic comparisons consistent across three target languages, the current study adopted the second way of categorizing Cantonese as an E-language, that is, Cantonese incorporates linguistic properties of both S- and V-languages. Thus, the second component in a Cantonese serial verb construction is regarded as a verb complement, an equivalent to the English satellite, rather than the directional verb (path verb).

### 4.4. Cross-linguistic differences in caused motion encoding: a comparison of English, Japanese and Cantonese

Caused motion, or self-agentive motion, refers to the situation where an agent exerts some external forces on the object which causes its direct movement (Talmy, 1985, 2000). It is a complex type of motion that allows a number of semantic elements to be encoded simultaneously in linguistic descriptions. The basic components of caused motion include path of motion (into, out of), causation or causal displacement (take, carry), and manner of motion. More specifically, there are several types of manners in caused motion: manner of the agent (i.e. walk), manner of cause (pull, drag) and manner of object (slide, roll), as illustrated in Example (28). In line with Talmy's typological distinctions of voluntary motion, caused motion fall into two broad categories based on the semantic distribution of path.

(28) Keio4 teui1zyu6 go3 muk6syun4 haang4dou2 go3 chou2 dei6
S/he push-GER a wood ship walk to a grassland
'S/he push a wood ship to a grassland.' (CAN9vol)

For example, for S-languages such as German and English, path of motion (e.g. into) is encoded outside of the verb root in a satellite whereas cause of motion, like manner, is an external co-event which can be conflated with motion in the main verb. In contrast, for V-languages such as Japanese and French, cause of motion is typically conflated with path in the main verb, leaving manner unexpressed (by default) or in peripheral devices (i.e. subordinations, gerunds, adverbial expressions).

#### 4.4.1 Linguistic encoding of caused motion in English

Caused motion in English is typically encoded via lexical causative verbs, that is, verbs that denote causation and specific manner of cause, such as push, pull and kick. Syntactically, encoding manner and path of motion into a compact form (i.e. cause +manner in the verb and path in satellite) is the most prototypical lexicalization pattern as it provides a more readily accessible construction, as shown in (29). Although it is grammatically correct for English speakers to conflate path in the main verb whereas manner/cause in peripheral phrases, as illustrated in (30), this construction is less conventional in communication and rarely used in oral descriptions (Talmy, 1985, 2000). This is because the use of causation verbs and compact syntactical structures allow speakers of English to express caused motion in a more economical way without adding additional subordinate constructions.

(29) A boy pushed [Manner + Cause] a box into [Path] a cave.

(30) A boy crossed [Path] the street pushing [Manner + Cause] a box.

### 4.4.2. Linguistic encoding of caused motion in Japanese

With regard to the lexicalization of caused motion in Japanese, research to date documents that as a typical V-language, Japanese most often conflates cause with path in a set of transitive path verbs (i.e. ageru 'make ascend', orosu 'make descend', dasu 'make exit' etc.), leaving manner of cause unexpressed (by default), or via subordinate forms (i.e. -te conjunctive marker), as illustrated in (31a) and (31b).

In example (31a), the transitive verb, or causative path verb, *ageru* 'make ascend' is used to indicate the agent's action on an object which causes its direct movement in an upward direction. In this expression, the manner of cause is not specified. However, example (31b) demonstrates that manner of cause in Japanese can be encoded as a co-event verb via a subordination form, when the interlocutor plans to stress this piece of information.

(31) a. Kara-TOP nimochi-o **ageta**[Path+Cause]
S/he TOP goods-ACC ascend.PST
'S/he moved up the goods.'

b. Kara-TOP nimochi o **oshi-te** [Manner of Cause] michi o watali-mashita [Path]
S/he-TOP goods-ACC pushing-CONJ road-ACC crossed. PST
'S/he crossed the road pushing the goods.'

In addition, although Japanese has a specific set of manner verbs for causation (i.e. oku 'put', toru 'take', 'hakobu' carry), the number of these verbs is very limited, not only when compared with head-external languages (S-languages) such as German and English, but also other head-internal languages (V-languages) like French and Italian (Levin, 1993; Verkerk, 2013). Furthermore, like voluntary motion, the boundary-crossing constraint also applies to caused motion in that cause or manner should be encoded via peripheral devices rather than main verbs when there is a categorical change of location (Aske, 1989; Slobin & Hoiting, 1994; Hickmann, 2003).

### 4.4.3. Linguistic encoding of caused motion in Cantonese

Like the linguistic encoding of voluntary motion, the expression of caused motion in Cantonese integrates the typological features of both S- and V-languages. One of the most prototypical expressions in Cantonese is the disposal construction. Like the BA construction in Mandarin Chinese, the disposal construction in Cantonese is marked by the disposal marker zoeng1 and followed by transitive verbs. The zoeng1 marker, also called 'differential object marking', is used to signal how the object is affected by the causing action and denotes 'how the object is disposed of.'

As shown in example (32) the disposal construction consists of a subject (Keio4 'S/he'), followed by the disposal marker zoeng1, a subsequent post-zoeng1 object noun phrase (toi2 'table'), and ends with a cause verb (bun1 'move'), an aspect marker (zo2 PST), as well as a directional complement (faan1 'back'). In addition, it can be further followed by a spatial noun phrase denoting the ground of motion. The zoeng1 construction conflates cause/manner of cause in the verb root whereas path in the satellite, very much like the constructions of motion in English. From this perspective, the zoeng1 construction in Cantonese reflects typological properties of typical S-languages in terms of the syntactic complexity and framing strategy.

(32) Keio4 zoeng1 zoeng1toi2 **bun1** zo2 **faan1** uk1kei5
S/he DM a table move-ASP return home
'S/he moved a table back home.'

Meanwhile, like Japanese, Cantonese also allows the conflation of cause with path in the main verb, without expressing manner of cause at all, or via subordinate forms, as demonstrated in (33) and (34). The directional verb soeng5 (make ascend) in (33) indicates that the agent causes the object to move in an upward direction without specifying the manner of causation, very similar to the usage of path verbs in Japanese, as illustrated in example (30).

(33) Keio4 **soeng5** zo2 saam1seong1fo1 hai2go3gaa3 (dou6)
S/he ascend-ASP three goods at the CL shelf (Localizer)
'S/he moved three boxes of goods up onto the shelf.'

(34) Keio4 **teui1 zyu6** go3 muk6leun4 heui3zo2 ho4bin1
S/he **push-DUR** a wheel went riverside
'S/he went to the riverside pushing a wheel.'

Given their typological status of English (S-language), Cantonese (E-language) and Japanese (V-language), cross-linguistic differences in manner and path salience raise the question of whether language-specific ways of talking about motion affect how motion is presented at a deeper level of cognition. According to Slobin's Manner Salience Hypothesis (2003, 2004, 2006), the codability of manner in lexicalization increases its accessibility in cognitive processing. That is, speakers' memory and attention are guided by cross-linguistic variations in the lexical and grammatical patterns, such that speakers tend to pay more attention to the linguistic element that is made more available and salient by that language.

Under this view, speakers of high-manner-salience language (i.e. S-languages) tend to attach more importance to manner of motion in cognition due to its high codability and prominent grammatical status (i.e. in the verb root). In contrast, as manner is not obligatory and often encoded via subordination (for boundary-crossing events in particular) in low-manner-salient-languages (i.e. V-languages), the frequent omission of manner in lexicalization reduces its cognitive salience in mental representations for speakers of V-languages. Thus, the current study hypothesizes that English monolinguals attach more salience to manner of motion compared with Cantonese monolinguals, whereas Japanese speakers present the lowest degree of manner salience in both the verbal measurement of event lexicalization and subsequent non-verbal event categorization.

# Chapter 5. Empirical evidence on motion event lexicalization and conceptualization in monolingual and bilingual speakers

As reviewed in Chapter 4, world languages differ typologically in motion event encoding. These varying degrees of manner and path salience in motion lexicalization have brought about the enquiry of whether language-specific ways of talking about motion affect how motion is presented at a deeper level of cognition.

This chapter will combine current theories on language and cognition (i.e. thinking-for-speaking and thinking-with-language hypothesis) with the latest empirical findings on motion event lexicalization and conceptualization. It starts with a systematic review of the development of lexicalization patterns in monolingual speakers, bilingual speakers and various types of L2 learners. Then it will address the questions of whether, and to what extent, bilinguals and L2 speakers are able to recalibrate their cognitive dispositions towards the conceptualization patterns of the target language as a result of L2 learning. At the end of this chapter, research gaps will be identified, and the rationale will be justified together with specific research questions.

## 5.1. Empirical Evidence on motion event lexicalization in monolingual and bilingual speakers

### 5.1.1. Cross-linguistic differences in motion event lexicalization: monolingual speakers

Previous studies show that the effect of thinking-for-speaking is language-specific with a wide range of elicitation methods, such as spontaneous speech (Choi, 2011a; Choi & Bowerman, 1991; Pourcel, 2009), storytelling (Slobin, 2000, 2006; Slobin & Hoiting, 1994) and descriptions of dynamic stimuli (Hendriks et al., 2008; Hickmann & Hendriks, 2010; Ji et al., 2011b; Ji & Hohenstein, 2014). Empirical evidence shows that in the lexicalization of motion, both children and adult speakers of S-languages tend to encode manner of motion more frequently and are more prone to interpret novel verbs as manner than V-language speakers (Brown & Gullberg, 2008; Hickmann & Hendriks, 2006; Hickmann, Taranne, & Bonnet, 2009; Hohenstein, 2005; Montero-Melis & Bylund, 2017). For example, in an experiment with dynamic stimuli of voluntary motion, Hickmann, Taranne, et al. (2009) reported that English-speaking participants (both adults and children) typically encoded manner in the verb root whereas path in the satellite. In contrast, French participants preferred to encode path in the main verb, leaving manner unexpressed, or via subordinate forms. Similar results were reported by Montero-Melis and Bylund (2017) in linguistic encoding of caused motion. Results showed that Swedish monolinguals (S-language) encoded manner of cause more often than Spanish monolinguals (V-language) in the description of dynamic video stimuli. However, these two groups did not differ from each other in path selection. Brown and Gullberg (2008) also reported that speakers of English and Japanese differed in when and where manner was encoded, that is, English speakers encoded manner more frequently than Japanese speakers via subordinate forms. These findings suggest that language specificities affect how spatial concepts are linguistically encoded at the formulation level in speech production (Soroli et al., 2019).

In fact, the acquisition of language-specific lexicalization patterns starts from a very early age. Although children normally begin with a default starting point with equal amount of attention being paid to manner and path of motion, the language-specific lexicalization patterns start to emerge from an average of three years old (Allen et al., 2007; Choi & Bowerman, 1991; Hickmann & Hendriks, 2006; Hickmann & Hendriks, 2010; Ji et al., 2011a; Maguire et al., 2010). These language-specific patterns are manifested in three dimensions: patterns of lexicalization (Choi, 2011a, 2017; Choi &

Bowerman, 1991; Hickmann & Hendriks, 2010), syntactic constructions (Allen et al., 2007), and discourse organization (Talmy, 2000; Slobin, 2004, 2006). For example, Choi and Bowerman (1991) reported in a longitudinal study that bilingual children showed sensitivity to typical patterns of two languages from 17-20 months old. Similar findings were found in Choi (2017) that from early on, English children (S-language) showed greater density in motion event descriptions compared with Korean and French children (V-language). As early as 17 months, Korean children began to encode path in the main verb whereas manner in adverbial forms. Similar lexicalization patterns were also reported by Hickmann, Hendriks, and Champaud (2009) in a corpus study with English- and French-speaking children. Results showed that the language-specific lexicalization patterns started to emerge as early as two years old. English children typically adopted the ‘Manner verb +Path satellite’ construction, while French children preferred to conflate path in the main verb whereas expressed manner less frequently. These findings are closely related to the current study as bilinguals with early L2 exposure may become sensitive to particular ways of linguistic encoding in both languages.

Although cross-linguistic research on motion event lexicalization mostly focuses on voluntary motion, studies exploring caused motion events remain limited (Choi & Bowerman, 1991; Hendriks et al., 2008; Ji et al., 2011b; Montero-Melis & Bylund, 2017). For example, among the limited number of available studies, language-specific properties in the lexicalization of caused motion are found among speakers of Korean and English (Choi & Bowerman, 1991), Spanish and Swedish (Montero-Melis & Bylund, 2017), and English and Chinese (Ji et al., 2011b). For example, Montero-Melis and Bylund (2017) reported that speakers of Spanish and Swedish differed greatly in how often they chose to encode manner of cause (C-manner) when verbalizing dynamic stimuli of caused motion. As a satellite-framed language (S-language), speakers of Swedish encoded C-manner with a high frequency whereas speakers of Spanish (V-language) tended to omit this information. These findings echoed with Ji et al. (2011b) that speakers of Chinese and English differed from each other not only in how often C-manner was encoded, but also where this information was encoded and what structures were used to encode this information (i.e. the semantic distribution and syntactic structures of C-manner). The mixed results suggest that like voluntary motion, typological contrasts are also observed in this more complex type of motion in the cross-linguistic research.

### 5.1.2. Cross-linguistic differences in motion event lexicalization: bilingual speakers

Moving to bilingual speakers, the lexicalization of motion events has been a central topic in bilingual and L2 acquisition research. Most studies along this line focus on whether L2 learners are able to restructure their L1-based lexicalization patterns and develop L2-based patterns of thinking-for-speaking. Some studies have demonstrated that bilinguals or L2 learners with typologically different languages may transfer certain L1-based lexicalization patterns into the L2 (Brown & Gullberg, 2008; Cadierno, 2008, 2010; Cadierno & Ruiz, 2006; Daller et al., 2011; Hendriks & Hickmann, 2015; Larrañaga, Treffers-Daller, Tidball, & Ortega, 2012). For example, speakers of V-languages (Japanese, French) tend to encode manner of motion less frequently in their L2 S-languages (English, Danish) compared with S-language monolinguals (Brown & Gullberg, 2008; Cadierno, 2010), whereas S-language speakers learning a V-language as an L2 may have difficulty in acquiring target lexicalization patterns of manner encoding (Cadierno & Ruiz, 2006; Hendriks & Hickmann, 2015; Hendriks et al., 2008). For example, using a storytelling task, Daller et al. (2011) reported evidence for L1-based transfer in German-French bilinguals when describing voluntary motion with a boundary-crossing situation. Results showed that violations of boundary-crossing constraints were found even in L2 learners with advanced proficiency. Similar results were reported by Inagaki (2001) that English learners of L2-Japanese accepted L1-based patterns of placing manner verbs prior to the directional PPs (e.g. walk to school), which were rejected by Japanese native speakers. The results provide evidence that it is difficult to restructure the L1-based patterns of 'thinking for speaking' towards the target language during L2 learning.

However, other studies report that bilinguals and L2 learners are able to restructure their L1-based patterns of 'thinking for speaking' to the target linguistic forms when talking about motion events in an L2 (Hendriks & Hickmann, 2011; Ji & Hohenstein, 2014), and the influence between L1- and L2-based concepts is bidirectional (Brown & Gullberg, 2011; Daller et al., 2011; Hohenstein et al., 2006). For example, Hendriks and Hickmann (2011) showed that intermediate and advanced English learners of French were able to acquire the L2-based patterns of conflating path in the main verb when describing voluntary motion. Hohenstein et al. (2006) reported that Spanish-English bilinguals' lexical choice of manner verbs fell in-between monolingual baselines, suggesting a convergence of L1- and L2-based concepts in the bilingual mind. In addition, several studies also document bidirectional transfer in Japanese-English (Brown & Gullberg, 2008, 2011), German-French (Berthele & Stocker, 2017), Spanish-Swedish (Bylund 2011), Hungarian-English (Vanek & Hendriks, 2015) and Turkish-German bilinguals (Daller et al., 2011; Treffers-Daller & Calude, 2015). These findings suggest that bilinguals or L2 learners with higher levels of language proficiency or consistent exposure to the target language tend to show a convergence of both L1- and L2-based patterns of lexicalization.

While cognitive restructuring in motion event encoding has been actively examined with diverse L2 populations, only few studies report evidence for the influence of L2-based patterns of lexicalization on the native language (Aveledo & Athanasopoulos, 2016; Wang & Li, 2019). Among the few studies, Aveledo and Athanasopoulos (2016) examined the effects of L2 learning on L1 event encoding with early Spanish-English bilinguals. Results showed that bilingual speakers opted for more manner verbs but fewer path verbs in their L1 descriptions compared with Spanish monolinguals, suggesting a backward influence from the L2 on the L1. Similar findings were reported by Wang and Li (2019) that early Cantonese-English bilinguals encoded manner of motion more frequently compared with Cantonese monolinguals when describing voluntary motion in their native language. And the degree of L2-based cognitive restructuring in motion event lexicalization was modulated by the amount of contact with each language.

The overall results suggest that cognitive restructuring is a dynamic process and susceptible to various linguistic and extra-linguistic factors. The linguistic factors are concerned with the target languages under investigation (with or without contrastive typological features), the modalities of language production (oral or written), types of stimuli (static or dynamic) and targeted L2 populations (early or late bilinguals). The extra-linguistic factors are multiple, starting from language learning environment and language dominance (Park, 2019; Stocker & Berthele, 2019), age of acquisition (Engemann, Harr, & Hickmann, 2012; Hohenstein et al., 2006), to L2 proficiency (Cadierno & Ruiz, 2006; Ji & Hohenstein, 2014; Treffers-Daller & Calude, 2015), frequency of language use (Daller et al., 2011; Park, 2019), as well different language contexts (Berthele & Stocker, 2017; Stocker & Berthele, 2019). This makes the domain of motion event cognition in bilingual speakers and language learners a promising research area that needs further investigation.

### 5.2. Empirical Evidence on motion event conceptualization in monolingual and bilingual speakers

#### 5.2.1. Cross-linguistic differences in motion event conceptualization: monolingual speakers

Moving beyond language use, cross-linguistic studies in the linguistic encoding of motion events start to question whether language-specific patterns in lexicalization affect how motion is presented at a deeper level of cognition. Studies on event cognition have been documented in children and adults (Allen et al., 2007; Aveledo & Athanasopoulos, 2016; Ji & Hohenstein, 2017; 2018; Hickmann et al., 2018), with different combination of language pairs (Athanasopoulos, Bylund, et al., 2015; Daller et al., 2011; Hickmann & Hendriks, 2010; Montero-Melis et al., 2017; Soroli & Hickmann, 2010), and by a wide range of non-verbal measurements, such as similarity judgement, recognition memory, attention allocation, reaction time, and gestures of motion (A. Brown, 2015; Filipović, 2011; Flecken, Carroll, et al., 2015; Montero-Melis et al., 2016; Papafragou et al., 2008; Von Stutterheim et al., 2012).

Numerous studies have investigated whether different degrees of manner salience in lexicalization affect event cognition. Some studies have reported that speakers of S- and V-languages categorize motion events along the same parameters regardless of the typological differences (Loucks & Pederson, 2011; Papafragou et al., 2002). However, other studies have demonstrated a clear language effect on non-verbal behaviour, such as event categorization, recognition memory and attention allocation, when speakers' access to language is not blocked in the decision-making process

(Gennari et al., 2002; Papafragou et al., 2008; Soroli, 2012; Soroli & Hickmann, 2010; Trueswell & Papafragou, 2010). For example, Soroli and Hickmann (2010) investigated whether cross-linguistic differences in motion event lexicalization modulated English and French speakers' categorical preferences by using a triads-matching paradigm. The results suggested that in line with language-specific encoding patterns, French speakers showed a preference for path in event categorization under both a verbal and a non-verbal condition, while English speakers presented a significant manner preference only in a verbal condition where linguistic encoding was provided prior to categorization. Similarly, Montero-Melis and Bylund (2017) examined the effect of language on cognition with Swedish (S-language) and Spanish (V-language) native speakers in different conditions. The results suggested that Swedish monolinguals were more likely to use 'same-manner' criteria for similarity assessments as long as they could access the target language. These results are in line with the thinking-for-speaking hypothesis that language effects on conceptualization only appear when participants are engaged in language-driven activities, such as language production and comprehension. There is consistent evidence that speakers of different languages exhibit different cognitive patterns during or immediately after overt language use (Montero-Melis & Bylund, 2017; Papafragou et al., 2008; Soroli & Hickmann, 2010).

However, other studies demonstrate that such language effects disappear under verbal interference. For example, some studies do not report language effects on cognition, such as recognition memory and categorization (Papafragou et al., 2002), whereas other studies suggest that language effects only appear when the access to linguistic labelling is permitted during cognitive processing (Gennari et al., 2002; Montero-Melis & Bylund, 2017; Papafragou & Selimis, 2010). The mixed results suggest that the language-specific encoding has a selective impact on non-verbal cognition, which is modulated by various factors, such as the degree of language involvement (explicit, implicit, and blocked), different experimental set-ups, and various target languages or cognitive functions under investigation.

Regarding different ways of coupling linguistic production with non-linguistic data (i.e. categorization, visual attention, and processing efficiency etc.), most studies utilize a classic triads-matching paradigm to tap into the potential bias in speakers' categorization or similarity judgement preferences. The triads-matching paradigm has been extensively used in language and cognition research as a classical measurement of participants' non-verbal behaviour (Bylund & Athanasopoulos, 2014b). In this task, participants are usually provided with visual materials that are not directly related to language and need to match the target stimuli (X) with one of the alternate stimuli (AB) based on the degree of similarity. The basic tenets of this research paradigm lies in that 1) similarity is the basis of categorization (Nosofsky, 1986); and 2) categorization is an indispensable part of human cognition (Harnad, 1987). Thus, this method requires higher level cognitive processes, which are post-perceptual and occur at a later stage of cognitive processing (i.e. categorical judgements, classification, reasoning), and often used as an explicit measurement to tap into subjects' overt preferences and decision strategies (Bylund & Athanasopoulos, 2014b; Ji, 2017; 2018; Soroli, Hickmann, & Hendriks, 2019).

The triads-matching paradigm has been successfully implemented in different cognitive domains such as colour, objects and substances, as well as in the domain of motion events (Bylund & Athanasopoulos, 2014b; Ji & Hohenstein, 2018; Park, 2019), and in combination with different types of experiment manipulation (i.e. pre-verbalization, verbal interference, or shadowing). The results show that on the one hand, effects of language on conceptualization were found when language was either explicitly or implicitly involved at the very moment of decision-making (Gennari et al., 2002; Hickmann, Engemann, Soroli, Hendriks, & Vincent, 2017; Lai et al., 2014; Papafragou & Selimis, 2010). However, such effects disappeared when a verbal interference (Ji & Hohenstein, 2017; Trueswell & Papafragou, 2010) or a task distraction was used (Filipović & Geva, 2012). For example, Hickmann et al. (2017) investigated how English and French speakers categorized voluntary motion under three different conditions: a non-verbal condition, a verbal condition with prior linguistic production, and a dual task condition where verbal interference was used. The results provided evidence for the effects of language on cognition under both verbal and non-verbal conditions, that is, English group were more likely to base their judgements on manner-choices compared with French group. However, such language effects disappeared when the verbal interference was introduced.

Similar results are reported by other studies that probed into recognition memory: whether different degrees of manner salience affect the memorization and recall of relevant linguistic elements (Filipović, 2011, 2018; Papafragou, 2002). The results further support that the effects of language on thought are visible under conditions when linguistic labelling was permitted prior to or during categorization.

In addition, several studies have used other types of online measurement such as eye-tracking and brain imaging to capture participants' cognitive behaviour during online processing (Athanasopoulos et al., 2010; Flecken, Athanasopoulos, et al., 2015; Guillaume et al., 2009). For example, many up-to-date studies applied an eye tracking and preferential looking scheme to examine participants' attention allocation during motion event perception and categorization (Hohenstein, 2005; Papafragou et al., 2008; Soroli & Hickmann, 2010; Trueswell & Papafragou, 2010; Von Stutterheim et al., 2012). The eye-tracking paradigm is one type of online measurements that aims to capture cross-linguistic variations in speaker's visual attention allocation in decision-making. It is regarded as a lower-level cognitive task and involves automatic and unconscious cognitive processes (Bylund & Athanasopoulos, 2014b; Papafragou et al., 2008; Soroli & Hickmann, 2010).

This paradigm has been widely applied in the motion domain, and usually coupled with both the linguistic (production) and non-linguistic task to examine how subjects across different languages allocate their attention to specific motion components (as measured by patterns of eye movements and eye fixation). For example, Papafragou et al. (2008) tested how native speakers of English (S-language) and Greek (V-language) expressed dynamic motion and how they allocated their visual attention to different motion components (i.e. manner and path of motion) while viewing an unfolding event. The results showed that in verbalization, speakers of both languages did not differ in the frequency of manner of and path selection. However, with regard to attention allocation, their eye movement patterns showed that 1) English speakers were more likely to first allocate their visual attention to the manner regions whereas Greek speakers fixated path areas; and 2) after the end of each video clip, speakers of each languages tended to re-examine the regions that were not typically encoded and highlighted by their language, that is, English speakers tended to look at the path area whereas Greek speakers began to direct their attention to manner of motion. Then the authors attributed these differences to the effects of language-specific features on cognition (i.e. online processing). Similar results were reported by Von Stutterheim et al. (2012), who investigated how languages with aspect marking biased speakers' attention towards trajectories or endpoints of target events. Based on the eye tracking data from children of seven languages, results suggested that speakers from non-aspect languages (i.e. Germany, Swedish) tended to encode endpoints more frequently in verbalization and allocated more attention to endpoints when viewing the target event, whereas aspect-language users (i.e. English, French) chose not to mention endpoints that often in lexicalization and attended more to the ongoing phase on the same event.

In addition, the eye-tracking paradigm is often coupled with non-linguistic tasks. For example, Hohenstein (2005) examined the developmental patterns of children (aged 3.5 to 7 years) in motion event categorization together with a preferential looking scheme. Participants' categorical preferences and eye movements to one of the two alternate stimuli were recorded. Results showed that children aged 3.5 showed an overall preference for path-match alternate in event categorization. However, differences were observed for children at the age of 7, when English children clearly preferred the manner-match alternate in event categorization and spent longer time looking at the manner-match variants than the same-aged Spanish speaking children. Similar results were reported by Soroli et al. (2015) that in event categorization of manner- and path-preferences, French-speaking participants directed their visual attention to the path-match alternate at the first instance and their duration of fixations on manner-match alternate was significantly shorter compared to English-speaking counterpart, especially in a verbal encoding condition. The results suggest that the language-specific variations on spatial cognition are multimodal and take on different forms. Thus, diverse methods need to be incorporated to capture these multi-dimensional cross-linguistic differences in language and cognition research.

More recently, reaction times have been used as another subtle online implicit measurement that can provide indirect evidence for cross-linguistic variations in the processing efficiency and reasoning patterns of motion event cognition (Ji, 2017; Ji &

Hohenstein, 2014, 2017, 2018; Soroli, Hickmann, Hendriks, Engemann, & Vincent, 2015; Von Stutterheim, 2003). According to the current literature, reaction times are usually coupled with the non-linguistic categorization task to tap into the language-specific effects on subjects' processing mode. In a recent study, Ji and Hohenstein (2018) examined how Chinese and English children (3-year olds and 8-year olds) and adult participants categorized and responded to caused motion in a similarity judgement task with verbal interference. Two types of measurements were used: a similarity judgement task that reflected participants' categorical choices and the reaction time that indicated their processing efficiency. The results suggested that participants showed an overall preference for path in event categorization regardless of age and language group. However, their reaction times to manner and path preferences patterned with the typological properties of each language: speakers of English speakers reacted more quickly to manner-match variants than path-match variants, while speakers of Chinese reacted equally quickly to manner- and path-match variants. The results provided evidence for the robust effects of language on thought even when the access to language was blocked via a verbal shadowing during implicit processing. The authors attribute the processing differences to the typological features of English (satellite-framed) and Chinese (equipollent-framed) that speaker of each language may adopt different processing strategies during event perception (manner-salience processing strategy vs. manner- and path salience processing strategy). Similar results were reported by Ji and her colleagues that the different processing efficiency in manner- and path-match preferences were also documented in different types of motion events (voluntary vs. caused motion), with Chinese and English monolingual speakers (Ji & Hohenstein, 2017, 2018), and L2-English learners (Ji, 2017).

The importance of including reaction times as an online measurement lies in that discrepancies may exist between participants' categorical preferences (as measured by the triads-matching paradigm) and automatic and implicit responses (as measured by RTs) in tasks such as the grammatical judgement tasks or sentence interpretation (Kempe, 1999). One example of immediate relevance is Kempe (1999) which examined how Russian and German speakers processed agentivity when presented with different types of linguistic cues (i.e. word order, case marking, and animacy etc.). In a picture-choice paradigm, participants listened to transitive sentences while making choices between two visible agents. The results suggested that both Russian and German speakers relied on case marking as a selection criteria to locate the agent of a target sentence. However, differences in their RT suggested different types of cues had different magnitude in affecting processing efficiency. That is, Russian speakers reacted significantly more quickly in processing agentivity than German speakers, suggesting that case marking served as a stronger cue in Russian that in German. Thus, reaction time can serve as a good indicator to tap into participants' processing efficiency during the decision-making process.

In a similar fashion, the RT in the domain of motion can also be used as one type of implicit measurement to indicate how fast the information of manner or path gets activated and retrieved from participants' memory and used for categorical judgements. The efficiency of information retrieval is closely related to its linguistic prominence, that is, if a semantic element is made more linguistically codable and more accessible in daily interactions, it will become cognitively more available, and takes less time and efforts to be retrieved from memory during cognitive processing (Lupyan, 2012; Slobin, 2004, 2006; Winawer et al., 2007).

### 5.2.2. Cross-linguistic differences in motion event conceptualization: bilingual speakers

Beyond studies of monolingual speakers, only few studies have probed into the effect of language on thought in bi- or multilingual speakers beyond language production. The core issue with bilingual speakers or L2 learners lies in whether learning an additional language with contrastive typological features will give rise to cognitive restructuring of the original conceptual categories and the potential factors that modulate this process (Bylund & Athanasopoulos, 2014b).

On the one hand, some studies suggest that the already established conceptual categories in the L1 are stable and resistant to change regardless of the L2 proficiency (Aveledo & Athanasopoulos, 2016; Cadierno, 2010; Filipovic, 2018). For example, Filipovic (2018) examined the lexicalization patterns and recall memory in causation events (intentional vs non-intentional) with late English-Spanish and Spanish-English bilinguals. Results showed that bilinguals continuously relied on their habitual L1-based patterns of thinking for speaking as an aid to facilitate recall memory, although L2-based labelling also got activated in language production.

On the other hand, other studies demonstrate that learning an additional language means acquiring a new way of thinking, especially when participants' access to language use is not blocked in cognitive processing (Athanasopoulos & Albright, 2016; Athanasopoulos, Bylund, et al., 2015; Athanasopoulos, Damjanovic, et al., 2015b; Brown & Gullberg, 2011; Filipović, 2011; Flecken, Carroll, et al., 2015; Hohenstein et al., 2006; Kersten et al., 2010; Lai et al., 2014; Montero-Melis et al., 2016). For example, by using a supervised learning paradigm, Kersten et al. (2010) examined how Spanish-English bilinguals classified novel motion events based on manner-path contrasts. Results showed that bilinguals with early exposure to English (< 6 years) patterned with English monolinguals in sorting manner-based events when tested in an English-instructed context compared with a Spanish-instructed context, indicating that immediate language use can activate the association between language-specific labels and their correlated mental representations. However, such co-activation was not observed in late bilinguals (> 6 years). On the contrary, Lai et al. (2014) examined how English-Spanish bilinguals categorized voluntary motion in different priming conditions. Results suggested that late bilinguals in a Spanish-priming condition were more likely to base their judgements on path of motion than those in an English-priming condition. However, early bilinguals presented a converged mode of thinking patterns regardless of language in operation. Similar findings were reported by Montero-Melis et al. (2016) who examined whether recent L2 exposure affected similarity assessments of caused motion in Swedish adult learners of L2 Spanish. Participants repeated L2-priming sentences with different degrees of manner salience prior to making their similarity arrangements. The findings showed that Swedish speakers preferred to base their arrangements on ‘same-path’ criteria when primed with path-biased sentences. In addition, Athanasopoulos, Bylund et al. (2015) further reported that late bilinguals of German-English switched their preferences between L1- and L2-based conceptual categories (ongoingness vs goal orientation) in event categorization as a function of language in operation. These findings indicate that conceptual representations in the bilingual mind are flexible and speakers can switch between different language-modulated patterns within a short time. It is suggested that event representations in bilinguals are context-bound and can be modulated by short-term language mediation (Athanasopoulos, Bylund, et al., 2015; Lai et al., 2014; Montero-Melis et al., 2016).

In sum, the overall findings of motion event research show that the effects of language on non-verbal cognition in L2 acquisition are context-bound and modulated by short-term language activation. In addition, the degrees of cognitive restructuring can be modulated by various extra-linguistic factors such as the onset of L2 acquisition (Filipović, 2011; Kersten et al., 2010; Lai et al., 2014), L2 proficiency (Bylund et al., 2013; Ji, 2017; Park & Ziegler, 2014), length of L2 exposure (Athanasopoulos, Damjanovic, et al., 2015b; Park, 2019), frequency of L2 use (Athanasopoulos, Damjanovic, et al., 2015b; Bylund & Athanasopoulos, 2014a; Bylund et al., 2013), as well as abilities for statistical learning (Treffers-Daller & Calude, 2015).

Based on the studies reviewed above, although a series of studies through the lens of 'thinking-for-speaking' and 'thinking-with-language' have provided clear evidence learning a new language means acquiring 'a new way of thinking' when the target language is involved in the decision-making process, little is known about whether the impact of language learning is strong enough to restructure the conceptualization patterns in their native language. In one study of immediate relevance, Wang and Li (2019) examined the extent to which early Cantonese-English bilinguals restructured their L1-based lexicalization and conceptualization patterns as a result of L2 learning by manipulating different language contexts. Specifically, bilinguals were assigned to a monolingual (L1) context where L1 was the only activated language, and a bilingual (L1 and L2) context where both L1 and L2 were kept activated. Results showed that bilinguals patterned with English monolinguals in both event lexicalization and conceptualization regardless of the language context, suggesting that early exposure to a second language has motivated speakers to converge to a single lexicalization and conceptualization pattern compatible with both languages. Thus, it is of theoretical and empirical importance to investigate further how these bilinguals tend to behave when learning an additional language with contrastive typological features.

### 5.3. Defining the Research Gap

Despite compelling empirical evidence from psychological experiments, several issues regarding the language-thought debate remain unresolved. First, the evidence obtained is largely based on the investigation of monolingual speakers or L2 learners through cross-linguistic comparisons. Research on speakers of more than two languages remains limited. With the exception of Bylund and Athanasopoulos (2014a) and Bylund et al. (2013), from a grammatical perspective, very little has been done to examine how speakers of more than two languages conceptualize motion events from a lexical perspective of manner versus path in the domain of motion. As multilingualism is a common linguistic phenomenon worldwide (Aronin & Singleton, 2012) and multilingual speakers display unique linguistic and cognitive features (Cook and Li, 2016), extending research on language-and-thought to multilingualism provides new insights into and have important implications for understanding the process and effects of additional language learning.

Second, although empirical evidence suggests that the effects of language learning on cognitive processing are context-dependent and malleable under various experimental manipulations, it is still unclear whether such effects have long-term and lasting consequences for mental representations across different times, contexts and modalities. For example, although some studies have reported that the experience of long-term L2 learning can lead to a restructuring of conceptual categories established in the L1. It remains unclear how long-term effects of language learning interact with short-term experimental manipulation and language mediations. In addition, most studies examining short-term effect of L2 exposure use L2 linguistic priming or biased instructions as ways to activate or manipulate the immediate language use. However, it remains unclear whether participants are aware of such priming effects and whether this pseudo-priming can properly reflect real-life situations (Montero-Melis et al., 2016). Although Filipović (2011) elicited participants' language productions as task manipulations, the activation of both languages at the same time in the lexicalization may counterbalance potential effects that language placed on cognition. Thus, it is of great theoretical and methodological importance to control different degrees of language activation in linguistic encoding more carefully, for example, by manipulating a monolingual where only the L1 is activated, and a bilingual context where both languages are activated to avoid the potential side effects of task manipulation in terms of the language mediation.

Third, most research along this line is conducted with late bilinguals or adult L2 learners with typologically contrastive languages (satellite- vs. verb-framed languages) (Athanasopoulos, Bylund, et al., 2015; Daller et al., 2011; Lai et al., 2014; Montero-Melis et al., 2016; Stocker & Berthele, 2019). However, little is known about how multilingual speakers of partially overlapping language systems (equipollent-framed language with properties of both satellite- and verb-framed languages) tend to behave: whether they establish distinct sets of lexical and conceptual representations and switch between them, or they have a single pattern of 'thinking-for-speaking' that is compatible with all languages.

### 5.4. Rationale of the current study and the research questions

The general motivation of the current study is to expand the sphere of event cognition from bilingualism to multilingualism by adopting a lexical perspective of manner versus path. It takes a first step in investigating how speakers of three typologically different languages gauge similarity of motion events when the L1-based labelling is activated during cognitive processes. Specifically, it examines the questions of how, and to what extent, the acquisition of an L2-English in childhood and an L3-Japanese in adulthood recalibrate the lexicalization (i.e. event structures and semantic distributions) and categorization patterns (i.e. categorical preferences and reaction time) associated with the L1-Cantonese by a verbal encoding and a non-verbal similarity judgment task. Participants' co-verbal processing efficiency in their decision-making is also measured by the reaction time.

In addition, to further explore how the long-term effect of language learning and short-term effect of language manipulation interact with each other, the current study combines these two lines of enquiry and aims to address how early Cantonese-English bilinguals and Cantonese-English-Japanese multilinguals lexicalize and categorize motion events in different language contexts. Specifically, the current study examines how bilinguals in a monolingual (L1) and a bilingual (L1 and L2/L3) context lexicalize and categorize motion events compared with monolinguals of each language as a function of the recent L2/L3 activation. It also addresses how other long-term factors such as language use with each language and L2/L3 proficiency tend to affect bilinguals' verbal and non-verbal behaviour. Four specific research questions were formulated as follows:

1. How do monolingual controls of Cantonese, English and Japanese lexicalize and categorize motion events when their respective L1 is kept active during the decision-making process? Is higher salience in manner lexicalization associated with more attention to manner in event categorization?

2. How do early Cantonese-English bilinguals in different language contexts (i.e. a monolingual and a bilingual context) lexicalize and categorize motion events compared with the monolingual controls of Cantonese and English?
3. How do Cantonese-English-Japanese multilinguals in different language contexts lexicalize and categorize motion events compared with Cantonese-English bilingual and monolingual controls of each language?
4. What linguistic factors (i.e. language context) and extra-linguistic factors (i.e. language proficiency, language use) may predict the degree of cognitive restructuring in the bi- and multilingual mind?

# Chapter 6. Methodology

## 6.1. Participants

A total of 150 university students took part in the study and divided into five language groups (N=30 each group). Native controls of Cantonese ($M_{age=}$ 22.1, SD=2.7), English ($M_{age=}$ 23.7, SD=1.9) and Japanese ($M_{age=}$ 24.6, SD=2.3) were recruited from local universities in China, UK and Japan, respectively. The monolinguals in the study refer to functional monolinguals with limited proficiency and minimal exposure to any foreign language other than their native language. The dominant language in their daily communication is the native language. To limit the exposure to English, monolinguals of Cantonese were not recruited from HK but from a local university in Shenzhen, China, where Cantonese is the mother tongue. Based on the language education system in China, most participants learn English as a foreign language in schools. School teaching of English is restricted in time and is usually offered in very large classes. Thus, few people who have learned English only through schools (instructed foreign language learning) would be able to function as bilinguals. Likewise, monolinguals of English and Japanese were from local universities in London and Tokyo, respectively. None of the participants of different languages regarded themselves as functional bilinguals based on the information from the language background questionnaire. For all monolinguals in the current study, their daily use of any foreign languages was close to zero. This operationalization of monolingual speakers is in line with previous studies (Athanasopoulos, Bylund, et al., 2015; Brown & Gullberg, 2011; Montero-Melis & Bylund, 2017; Park, 2019), as it is practically impractical to find speakers ofout any exposure or knowledge of any second and/or additional language in today’s bi- and multi-lingual world (Cook, 1992, 2003; Cook & Li, 2016).

Cantonese-English bilinguals ($M_{age}$=20.7, SD=2.1) and Cantonese-English-Japanese multilinguals ($M_{age}$=21.2, SD=1.8) were from Hong Kong with both Cantonese and English as official languages. According to the language education policy in HK, students normally start the L2-English learning from an average age of 3 as early-bilinguals ($M_{age}$=3.7, SD=1.5) and pick up a third language ($M_{age}$=19.2, SD=1.4) as either Major or Minor at university. Due to early exposure to and active use of the L2-English both at school and at home, speakers have already achieved a high level of proficiency in English. Based on the context of Hong Kong, bilingual speakers in the current study refer to balanced bilinguals with high frequency of language use, and an equal level of proficiency in both of their languages.

In line with previous studies and a large body of L2 acquisition research (Athanasopoulos, Damjanovic, et al., 2015b; Montero-Melis et al., 2016; Park & Ziegler, 2014), participants' language proficiency was self-evaluated with a language history questionnaire. In order to take measurements of language proficiency into consideration, in the current study, two forms of self-reported scores were used: scores of any standardized proficiency tests taken within the last two years and self-rated scores of current proficiencies of English for bilingual speakers, and English and Japanese for multilingual speakers. According to the Common European Framework of Reference for Language (Council of Europe, 2011), bi- and multilinguals' English proficiency, and multilinguals' Japanese proficiency was above the upper intermediate level (B2), as measured by their IELTS/TOEFL (for English), and JLPT (for Japanese) scores, respectively. In addition, participants also needed to evaluate their current proficiency in all languages they know in terms of four different categories, namely speaking, listening, reading and writing, based on a seven-point scale where 7 is the maximum rating. The lowest score denotes a 'very poor' level whereas the highest score indicates the level of ‘highly advanced’. Only self-evaluated proficiency scores were used for the statistical analysis. According to the self-rated scores, bilinguals’ proficiency in English (M=6.21; SD=0.46) and multilinguals’ proficiency in English (M=6.18; SD=0.35) and Japanese (M=6.06; SD=0.56) were above the upper intermediate level (B2), as measured by their self-rating scores. Detailed information (Mean and SD) on proficiency levels is presented in Table 2 below. Thus, bilingualism and multilingualism in the current study are defined as an alternate of two or more languages of advanced level of proficiency.

Table 2. Summary of bi and multilingual speakers’ language proficiency

| Language background | English proficiency of CE-bilinguals | English proficiency of CEJ-multilinguals | Japanese proficiency of CEJ multilinguals |
|---|---|---|---|
| Speaking | 6.13 (0.62) | 6.30 (0.65) | 6.10 (0.55) |
| Listening | 6.27 (0.58) | 6.24 (0.45) | 6.07 (0.69) |
| Reading | 6.43 (0.73) | 6.30 (0.72) | 6.10 (0.71) |
| Writing | 5.97 (0.81) | 5.88 (0.86) | 5.93 (0.83) |
| Overall | 6.21 (0.46) | 6.18 (0.35) | 6.06 (0.56) |

Following Bylund and Athanasopoulos (2014a) and Bylund et al. (2013) that in the context of multilingualism, it is necessary to track the amount of language contact with all languages that speakers have, as multilinguals may have more than two L1s or L2s. Thus, the current study measured bi- and multilinguals’ language contact with Cantonese, English or Japanese, respectively. Participants were asked to self-report the time they spent with each language on doing different daily activities (e.g. watching television, surfing the Internet reading for fun/school, and writing for school etc.) within the last three months. Detailed information on multilinguals’ language contact with each language is presented in Table 3. According to participants’ self-estimation, the dominant languages for Cantonese-English bilinguals are Cantonese and English, whereas for Cantonese-English-Japanese multilinguals, the dominant languages are Cantonese and Japanese.

Table 3. Summary of bi and multilingual speakers' contact with each language

| Language background | Amount of Cantonese use | Amount of English use | Amount of Japanese use |
|---|---|---|---|
| CE-bilinguals | 4.78 (1.09) | 6.07 (1.90) | NA |
| CEJ-multilinguals | 3.91(0.92) | 3.22 (1.80) | 5.97 (1.96) |

Note: The frequency of language use in Cantonese, English and Japanese was estimated by hours/day within the last three months.

### 6.2. Materials

To investigate the interplay between language and cognition in speakers of different languages, triangulation of methods was used to connect different types of measurements: verbal encoding, non-verbal categorization and co-verbal processing efficiency of the dynamic stimuli, for a full picture of the effect of language learning on cognitive processing.

Altogether two tasks were used: a language elicitation task and a non-linguistic similarity judgement task. The linguistic encoding task was designed to examine the effects of language-specific categories on speaking itself in speech planning (i.e. verbal evidence), that is, whether speakers of different languages select and structure information differently depending on various linguistic resources available to them. The non-linguistic categorization task further explored whether language-specific differences in linguistic encoding have far-reaching consequences for how speakers perceived and categorized motion events when the access to language is not blocked in the decision-making process (i.e. as measured by the non-verbal evidence of categorical perception and co-verbal processing efficiency).

The involvement of the reaction time as a co-verbal measurement for the processing of the dynamic stimuli is due to the following reasons. First, as proposed by Tokowicz and MacWhinney (2005), reaction time is one subtle type of measurement which directly reflects participants' automatic, non-reflective, and implicit responses in various cognitive processes. This type of measurement has been widely applied in various cognitive domains to capture the language-specific effects on simple, sub-conscious and perceptual decisions (Flecken, Athanasopoulos, et al., 2015; Guillaume et al., 2009; Winawer et al., 2007). Second, as reviewed in Chapter 5, different types of measurements may end up with different results on the dynamic interplay between language and cognition. In the current case, the three target languages under investigation are integrated with partially overlapping language systems (i.e. the equipollent-framed language Cantonese has both satellite- and verb-framed properties). And therefore, in addition to the overt selection of manner- or path-preferences, processing time serves as a subtle and precise measurement to capture cross-linguistic differences that are not easily manifested in speakers' categorical preferences and overt selections. Last but not least, as cognitive restructuring is a dynamic process, it needs to be investigated across tasks such that one can compare the results across tasks to establish genuine individual differences. In this way, a task that combines different types of measurements is like 'multiple shots', which can generate a full picture of one's cognitive profile.

### 6.2.1. Task 1: Linguistic encoding of voluntary and caused motion

Following other well-established studies that use dynamic videos stimuli as the elicitation materials in language production (Hendriks & Hickmann, 2015; Hickmann et al., 2018; Hohenstein et al., 2006; Ji et al., 2011b, 2011c; Soroli & Hickmann, 2010), animated cartoons were specially made for the current study to elicit participants' oral expressions. Altogether 72 sets of test items and 36 sets of control items were included for both voluntary (N=36) and caused motion (N=36). Each cartoon was 6 seconds long. For voluntary motion, the stimuli consisted of 54 sets of animated cartoons with 36 test items and 18 control items. The test items depicted a boy performing a voluntary motion with various types of manner and path against different backgrounds and settings (i.e. along the river, in the forest, and up the snow mountain etc.). Both manner and path in the test items were presented with equal frequency and salience (i.e. A boy walks up a snow mountain). However, the control items minimized the path of motion but only highlighted manner of motion (i.e. A boy is crawling). And all the control items were unbounded events (i.e. events without a clear boundary). The control items were presented against a different background colour in a white background. All 54 items illustrated a boy performing different types of voluntary motion along with one certain type of path and one certain type of manner.

The model used for the linguistic encoding was originally developed by Hickmann and Hendriks (2010). Different from the previous model, the current stimuli included more diverse types of manner and path. For path types, a total of 8 types of path were included, falling into three categories: vertical path (up and down), deictic path (along, towards, away from) and path of a boundary-crossing (across, into, out of). In addition, altogether 11 types of manners were included, ranging from general manners (i.e. walk, run, march), manners with specific movements (i.e. jump, hop, crawl, swim), to manners with instruments (i.e. cycle, skateboard, roller-skate, surfing). The purpose of including different types of manner and path was to avoid the potential effect that a certain type of manner or path may attract more attention compared with other types. The stimuli were fully randomized and counterbalanced across participants in each language group. A whole list of stimuli used in voluntary motion is presented in **Table 4 (test items)**. The control items are presented in **Appendix A.**

In a similar fashion, the linguistic encoding of caused motion was elicited by a total of 54 animated cartoons with 36 test items and 18 control items. The test items depicted a boy (the agent) performing a certain action (i.e. push, drag, pull, kick) on the object which directly caused its movement or displacement (i.e. roll or slide) along a certain trajectory (i.e. into, out of, up and down etc.). Each animation has a clear destination (goal of motion). In addition, the agent moved together with the object throughout the course by the manner of walking in all items. For instance, #item 1, '*A boy is pushing a box up a snow hill'* denoted a boy pushing a box up the hill while walking and the box was sliding up with the boy pushing behind it. The path of the agent and object remained the same as used in voluntary motion. All 54 items illustrated a boy performing different types of caused motion along one certain path and with one certain manner of cause (C-manner).

Following Hickmann and Hendriks (2010), four specific types of manner of cause: pull, push drag and kick together with eight types of path: up and down, into and out of, across and along, towards and away from were covered in the stimuli, with half of the items illustrating non-boundary-crossing events while another half denoting boundary-crossing events. The stimuli were fully randomized and counterbalanced across participants in each language group. A whole list of stimuli is presented in **Table 5 (test items)**. Like voluntary motion, the control items in caused motion also minimized the path of motion but only highlighted manner of cause (i.e. A boy is pushing a box). And all the control items depicted unbounded events (i.e. events that without a clear boundary), as presented in **Appendix B.**

For the linguistic encoding of both voluntary and caused motion, the purpose of using control items with a different background colour was threefold. First, control items (unbounded events) were used to distract participants from using same lexicalization patterns throughout the task. Second, control items were used to test whether bi- and multilingual speakers had already mastered the related vocabulary to describe various types of manner and path in the target language. Last but not least, control items were used as an inherent part of the experimental design to establish a bilingual context, where participants needed to describe all test items in one language whereas control items in another language (see 'Procedure' for detailed information). In this way, speakers' responses to the same test items in a monolingual and bilingual context can be compared with each other. Thus, for the purpose of easy discrimination, control items were presented against a white background.

### 6.2.2. Task2: Non-linguistic categorization of voluntary and caused motion

For the non-linguistic categorization of voluntary motion, a total of 54 animated video clips with 12 sets of test triads and 6 sets of control items were used as stimuli. The test triads shared the same content with the stimuli used in linguistic encoding. This was to make sure that participants had watched and described all scenes prior to event categorization, which aimed to boost participants' language involvement during their decision-making process.

Each triad consisted of three video clips: a target video illustrated a boy performing a voluntary motion event (i.e. A boy walks up a hill), and two alternative videos with either manner or path contrasts. For manner-match alternates, manner of motion was kept consistent whereas the path was changed (i.e. A boy walks DOWN a hill). In contrast, for path-match alternates, path of motion was kept the same whereas manner was altered (i.e. A boy RUNS up a hill). Following Ji (2019) and Ji and Hohenstein (2018), the background settings were kept consistent across each triad to minimize any potential side-effects of event surroundings. This task aimed to examine participants' overt preferences and grouping criteria (whether they were manner- or path-oriented) when categorizing voluntary motion. A whole list of stimuli used in voluntary motion is presented in **Table 4 (test items),** same as the stimuli used in linguistic encoding of voluntary motion. An illustration of test items for voluntary motion is presented in **Appendix E**.

All triads were displayed in a fully-randomized order. The target event appeared first, followed by its manner- and path-match alternates displayed simultaneously on the same screen. The presentation order of each triad was counterbalanced across participants in each group. The placement of manner- and path-match alternates on the screen (right-hand side or left-hand side) was counterbalanced in a fixed order to avoid a potential side effect.

Following Loucks and Pederson (2011) In order to mask the contrast of interest and disturb the regular patterns in the experiment, 6 sets of control items were designed for the categorization of voluntary motion. Among these control items, half of them contrasted Ground with manner of motion, while the other half contrasted Ground with path of motion. The control items are presented in **Appendix C**. An illustration of the control items is presented in **Appendix G**.

Similar as voluntary motion, the stimuli in caused motion consisted of 18 sets of animated videos, with 12 sets of test triads 6 sets of control items. The test items had the same content with the stimuli used in linguistic encoding. This was to make sure that participants had described all scenes prior to event categorization. Each triad contains three animated videos: a target video (e.g. A boy pushes a box into the room), and its two alternates with manner and path as the contrast of interest. For example, for C-manner-match alternate, manner of cause remained the same while path was changed (e.g. A boy pushes a box OUT OF the room) whereas for path-match alternate, path kept the same whereas manner of cause was different (e.g. A boy PULLS a box into the room). This design aimed to examine the participants' preferences between manner of cause and path when categorizing caused motion. In order to keep manner-path as the only contrast of interest, other semantic components in the caused motion (Figure, Ground, and Goal) remain consistent across each test triad. Altogether six sets of manner contrasts (push-pull; push-drag; push-kick pull-kick, pull-drag and drag-kick) and four sets of path contrasts were used (up-down, into-out of, across-along, towards-away from) in the stimuli. And the manner of the object were in two forms: roll and slide, which were counterbalanced across each triad. All stimuli were horizontal motions and the direction of agent's movement (i.e. from left to right or from right to left of the screen) was counterbalanced across each triad. A whole list of stimuli used in voluntary motion is presented in **Table 5 (test items),** same as the stimuli used in the linguistic encoding of caused motion.

Following Loucks and Pederson (2011), 6 sets of control items were introduced to mask the contrast of interest and distract participants from strategically using the same pattern throughout the whole course. Thus, half of these control items contrasted manner of cause with Ground, while the other half contrasted path with Ground. The control items are presented in **Appendix D**. An illustration of the control items for caused motion is presented in **Appendix H**.

Table 4. Dynamic stimuli for voluntary motion in linguistic encoding and similarity judgment task-**Test items**

| Item | Target | Manner-match alternate | Path-match alternate |
|---|---|---|---|
| 1. | Walk down a snow hill | Walk up a snow hill | Crawl down a snow hill |
| 2. | Walk out of a room | Walk into a room | Jump out of a room |
| 3. | Cycle towards a river | Cycle away from a river | Skateboard towards a river |
| 4. | Run down the stairs | Run up the stairs | Hop down the stairs |
| 5. | Cycle out of a castle | Cycle into a castle | Roller-skate out of a castle |
| 6. | Jump away from a tree | Jump towards a tree | Crawl away from a tree |
| 7. | Swim across a river | Swim along a river | Surf across a river |
| 8. | Cycle up a slope | Cycle down a slope | Skateboard up a slope |
| 9. | March across a street | March along a street | Hop across a street |
| 10. | Crawl into a room | Crawl out of a room | Run into a room |
| 11. | Walk towards a house | Walk away from a house | Hop towards a house |
| 12. | Cycle across a street | Cycle along a street | Roller-skate across street |

Table 5. Dynamic stimuli for caused motion in linguistic encoding and similarity judgment task-**Test items**

| Item | Target | Manner-match alternate | Path-match alternate |
|---|---|---|---|
| 1. | Push a box up a snow mountain | Push a box down a snow mountain. | Drag a box down a snow mountain. |
| 2. | Push a wheel across an icy river. | Push a wheel around an icy river. | Drag a wheel across an icy river. |
| 3. | Pull a suitcase down a slope. | Pull a suitcase up a slope. | Drag a suitcase down a slope. |
| 4. | Pull a toy car out of a room. | Pull a toy car into a room. | Drag a toy car out of a room. |
| 5. | Push a chair into a room. | Push a chair out of a room. | Pull a chair into a room. |
| 6. | Push a wheel towards a cave. | Push a wheel away from a cave. | Pull a wheel towards a cave. |
| 7. | Push a ball across an icy river. | Push a ball along an icy river. | Kick a ball across an icy river. |
| 8. | Push a barrel of bear up a hill. | Push a barrel of bear down a hill. | Kick a barrel of bear up a hill. |

| Item | Target | Manner-match alternate | Path-match alternate |
|---|---|---|---|
| 9. | Drag a toy car away from a cave. | Drag a toy car towards a cave. | Kick a toy car away from a cave. |
| 10. | Drag five woods towards a fire. | Drag five woods away from a fire. | Kick five woods towards a fire. |
| 11. | Pull a ball across a street. | Pull a ball along a street. | Kick a ball across a street. |
| 12. | Pull a ball out of a cave. | Pull a ball into a cave. | Kick a ball out of a cave. |

#### 6.2.3. Language background questionnaire

The background questionnaire used in the current study was adapted from Language Background Questionnaire (LHQ3) (Li, Zhang, Tsai, & Puls, 2014), which was widely used as an enhanced tool in assessing the language learning background and individual differences in multilingual speakers and L2 learners. The questionnaire used in the current study was manipulated as an online version in a Google Docs (doc.google.com) format to collect participants' language learning history and background closely related to the current study such as the age of acquisition, language proficiency and frequency of language use. Detailed information of the language background questionnaire is presented in **Appendix I**.

### 6.3. Procedure

#### 6.3.1. Pre-test section

Participants were tested individually by the experimenter in a quiet room at their universities. All the stimuli were displayed and run by the software Superlab 5.0 on a MacBook laptop. In line with Montero-Melis and Bylund (2017) and other well-established studies based on the hypothesis of 'thinking-for-speaking' (Filipović, 2011, 2018; Flecken, Carroll, et al., 2015; Gennari et al., 2002; Lai et al., 2014; Papafragou et al., 2008; Trueswell & Papafragou, 2010; Von Stutterheim et al., 2012), participants described all stimuli in an overt verbal encoding task immediately prior to the subsequent similarity judgement task. This operationalization was to maximally boost the engagement of target language (s) in the decision-making process.

A training session was given at the beginning of each experiment to get participants familiarized with the test procedures. In the training section, four sets of motion event descriptions (2 for voluntary motion and 2 for caused motion) and four sets of similarity judgement task (2 sets for voluntary motion and 2 sets for caused motion) were given before the test section.

### 6.3.2. Test section

In the test section, participants orally described motion videos immediately prior to their non-linguistic evaluation of the same scene. In the first task, participants in each group were asked to watch the cartoon stimuli and describe 'what happened' in each video. Participants were allowed to play the video clips as many times as they wanted. They could also pause between different clips for enough responding time.

Monolinguals of each language (N=30 for each) were instructed and narrated in their native language. To create different language contexts (i.e. a monolingual and bilingual context) for bilingual and multilingual speakers, short-term experimental manipulations were used. In the monolingual context, the L1 was the only activated language. In the bilingual context, L1 and L2 were activated for bilingual speakers, while L1 and L3 were activated for multilingual speakers.

More specifically, bilinguals (N=30) were randomly assigned to a monolingual and bilingual context (N=15 for each). Bilinguals in a monolingual context were asked to describe all test and control items only in their L1-Cantonese. In order to trigger a bilingual context and meanwhile avoid a direct translation effect (Berthele & Stocker, 2017; Stocker & Berthele, 2019), bilinguals in a bilingual context were instructed to narrate all test items in Cantonese whereas the control items in English. This operationalization could ensure that both Cantonese and English were activated in linguistic encoding, and meanwhile, the translation effects were eliminated as participants used different languages to describe the same items. For the purpose of easy discrimination, the control items were presented with a different background colour (in a white background) and participants in a bilingual context were informed before the experiment that items with a white background should be narrated in English. However, responses to the test items should be made in Cantonese.

Following the same reasoning, multilinguals (N=30) were randomly assigned to a monolingual and bilingual context (N=15 for each). Multilinguals in a monolingual context were asked to describe all test and control items in the L1-Cantonese. Multilinguals in a bilingual context were instructed to narrate all test items in Cantonese whereas the control items in L3-Japanese. This operationalization was to ensure that both Cantonese and Japanese got activated. The control items were presented with a different background colour and participants in a bilingual context were informed that items with a white background should be narrated in Japanese.

To keep the language context consistent across language groups, task instructions were given in Cantonese for bi- and multilinguals in a monolingual context whereas in English for bilinguals in a bilingual context, and in Japanese for multilinguals in a bilingual context. The stimuli were fully randomized and counterbalanced in terms of language context and participant groups. All responses were audio-recorded and transcribed following the coding guidelines in Section 6.4. below.

Right after the linguistic encoding, participants moved on to a subsequent similarity judgement task where their overt selections and reaction time were recorded. In terms of task instructions, monolinguals were instructed in their L1s. Bi- and multilinguals in a monolingual context were instructed in Cantonese to keep the L1 as the only active language. Bi- and multilinguals in a bilingual context were instructed in English and Japanese respectively so that both L1 and L2 or L3 were kept active during event conceptualization.

Participants were instructed that the stimuli were presented in a synchronized order where the target video played first at the bottom of the screen. Then the target disappeared right after its completion, followed by its two simultaneous alternates playing side by side at the top of the same screen. A half-second black screen was placed between the target video and its two alternates within each triad and a one-second black screen was placed between triads. The presentation order of each triad was counterbalanced across participants in each group. The location of manner- and path-match alternate on the screen (right-or left-side) was counterbalanced across the stimuli in a fixed order. Participants needed to decide which alternate video was more similar to the target by pressing one of the two keys: ‘A’ and ‘L’ respectively on the keyboard. They were told that there was no correct answer and required to make their decisions as soon as possible as their reaction time to manner- and path-match preferences in the decision-making process was automatically recorded.

After the experimental session, participants were instructed to complete a language background questionnaire. After completing the questionnaire, all sections were finished and participants were rewarded by £10 in cash or Amazon card.

## 6.4. Coding outlines for both linguistic encoding and non-linguistic categorization

### 6.4.1. Coding outlines for the linguistic encoding task

The linguistic data was transcribed from digital video recordings by three native speakers. Only the test items were transcribed and coded for the analysis. The coding was conducted at the casual level and participants' responses were first segmented into clauses. Following Berman and Slobin (1994), a clause is defined as either syntactically simplex or complex sentence containing a unified predicate expressing a single situation. Sentences with infinitives, participles that function as modal verbs (i.e. I think he went) or aspectual markers were not segmented (i.e. A boy wants to go). Matrix sentences with coordinating conjunctions and juxtaposition were considered as two separate clauses.

Responses may involve a sentence containing a single clause as in (35), or a sentence containing a matrix clause with one or more subordinate clause as in (36). When the response contained two or more clauses with coordination or juxtaposition, as in (37), only one of them was included as the target response.

（35）A boy pushed a toy car up to the top of a sand hill. (ENG12CAU)

（36）A boy went down a snow mountain pushing a large suitcase. (ENG13CAU)

（37）A boy is pushing a metal chair into his bedroom and he stops at the foot of the bed (ENG1CAU).

（38）A boy pushed a box and he crossed a street (Potential).

Following Ji et al. (2011b) and Hickmann et al. (2018), the principle of 'one and only one target responses for each item' was respected throughout the data coding.

Descriptions without a specific focus on motion were excluded from the analysis (e.g. The river was frozen). For responses with more than one clause, the principle of ‘semantic richness’ was used to select the target response. For example, in example (38), the first coordinate clause expresses two semantic components (push=Cause + Manner) whereas the second clause contains only one (cross=Path). Thus, the first clause was selected as the target response. For the remaining responses with two equally rich sets of semantic information (i.e. one clause with Manner whereas the other with Path), priority was given to path of motion as it is regarded as the most prominent and basic element to denote a motion event (Talmy, 1985, 2000, 2012).

Following Slobin (2006) and Özçalışkan and Slobin (2003), the degree of manner salience can be assessed by (a) the frequent use of manner expressions across different language contexts (i.e. oral narrative, news reporting, etc.) and (b) the lexical size and diversity of manner expressions in a language. And the major difference between a satellite- and verb-framed language lies in the number of manner and path verbs used in a certain language. Linguistic data in the current study was analysed in terms of three linguistic aspects: 1) the frequency of manner/C-manner and path encoding, 2) the semantic distribution of manner/C-manner and path, and 3) construction types and framing strategies in motion events. Detailed information of data coding is presented in Table 6.

Then the semantic encoding of each clause was conducted right after the systematic sentence segmentation. Within each clause, the semantic components of voluntary and caused motion was first identified. For voluntary motion, two based elements, Manner and Path, were coded in each target response as illustrated in example (39). For caused motion, three basic elements, cause, manner of cause and path were coded as shown in example (40).

（39） A boy walked [Manner of motion] up [Path of motion] a hill.

（40） A boy pushed [Manner + Cause] a box up [Path of motion] a hill.

The segmented and coded descriptions were further analysed from three perspectives: 1) whether the target elements were selected and encoded 2) where the target elements were encoded (i.e. semantic distribution), that is, whether it was encoded in the main verb or via other peripheral devices and 3) specific construction types and framing strategies of motion events. Take the frequency of manner encoding as an example, it was coded as ‘0’ when manner was absent and as ‘1’ when it was selected. This means that instead of coding manner or path for multiple times, the current study focused on whether or not the target element (i.e. manner or path) was expressed within each utterance. The dependent variable was operationalized as binary (the absence or presence of the target element) rather than continuous. Coding the frequency of manner or path selection as a binary dependent variable can better reflect various degrees of manner and path salience speakers attached to manner/path across different languages. The same way of data coding was also adopted by other well-established studies (Montero-Melis & Bylund, 2017; Montero-Melis et al., 2017; Park, 2019; Stocker & Berthele, 2019).

Table 6. Coding guidelines for the linguistic encoding of voluntary and caused motion

| **Type of motion events** | | **Coding outline** |
|---|---|---|
| Voluntary motion and caused motion events | Test items | 1. **Information selection: the frequency of Manner and Path encoding**.<br>The absence (code=0) or presence (code=1) of the target element was calculated within each clause. |

| | | |
|---|---|---|
| | | *2.* **Semantic distribution and information locus.**<br>The semantic distributions of Manner and Path across each utterance was examined. And there were two possible loci: in the Verb or the satellite (OTH). Similar to the frequency of manner or path encoding, the focus was placed on where the target element was encoded at the clausal level. |
| | | 3. **Types of motion event constructions and framing strategies.**<br>Different types of motion event constructions (i.e. Manner verb + Path satellite, Manner-only or Path-only) were examined. Then different construction types were grouped into two basic types of framing strategy (satellite- or verb-framed) for the statistical analysis. In line with the previous coding methods, the absence (code=0) or presence (code=1) of a certain framing strategy was operationalized as a binary variable in the analysis. |
| | Control items | The control items were not included in the semantic scheme. As mentioned in the section of Materials, the involvement of the control items was threefold: 1) to distract participants from using same lexicalization patterns throughout their descriptions, 2) to test whether bi- and multilingual speakers have mastered the vocabulary to describe various types of manner in the target language |

| | | |
|---|---|---|
| | | and 3) to establish a monolingual and a bilingual context |
| | | One point was given when participants used the target manner/C-manner lexicons in the oral descriptions (N=18 for both voluntary and caused motion). |

### 6.4.2. Particular coding issues regarding linguistic data of English, Japanese and Cantonese

The linguistic encoding of each language was followed by language-specific coding guidelines adopted by other well-established studies in the domain of motion. The coding guidelines for the English data were adapted from Hickmann and Hendriks (2006) and Hickmann et.al (2009) and these guidelines were frequently used for the linguistic encoding of motion events. For Japanese data, the coding scheme was based on Brown and Gullberg (2008, 2010, 2013). The coding outlines for Cantonese was mostly based on the guidelines of Yiu (2013, 2014). Part of the coding was also based on the adapted guidelines from the Mandarin Chinese data (Ji & Hohenstein, 2014), due to the partial similarities between Cantonese and Mandarin Chinese. Specific coding guidelines of each language are presented in line with the language-specific coding examples as illustrated in Chapter 4.

It is worth mentioning that following the language-specific coding guidelines, deictic verbs such as ‘come’ and ‘go’ were regarded as one specific form of path and coded as path of motion in all three languages. In addition, for the specific motion ‘climb’, it was coded as a path verb in Japanese (i.e. noboru) as it can only denote the meaning of ‘an upward direction’. However, it was coded as a manner verb in English and

Cantonese as it has a border sense of meaning (the manner of motion) and can be paired up with directional particles (e.g. ‘climb up’ and ‘climb down’).

In order to establish the reliability of data coding (i.e. the inter-rater reliability), 15% of the whole data set was re-segmented and re-coded by a second coder. For English data, 100% agreement was reached on data segmentation, identification and semantic encoding. In sum, 95% agreement was researched for Japanese data, and 97% agreement was obtained for the Cantonese data. Disagreement was settled by accepting the judgement of language experts in each language. Please note that the disagreement was only found in a very limited number of cases.

### 6.4.3. Coding outlines for the similarity judgment task

In the similarity judgement task, two types of variables were used to measure the data: 1) a categorical variable of overt selection (i.e. manner- or path-match choice), and 2) a continuous variable of reaction time (RT). For the overt selection, participants’ choices in their similarity judgements were coded as a binary dependent variable where ‘0’ represented participants’ choice for path-match alternate, and ‘1’ for manner-match alternate. Detailed information for the categorical choice is given in Table 7.

Participants’ RTs to manner- and path-match variants in each triad was measured in milliseconds as a continuous variable and calculated from the onset of displaying of the alternate video until participants made their decisions. Theoretically, the longest RT to each triad was 6 seconds (the same length as the video clip). As participants were required to make their decisions as fast as they can, most of the participants already had their decisions made before the video finished playing. The RT data was first cleaned by trimming the outliers of extremely long and short values with plus and minus two standard deviations (SD) from the mean. After the data trimming, 95% of the original data was kept for the final analysis.

Table 7. Coding outlines for overt selection in non-linguistic categorization task

| **Type of Motion** | | **Options** | **Coding method** |
|---|---|---|---|
| Voluntary motion events | Manner-Path contrast | Manner-match alternate (1) | As there were only two options available in this task (manner- and path-match), and choosing one option means rejecting the other. Thus, the total proportion of manner choices was calculated. |
| | | Path-match alternate (0) | |
| Caused motion events | Cause-Path contrast | Cause-match alternate (1) | As there were only two options in this task (C-manner- and path-match), and choosing one option means not choosing the other. Thus, the total proportion of C-manner choices were calculated. |
| | | Path-match alternate (0) | |

Mixed-effects models were used in the current study as it has been regarded as a powerful tool in SLA research for the analysis of different types of second language learning data (Baayen, Davidson, & Bates, 2008; Baten, Hofman, & Loeys, 2011; Cunnings, 2012). The benefits of the mixed-effects models are as follows: 1) the mixed-effects models are better at dealing with multi-level data with hierarchical structures; 2) the random effects in mixed-effects models can include both participant-level and item-level variances and take into consideration various levels of individual differences; and 3) mixed-effects models are better at dealing with missing data and unbalance dataset. In the current case, as participants were collected from different classes of different schools from different areas, the structure of the dataset was multi-level and hierarchical in nature. In addition, as participants were asked to describe a total of 72 independent sets of video stimuli in the linguistic encoding and 24 independent sets of similarity judgements in event categorization, the influence of test items and individual differences in participants could be well taken care of by using this mixed-effects modelling.

# Chapter 7. Experimental results of motion event lexicalization and categorization in Cantonese, English and Japanese monolingual speakers

Chapter 7 basically focuses on the experimental results of motion event lexicalization and categorization for monolingual speakers of each language. The results will be reviewed in terms of two types of motion: Section 7.1 presents the results of voluntary motion, and Section 7.2 discusses about the results of caused motion. Within each section, the results will be reviewed in terms of two experimental tasks: a verbal encoding task for event lexicalization (i.e. event structures and constructions) and a non-verbal task for event categorization (i.e. overt preferences and reaction time). At the end of this section, the typological status of Cantonese, English and Japanese will be reviewed and discussed from both a linguistic and cognitive perspective. The discussion will be supported by the qualitative analysis of participants' linguistic descriptions. Chapter 7 aims to present a comprehensive picture on whether monolinguals of each language followed lexicalization and categorization patterns typical of their languages and provide a baseline for further analysis (i.e. bi- and multilingual behaviour).

## 7.1. Lexicalization and categorization of voluntary motion in Cantonese, English and Japanese monolingual speakers

### 7.1.1. Linguistic encoding of voluntary motion

#### 7.1.1.1 Frequency of Manner and Path encoding across three monolingual groups in voluntary motion

Altogether 3240 target descriptions were included for the final analysis. Participants' linguistic encodings of voluntary motion were first calculated by the frequency of manner and path selection. Their responses to each stimulus were transformed into percentages and the mean percentage of each group was compared in terms of participant group. As shown in Figure 2, participants in each group presented a high tendency of path encoding, with a ceiling effect across three groups (English: M=95.56%, SD=4.25%; Cantonese: M=93.70%, SD=7.68%; Japanese: 95.83%, SD=6.60%). However, with regard to manner encoding, there was a hierarchical decrease across different language groups (English: M=98.70%, SD=4.04%; Cantonese: M=79.07%, SD=11.31%; Japanese: M=67.59%, SD=9.46%).

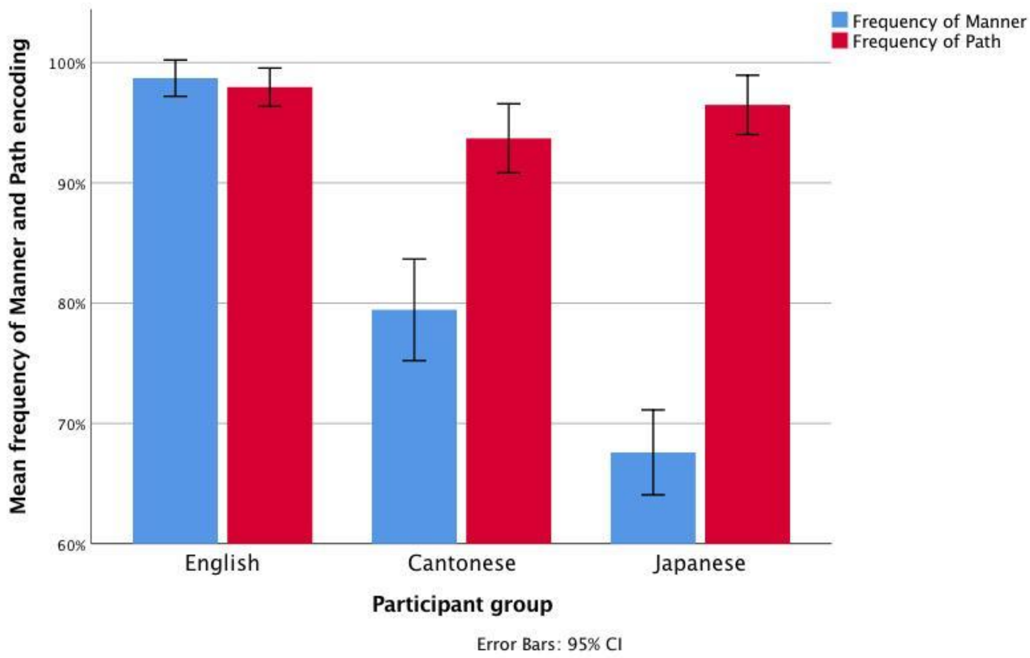



Figure 2. Mean percentage of manner encoding across monolingual groups in voluntary motion

To assess whether speakers from different groups differed in their likelihood of manner and path encoding, two separate logistic mixed-effect models [1] [2] were built with the lme4 package (Bates et al., 2014) in R (R Development Core Team, 2018). Within each model, the binary dependent variable was whether the target semantic element (i.e. manner and path of motion) was encoded (code=1) or not (code=0). The fixed effect was participant group (three levels: English monolinguals vs. Cantonese monolinguals vs. Japanese monolinguals, N=30 for each group). The random effects were random intercepts for participant and item. For path encoding, results showed that the inclusion of group did not significantly increase the model fit compared with the null model ($\chi^2$ (2) =2.79, p=0.247), indicating that group was not a main effect. In other words, participants across different groups were equally likely to encode path of motion when describing voluntary motion.

Table 8. Coefficients of the logistic mixed-effects model for the frequency of manner encoding in voluntary motion

| Fixed effects | Estimate | SE | Wald z | Pr (>\|t\|) |
|---|---|---|---|---|
| **Intercept** | 2.015 | 0.308 | 6.534 | < .001 *** |
| **Participant group** | | | | |
| English | 3.634 | 0.403 | 9.005 | < .001 *** |
| Japanese | -0.885 | 0.242 | -3.644 | < .001 *** |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.662 | 0.814 | | |
| Item (Intercept) | 2.188 | 1.479 | | |

Note: The intercept represents the condition of Cantonese monolinguals as a baseline for between-group comparisons.

However, for the frequency of manner encoding, including participant group as the fixed effect significantly optimized the model compared with the null model ($\chi^2$ (3) =126.93, p<.001), indicating that group was a main effect. Then forward coding was used to compare the likelihood of manner encoding with the next group. As presented in Figure 1, English monolinguals encoded significantly more manner of motion than Cantonese monolinguals ($\beta_{English\text{-}Cantonese}$ = 3.63, SE = 0.40, Wald z =9.01, p < .001), whereas Japanese monolinguals encoded manner least frequently among the three language groups ($\beta_{Cantonese\text{-}Japanese}$ = 0.88, SE = 0.24, Wald z =3.64, p < .001). The findings indicated that in terms of manner and path selection, participants across different language groups demonstrated a clear tendency of encoding path of motion in their oral description when describing voluntary motion. However, three groups differed in how frequently manner of motion was encoded: English group (S-language) showed the highest frequency of manner encoding whereas Japanese group (V-language) had the lowest frequency of manner encoding. Cantonese group was in the middle between S- and V-language group in manner selection. Detailed information of the coefficients for each parameter is presented in Table 8.

#### 7.1.1.2 Semantic distribution of Manner and Path across monolingual groups

Based on the frequency of manner and path encoding, I further explored where each element was semantically distributed within each utterance. As mentioned in the coding outlines, there are two possible loci for the encoding of manner and path: either in the form of main verb (finite verb), or the satellite (infinite verb), such as the subordination or gerund etc. As shown in Table 9, three languages differed from each in terms of information locus, for both manner and path. Results indicated that as a typical S-language, English predominantly encoded manner of motion in the main verb (M=92.9%, SD=11.6%) whereas path in the verb particle (M=91.85%, SD=10.69%). On the other hand, Japanese typically encoded path of motion in the main verb (M=82.51%, SD=8.48%), whereas manner of motion in the form of subordination (M=52.96%, SD=10.81%). As an E-language, Cantonese showed great flexibility in the semantic distribution of manner and path. For manner encoding, Cantonese encoded manner in either the main verb (M= 55.56%, SD=11.67%) or outside of the verb in a subordinate form (M=29.26%, SD=7.43%). Following a similar trend, path of motion was encoded as either directional verbs (M=45.74%, SD=14.27%) or satellites (M=57.71%, SD=16.11%).

Table 9. Percentages of semantic distribution of manner/path in verb (V) or outside of it (OHT) in voluntary motion.

| | English (%) | | | Cantonese (%) | | | Japanese (%) | | |
|---|---|---|---|---|---|---|---|---|---|
| Component | V | OTH | Total | V | OTH | Total | V | OTH | Total |
| Manner | 93.15 | 11.29 | 98.70 | 55.56 | 29.26 | 79.07 | 15.63 | 52.96 | 67.59 |
| Path | 5.56 | 91.85 | 95.56 | 45.74 | 57.71 | 93.70 | 82.50 | 40.92 | 95.83 |

Notes: The sum of the first two columns within each language group doesn't always add up to the total proportion as the manner or path of motion can be double-encoded in V and OTH at the same time (e.g.: The boy is [jumping]verb downstairs [with one leg] OTH).

The semantic distribution of manner and path is in line with the typological status of Cantonese, an equipollent-framed language standing midway on the continuum of S- and V-languages. Take the path encoding for example, it can be either encoded in the main verb as V-languages or in a satellite as S-languages. In addition, the occurrence of path at both Verb and OTH loci is not rare in Cantonese. Examples of language-specific constructions are illustrated in detail below.

(41) **a. English: Manner in the main verb**

A boy is **walking** **[Manner verb]** up a snowy hill. (ENG12vol)

**b. English: Manner in OTH (infrequent)**

A boy is going down some stairs, **jumping with one leg** **[Manner OTH]**. (ENG17vol)

(42) **a. English: Path in OTH**

A boy is walking **from** right **to** left **across** the snow. (ENG11vol)

**b. English: Path in the Verb (infrequent)**

A boy **crossed** **[Path verb]** the street on the skateboard. (ENG10vol)

(43) **a. Japanese: Manner in the main verb (infrequent)**

Kara-wa yama-no ue-ni **hasit-ta** **[Manner verb]**. (JAP19vol)

He-TOP mountain top-TO **run-PST**

'He ran up to the top of the mountain.'

**b. Japanese: Manner in OTH**

| | | | |
|---|---|---|---|
| Kara-wa | kawa-o | **oyoi-de** **[Manner in gerund]** | watat-ta. (JAP1vol) |
| He-TOP | river-ACC | **swimming-GER** | cross-PST. |

'He crossed the river swimming.'

(44) **a. Japanese: Path in the verb**

| | | |
|---|---|---|
| Kara-wa | yuki yama-o | **nobolimashita** **[Path verb].** (JAP2vol) |
| He-Top | snow mountain-ACC | **ascend-PST** |

'He went up a snow mountain.'

**b. Japanese: Path in OTH**

| | | | | |
|---|---|---|---|---|
| Kara-wa | ie-**kara** **[Path satellite]** | eki-**made** **[Path satellite]** | hasit-te | it-ta. (JAP12vol) |
| He-TOP | home-**SOURCE** | station-**GOAL** | run-TE | go-PST |

'He ran from home to the station.'

(45) **a. Cantonese: Manner in the Verb**

| | | | |
|---|---|---|---|
| Go3 naam4 zai2 | **haang4** **[Manner verb]** | zo2 | jap6 seoi6 fong2 (CAN5vol) |
| A boy | **walk** | ASP | into the bedroom |

'A boy walked into the bedroom.'

**b. Cantonese: Manner in OTH**

| | | | | |
|---|---|---|---|---|
| Naam4 zai2 | **caai2-zyu6** **[Manner adjunct]** | daan1ce1 | lok6-zo2 | saan1 (CAN4vol) |
| A boy | **cycling-DUR** | bicycle | descend-ASP | hill |

'A boy descended the mountain cycling.'

(46) **a. Cantonese: Path in the Verb**

| | | | |
|---|---|---|---|
| Go3 naam4 zai2 | **seong5** **[Path verb]** | gan2 | lau4tai1. (CAN1vol) |
| A boy | **ascend** | ASP | stairs |

'A boy is ascending the stairs.'

**b. Cantonese: Path in OTH**

| | | | | |
|---|---|---|---|---|
| Go3 naam4 zai2 | haang 4 | gan2 | **lok6** **[Path satellite]** | lau4tai1 (CAN2vol) |
| A boy | walk | ASP | **down** | stairs |

'He is walking down the stairs.'

#### 7.1.1.3. Framing strategies of voluntary motion across three monolingual groups

To further explore the cross-linguistic differences in lexicalization of voluntary motion for monolinguals of English, Cantonese and Japanese, a qualitative analysis was conducted regarding construction types and framing strategies in each language. Regarding construction types, English monolinguals predominantly used ‘Manner verb + Path satellite’ construction, a typical satellite-framing strategy, whereas the verb-framing ‘Manner adjunct + Path verb’ construction was hardly used in their oral expressions, as illustrated in example (47) and (48) below.

(47) **Satellite-framing strategy: Manner verb + Path satellite**

A boy is cycling [manner in verb] up [path in satellite] the hill.

(48) **Verb-framing: Manner adjunct + Path verb**

A boy crossed [path in the verb] the road on a skateboard [manner in OTH].

In contrast, Japanese monolinguals most frequently used a ‘Path verb + Manner adjunct (optional)’ construction, a typical verb-framing strategy. However, the application of satellite-framing strategy ‘Manner verb + Path satellite’ was not frequently observed in their oral descriptions due to the typological constraints on the use of manner verbs (i.e. the boundary-crossing constraint and the constraints on the occurrence of manner verbs and goal PPs, as reviewed in Chapter 4). Thus, in most cases, manner in Japanese was either omitted, or expressed via a subordination form, as illustrated in example (49) and (50), while the satellite-framing strategy was used in much fewer cases, as illustrated in example (51).

(49) **Verb-framing strategy (type 1): Path verb only**

Kara-wa yuki yama-o **nobolimashita [Path verb].** (JAP2vol)

He-Top snow mountain-ACC **ascend-PST**

'He went up a snow mountain.'

(50) **Verb-framing strategy (type 2): Manner adjunct + Path verb**

Kara-wa kawa-o **oyoi-de [Manner in gerund]** watat-ta. (JAP1vol)

He-TOP river-ACC **swimming-GER** cross-PST.

'He crossed the river swimming.'

(51) **Satellite-framing strategy: Manner verb + Path satellite**

Kara-wa yama-no-uchi-**kara [path satellite]** ie-**made [path satellite]** hasit-ta **[Manner verb].**

He-TOP mountain foot-from top-TO run-PST

'He ran from the foot of the mountain to the top of the mountain.'

In contrast, as an E-language where path of motion can stand alone as an independent element, Cantonese monolinguals used both satellite- and verb-framing as primary strategies in their linguistic encoding. Example of satellite-framing is illustrated in example (52). As for the verb-framing, there are two sub-types in it. In the first type, manner of motion was not expressed, as shown in example (53). In the second type, manner of motion was encoded in a gerund as illustrated in example (54). Due to different structures available for Cantonese speakers, they could freely choose from different structures based on the degree of manner salience speakers attached to each motion element.

(52) **Satellite-framing strategy: Manner verb + Path satellite**

Go3 naam4 zai2 hai6dou6 **aap2 jia2 tiu3 [Manner verb] yau4 jo2 ji3 yau6 [Path satellite]**

A boy there jump from right to left

'A boy is jumping from left to right.'

(53) **Verb-framing strategy (type 1): Path verb only**

Go3 naam4 zai2 jap6 **heui3** [Path only] gaan1 fong2

A boy entered go (deictic) the room

'A boy entered the room.'

(54) **Verb-framing strategy (type 2): Manner adjunct + Path verb**

Go3 naam4 zai2 **caai2-zyu6** **[Manner adjunct]** daan1ce1 lok6-zo2 saan1

A boy cycling DUR descend-PST the mountain

'A boy descended the mountain cycling.'

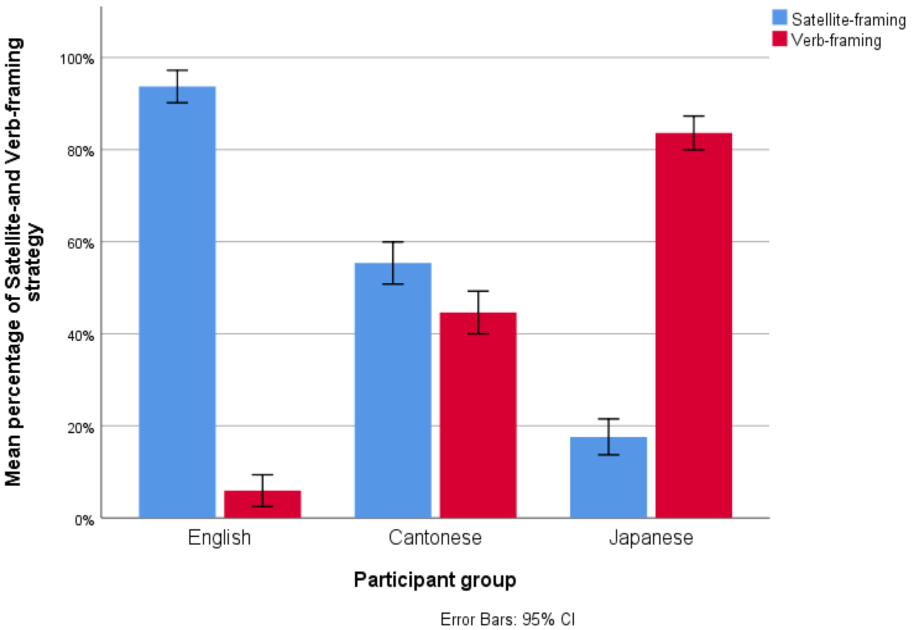



Figure 3. Mean percentage of different framing tendencies (satellite-framing and verb-framing) across monolingual groups in voluntary motion

Quantitative analysis provided further evidence that English monolinguals expressed more manner and adopted a predominant satellite-framing strategy (satellite-framing: M=93.70%, SD=9.43%), while Japanese monolinguals predominantly used verb-framing strategy (satellite-framing: M=17.59%, SD=10.45%). Meanwhile, both the satellite- and verb-framing strategy were frequently examined in Cantonese monolinguals' event descriptions (satellite-framing: M=55.37%, SD=12.25%), as illustrated by Figure 3.

To further explore whether there was a statistical significance in framing strategies used by each monolingual group, a mixed-effects logistic model was fitted. As the total amount of satellite-framing and verb-framing strategies was added up to one, the former was used as the dependent variable. The final model was built [3] with the presence or absence of satellite-framing as the binary dependent variable. The fixed effect was participant group and the random effects were intercepts for participant and item. Results showed that Cantonese monolinguals used more satellite-framing strategies than Japanese monolinguals ($\beta_{\text{Cantonese-Japanese}} = 2.05$, SE = 0.24, Wald z =8.67, $p < .001$). Meanwhile, English monolinguals used satellite-framing strategies the most frequently among the three language groups ($\beta_{\text{Cantonese-English}} = -3.03$, SE = -0.27, Wald z =-10.91, $p < .001$).

Table 10. Coefficients of the logistic mixed-effects model for the frequency of satellite-framing in voluntary motion

| **Fixed effects** | **Estimate** | **SE** | **Wald z** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 0.253 | 0.187 | 1.361 | 0.176 |
| **Participant group** | | | | |
| English | 3.028 | 0.278 | 10.908 | < .001 *** |
| Japanese | -2.049 | 0.236 | -8.667 | < .001 *** |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.644 | 0.803 | | |
| Item (Intercept) | 0.330 | 0.575 | | |

Note: The intercept represents the condition of Cantonese monolinguals as a baseline for between-group comparisons.

### 7.1.2. Non-linguistic categorization of voluntary motion

#### 7.1.2.1. Categorical preferences of manner- and path-match alternates in voluntary motion across three monolingual groups

Right after linguistic encoding, a non-linguistic categorization task was manipulated to further explore whether participants' linguistic preferences might exert some influence on their non-linguistic cognition. Altogether two types of measurements were used: a categorical measurement of participants' manner- and path-match preferences, and a continuous measurement of reaction time for their processing efficiency in the decision-making process.

Regarding participants' manner- and path-match preferences in subsequent categorization (Figure 4), English monolinguals (M=68.89%; SD=19.44%) had an overall manner-match preference compared with Cantonese monolinguals (M=44.44%; SD=20.91%), whereas Japanese monolinguals had the lowest frequency of manner preferences (M=28.61%; SD=23.84%), as shown in Figure 4.

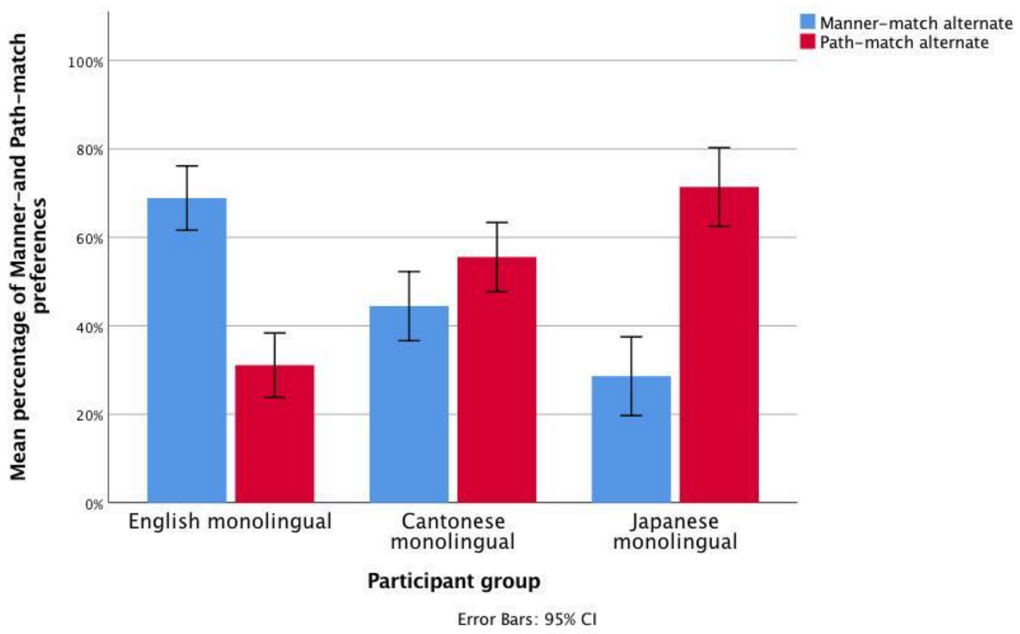



Figure 4. Mean percentage of manner- and path-match preferences across monolingual groups in voluntary motion

A mixed-effects logistic model [4] was fitted to detect whether the observed differences were statistically significant. As the proportion of manner- and path-match preferences was added to 1, the former was used as the dependent variable throughout the statistical analyses. The analysis modelled participants' manner-match preferences as a binary dependent variable. The fixed effect was participant group. The random effects were crossed-random intercepts for participant and item. Including participant group as the fixed-effect significantly optimized the model ($\chi^2$ (3) =21.55, p<.001) compared with the null model, indicating that group was a main effect in participants' similarity judgements. The lack of significance in the overall intercept ($\beta_0$ = -0.30, SE = 0.28, Wald z =-1.08, p = 0.28) indicated that participants across language groups did not have an overall preference for either manner- or path-match alternates. Then forward difference coding was given to compare the likelihood of manner-match selection in this group with the next group. Results confirmed English monolinguals preferred more manner-match alternates than Cantonese monolinguals ($\beta_{English-Cantonese}$= 1.29, SE = 0.31, Wald z =4.24, p<.001) whereas Cantonese monolingual preferred more manner-match preferences compared with Japanese monolinguals ($\beta_{Cantonese-Japanese}$= 0.87, SE = 0.31, Wald z =2.83, p = 0.004), with more details presented in Table 11.

Table 11. Coefficients of the logistic mixed-effects model for manner- and path-match preferences in voluntary motion

| **Fixed effects** | **Estimate** | **SE** | **t value** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 7.656 | 0.031 | 245.694 | < .001 *** |
| **Participant group** | | | | |
| English | 1.295 | 0.305 | 4.235 | 2.28e-05 *** |
| Japanese | 0.866 | 0.306 | -2.828 | 0.004** |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.922 | 0.960 | | |
| Item (Intercept) | 0.420 | 0.648 | | |

Note: The intercept represents the condition of Cantonese monolinguals as a baseline for between-group comparisons.

The findings suggested that speakers of different groups adopted different strategies when categorizing voluntary motion: English speakers were more prone to use ‘manner-match’ criteria when grouping motion events, while Japanese speakers preferred to use ‘path-match’ criteria. As expected, Cantonese speakers used both manner- and path-match as criteria in their categorization, which was backed up by language-specific features of Cantonese, that is, Cantonese exhibits the typological features of both S- (i.e. manner-salience) and V-languages (i.e. path-salience).

#### 7.1.2.2. RT to manner- and path-match alternates in voluntary motion across three monolingual groups

Furthermore, to better explore the degree of differences in participants’ decision-making process, their RTs to manner- and path-match preferences were measured as the continuous variable and used to indicate participants’ processing efficiency in voluntary motion. The mean RT to manner- and path-alternate across each participant group is presented in Table12.

Table 12. Mean RT (in millisecond) to manner- and path-match alternate in voluntary motion

| Participant Group | Mean RT in Manner-match | Mean RT in Path-match |
|---|---|---|
| English | 2187 (SD=735) | 2508 (SD=989) |
| Cantonese | 2393 (SD=956) | 2214 (SD=958) |
| Japanese | 2349 (SD=1207) | 2156 (SD=853) |

To further explore the statistical significance, a mixed-effects model [5] was built with RT as the continuous dependent variable. Fixed effects included participant group (three levels: English, Cantonese and Japanese), preference type (two levels: manner- and path-match preference) and their interaction. The random effects included the crossed random intercepts for participant and item. The dependent variable (RT) was log-transformed in order to meet the assumption of the normality of residuals. Details of fixed-effect parameters are presented in Table 13. The results suggested that with regard to the fixed effects, neither participant group nor preference type was a main effect. However, there was a main effect of participant group and a participant group by preference type effect. This indicated that participants in each group differed in their processing efficiency when selecting manner- or path-match alternates. In addition, the variance of crossed random intercept and slope provided further evidence that the interaction between preference types and participant group remained consistent and systematic across individual subjects and items.

Table 13. Fixed effects on RT as a function of participant group and preference type in voluntary motion

| **Fixed effects** | **Estimate** | **SE** | **t value** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 7.682 | 0.043 | 178.567 | < 2e-16 ** |
| **Preference type** (Manner) | 0.014 | 0.034 | 0.411 | 0.681 |
| **Participant group** | | | | |
| English | -0.049 | 0.064 | -1.772 | 0.079 |
| Japanese | -0.106 | 0.059 | -3.948 | <.001*** |
| **Preference type: Group** | | | | |
| Manner: English | -0.107 | 0.051 | -2.133 | 0.033 |
| Manner: Japanese | 0.145 | 0.053 | 2.746 | 0.006 |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 4.071e-02 | 0.202 | | |
| Item (Intercept) | 9.116e-05 | 0.009 | | |

Note: The intercept represents the log-transformed RT when the preference type is Path-match alternate and participant group is Cantonese.

To further address the interaction between participant group and preference types, three separate mixed-effects models were built with log-transformed RT as the dependent variable and preference type as the fixed effect to address the within group difference. The random effects included crossed random intercepts for participant and item. The intercept for each model set path-match alternate as the benchmark for comparison. For English monolinguals ($\beta_0$ = -0.09, SE = 0.036, t =-2.61, p=0.009), their mean RT to manner-match alternate was faster than path-match alternate. However, for Japanese monolinguals, their mean RT to path-match alternate was faster than manner-match alternate ($\beta_0$ = 0.159, SE = 0.039, t =4.00, p < .001). Meanwhile for Cantonese monolinguals, there is no statistical difference in processing efficiency of manner- and path-match alternate ($\beta_0$ = 0.032, SE = 0.035, t =0.90, p=0.369), as illustrated by Figure 5.

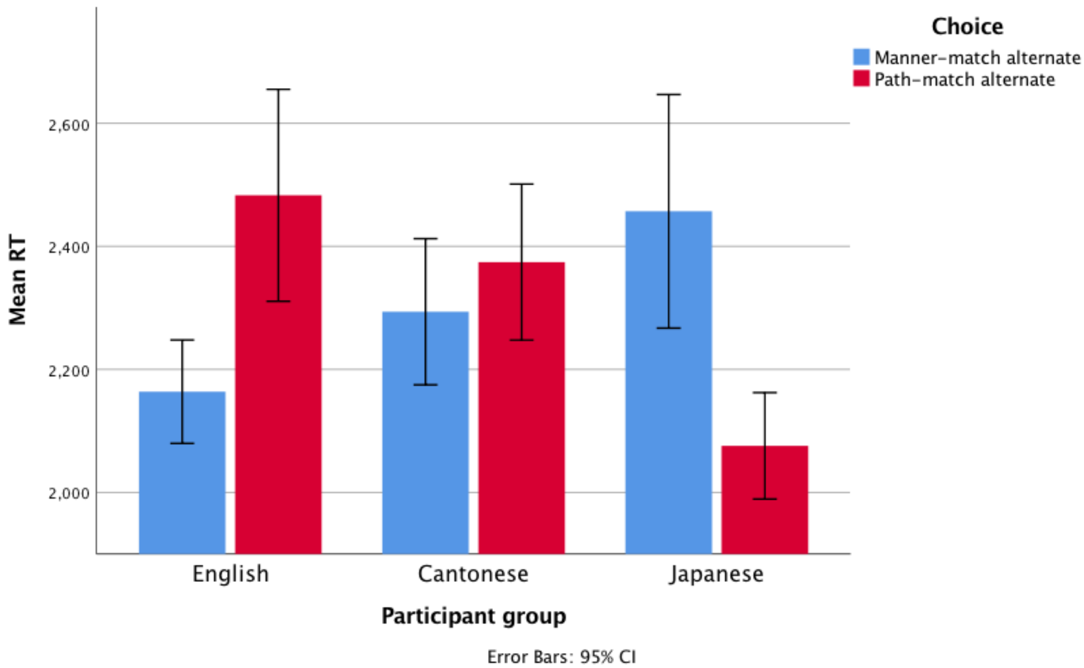



Figure 5. Mean RT to manner- and path-match alternate across participant groups in voluntary motion

The overall results indicated that in terms of both overt selection and processing efficiency of voluntary motion, monolingual controls of Cantonese, English and Japanese exhibited language-specific properties, indicating far-reaching consequences of language learning on event perception and conceptualization.

## 7.2. Lexicalization and categorization of caused motion in Cantonese, English and Japanese monolingual speakers

### 7.2.1. Linguistic encoding of caused motion

#### 7.2.1.1. Frequency of C-Manner and Path encoding across three monolingual groups in caused motion

Altogether 3240 target descriptions (36 items for each participant) were included for the final analysis. Participants' selection of C-manner (Manner of Cause) and path of each stimulus was calculated and transformed into percentages. The mean frequency of C-manner and path encoding was compared in terms of participant group. On the one hand, participants in each group presented a high tendency of path encoding, with a ceiling effect across three groups (English: M=97.31%, SD=3.22%; Cantonese: M=95.83%, SD=5.19%; Japanese: 97.22%, SD=3.34%). However, with regard to C-manner (manner of cause) encoding, a hierarchical decrease was observed across the three language groups (English: M=98.79%, SD=2.89%; Cantonese: M=86.94%, SD=9.51%; Japanese: M=73.79%, SD=10.89%), as shown in Figure 6.

To assess whether speakers from three monolingual groups differed in their likelihood of C-manner and path selection and encoding, two separate logistic mixed-effect models were built [6] [7]. Within each model, the binary dependent variable was whether the target semantic element (i.e. C-manner or path of motion) was encoded (code=1) or not (code=0). The fixed effect was participant group (three levels: English monolinguals vs. Cantonese monolinguals vs. Japanese monolinguals, N=30 for each group). The random effects were random intercepts for participant and item. For the frequency of path encoding, results suggested that the inclusion of group did not significantly increase the model fit compared with the null model ($\chi_2$ (2) =4.68, p=0.09), indicating that participant group was not a main effect. In other words, participants across different groups were equally likely to encode path in caused motion.

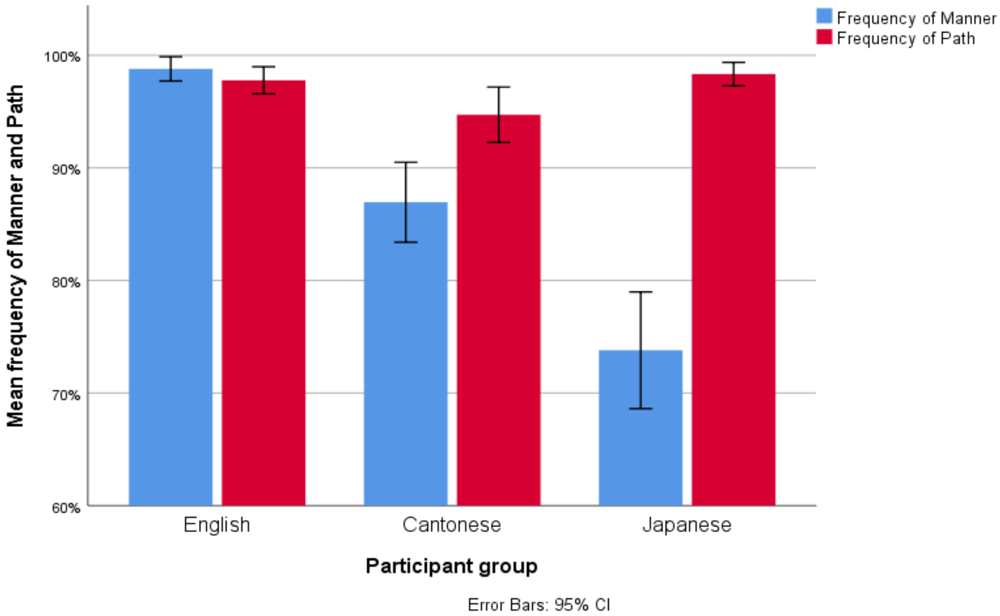



Figure 6. Mean percentage of manner encoding across monolingual groups in caused motion

However, for the frequency of C-manner encoding, including participant group as the fixed effect significantly optimized the model compared with the null model ($\chi_2$ (3) =359.38, p<.001), indicating that participant group was a main effect. Then forward coding was used to compare the likelihood of C-manner encoding with the next group. As show in Figure 6, English monolinguals encoded significantly more C-manner of motion than Cantonese monolinguals ($\beta_{\text{English-Cantonese}}$ = 2.59, SE = 0.29, Wald z =8.88, p < .001), whereas Japanese monolinguals encoded manner least frequently among the three language groups ($\beta_{\text{Cantonese-Japanese}}$ = 0.93, SE = 0.12, Wald z =7.91, p < .001). Detailed information on the statistical models is given in Table 14.

The findings indicated that in terms of the selection of path of motion, participants across different language groups chose to encode path in their oral descriptions as it served a central and core component. However, three groups differed significantly in how frequently C-manner of motion was encoded: English group (S-language) showed the highest frequency of C-manner encoding whereas Japanese group (V-language) had the lowest frequency of manner encoding. Like the linguistic encoding patterns examined in voluntary motion, speakers of Cantonese were situated in the middle between speakers of S- and V-languages in terms of the overall frequency of C-manner selection and encoding.

Table 14. Coefficients of the logistic mixed-effects model for the frequency of C-manner encoding in caused motion

| **Fixed effects** | **Estimate** | **SE** | **Wald z** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 2.115 | 0.172 | 12.334 | < .001 *** |
| **Participant group** | | | | |
| English | 2.587 | 0.291 | 8.882 | < .001 *** |
| Japanese | -0.933 | 0.118 | -7.911 | < .001 *** |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.356 | 0.597 | | |
| Item (Intercept) | 0.292 | 0.540 | | |

Note: The intercept represents the condition of Cantonese monolinguals as a baseline for between-group comparisons.

#### 7.2.1.2. Semantic distribution of C-Manner and Path across three groups

Based on the frequency of C-manner and path encoding, this section further explored where each element was semantically distributed within each utterance. As mentioned in the coding outlines, there are two possible loci for C-manner and path encoding, the form of the main verb, or outside of the main verb in a satellite (i.e. subordination or gerund). As shown in Table 15, three languages differed from each in terms of information locus, for both semantic distribution of C-manner and path. The results indicated that as a typical S-language, English predominantly conflated cause with manner in the main verb (M=98.05%, SD=3.51%) whereas path of motion in the satellite (M=95.93%, SD=4.99%), a typical construction type for S-languages. On the other hand, as a typical V-language, Japanese characteristically conflated cause with path in the main verb (M=73.52%, SD=20.79%), leaving manner of cause unexpressed, or via subordination forms (M=57.59%, SD=17.73%). Located in the middle of the S- and V-language continuum, Cantonese exhibited great flexibility in where C-manner and path were encoded in caused motion. For manner encoding, Cantonese encoded C-manner in either the form of the main verb (M= 48.33%, SD=21.35%), or outside of the verb in a subordinate form (M=38.05%, SD=21.66%). For the semantic distribution of path of motion, it was encoded in either directional verbs (M=42.76%, SD=26.27%) or satellites (M=54.44%, SD=45.48%). The semantic distribution of C-manner and path was in line with the typological status of Cantonese that as an equipollent-framed language, it incorporated typological features of both S- and V-languages and stands midway on the continuum of S- and V-languages.

Table 15. Percentages of semantic distribution of C-manner/path in verb (V) or outside of it (OHT) in caused motion.

| | **English (%)** | | | **Cantonese (%)** | | | **Japanese (%)** | | |
|---|---|---|---|---|---|---|---|---|---|
| Component | V | OTH | Total | V | OTH | Total | V | OTH | Total |
| C-Manner | 98.06 | 1.85 | 98.79 | 48.33 | 38.05 | 86.94 | 16.39 | 57.59 | 73.79 |
| Path | 1.76 | 95.93 | 97.31 | 42.77 | 54.44 | 95.83 | 73.52 | 49.35 | 97.22 |

Notes: The sum of the first two columns within each language group doesn't always add up to the total proportion as the manner or path of motion can be double-encoded in V and OTH at the same time (e.g.: The boy is [jumping]$_{verb}$ downstairs [with one leg] $_{OTH}$).

Qualitative analysis of language-specific properties of motion event encoding was conducted from two perspectives: the semantic distribution of C-manner and Path, and the corresponding construction types. Examples of specific motion event constructions in each language are illustrated below.

In English, C-manner is most frequently conflated with motion in the main verb, as illustrated by example (55). However, it is grammatically incorrect for English speakers to conflate cause with path due to the grammatical constraints.

(55) a. English: Conflating C-manner with motion in the main verb

A boy is **pushing** **[C-manner in the main verb]** a box uphill**.** (ENG11cau)

b. English: Conflating C-manner in OTH (infrequent)

A boy crossed the road **pushing** **[C-manner in OTH]** a box. (ENG18cau)

On the contrary, Japanese speakers frequently conflated C-manner with path of motion, or via the 'te-form' of subordination, as illustrated in (56) and (57). In addition, as Japanese is a case-marking language and allows the use of appositions, path of motion can also be often encoded in other positions.

(56) **a. Japanese: C-manner in the main verb (infrequent)**

Kara-wa sūtsukēsu-o ukiyama-no-ue ni **oshita** **[C-manner in the main verb].**

He-TOP suitcase-ACC mountain top GOAL push PST.

'He pushed a suitcase to the top of the mountain.' (JAP11cau)

**b. Japanese: C-Manner in the form of OTH**

Kara-wa sūtsukēsu-o **oshite** **[C-manner in OTH]** michi-o olimashita **[C-manner in OTH].**

He-TOP suitcase-ACC **pushing-GER** street-ACC cross PST.

'He crossed the street pushing a box.' (JAP12cau)

(57) **a. Japanese: Conflating Cause with Path in the main verb**

Kara-wa sūtsukēsu-o dokuci-ni **ireta** **[C-path in the main verb].**

He-TOP suitcase-ACC cave-GOAL make enter-PST

'He put a suitcase into the cave.' (JAP1cau)

**b. Japanese: Path in OTH**

Kara-wa sūtsukēsu-o dokuci-ni **ireta** **[C-path in the main verb].**

He-TOP suitcase-ACC cave-GOAL make enter-PST

'He put a suitcase into the cave.' (JAP20cau)

For Cantonese, C-manner can be frequently encoded both in the main verb form or subordination. Same is true for path that it could be either encoded in the main verb as V-languages or in a satellite as S-languages. In addition, the occurrence of path at both Verb and OTH loci is not rare in Cantonese. Examples are given in (58) and (59).

(58) **a. Cantonese: C-Manner in the Verb**

Keio4 zoeng1 go3 muk6seung1 **teui1 [C-manner in the verb]** dou2 uk1deng2

S/he DM a box push up to the top of the roof

'S/he pushed a wood box up to the top of the roof.' (CAN20cau)

**b. Cantonese: C-Manner in OTH**

Keio4 **teui1zyu6 [C-manner in OTH]** go3 muk6syun4 hang4dou2 go3 chou2 dei6

S/he **push-DUR** a wood ship walk to a grassland

'S/he walked to a grassland pushing a wood ship.' (CAN9cau)

(59) **a. Cantonese: C-Path in the Verb**

Keio4 zoeng1 zoeng1toi2 **bun1** zo2 [Path in verb] **faan1** uk1kei5

S/he DM a table move-ASP return home

'S/he moved a table back home.' (CAN21cau)

**b. Cantonese: Path in OTH**

Keio4 zoeng1 go3 che1 **yao3 [Path in OTH]** saan1po1 laai1**dou3** saan1po1 mei5

S/he DM a toy car from the top push-PST bottom of the mountain

'S/he pulled a toy car from the top to the bottom of the mountain.' (CAN29cau)

### 7.2.1.3. Framing strategies of caused motion across three monolingual groups

To further explore the differences of motion event encoding in English, Cantonese and Japanese, a qualitative analysis was conducted regarding the construction types and framing strategies in each language. For example, English monolinguals used satellite-framing as the dominant strategy for the encoding of caused motion whereas verb-framing was hardly used. Examples are given in (60) and (61).

(60) Satellite-framing: A boy is pushing a box [C-manner in verb] up [Path in satellite] the hill.

(61) Verb-framing (infrequent): A boy crossed [Path in the verb] the road pushing a box [C-manner in OTH].

In contrast, Japanese monolinguals most often adopted a verb-framing strategy in the description of caused motion whereas satellite-framing is occasionally used.

(62) **Satellite-framing strategy (infrequent)**

Kara-wa sūtsukēsu-o ukiyama-no-ue ni **oshita** **[C-manner in the main verb].**

He-TOP suitcase-ACC mountain top GOAL push PST.

'He pushed a suitcase to the top of the mountain.' (JAP11cau)

(63) **Verb-framing strategy**

a. Kara-wa sūtsukēsu-o dokuci-ni **ireta** **[C-path in the main verb].**

He-TOP suitcase-ACC cave-GOAL make enter-PST

'He put a suitcase into the cave.' (JAP1cau)

b. Kara-wa sūtsukēsu-o **oshite** **[C-manner in OTH]** michi-o olimashita **[C-manner in OTH].**

He-TOP suitcase-ACC **pushing-GER** street-ACC cross PST.

'He crossed the street pushing a box.' (JAP12cau)

Cantonese, an E-language where manner of cause can be conflated with path in the main verb, Cantonese monolinguals used both satellite- and verb-framing as primary strategies in linguistic encoding. Similar to English, the example of satellite-framing is illustrated in (64). Similar to Japanese, there are two sub-types in verb-framing strategies. In the first type, manner of motion was not expressed, as shown in (65). In the second type, manner of motion was encoded in a gerund as illustrated in (66).

(64) **Satellite-framing: C-Manner in the verb, Path outside in OTH**

Keio4 zoeng1 go3 muk6seung1 **teui1** **[C-manner in the verb]** dou2 uk1deng2

S/he DM a box push up to the top of the roof

'S/he pushed a wood box up to the top of the roof.' (CAN20cau)

(65) **Verb-framing (type 1): C-Path in verb without expressing manner**

Keio4 **jap6 zo2** jat1 go3 bo1

S/he **enter-ASP** a ball

'S/he made a ball into a cave.' (CAN21cau)

(66) **Verb-framing (type 2): Path in verb and C-manner in a gerund**

Keio4 **teui1zyu6** **[C-manner in OTH]** go3 muk6syun4 hang4dou2 go3 chou2 dei6

S/he **push-DUR** a wood ship walk to a grassland

'S/he walked to a grassland pushing a wood ship.' (CAN9cau)

As shown in Figure 7, English monolinguals most typically conflated cause with manner in the main verb form whereas path in the satellite, and adopted a predominant satellite-framing strategy (satellite-framing: M=98.08%, SD=3.28%). On the contrary, Japanese monolingual most often conflated cause with path in the main verb, leaving manner of cause unexpressed, or expressed via a subordination 'te-' form. In other words, Japanese speakers predominantly used a verb-framing strategy (satellite-framing: M=27.31%, SD=20.20%). Meanwhile, both satellite-framing and verb-framing strategies were used frequently in Cantonese monolinguals' descriptions of caused motion (satellite-framing: M=57.22%, SD=46.00%).

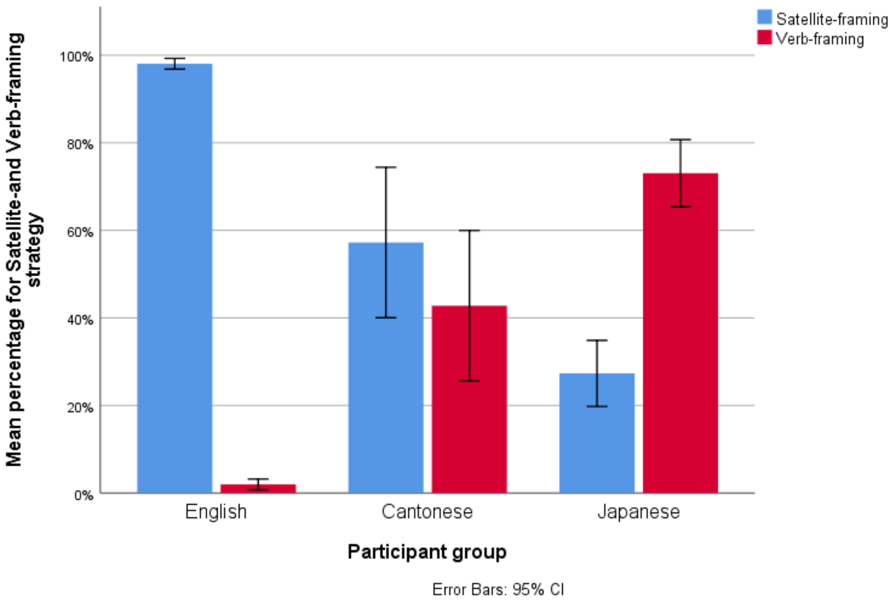



Figure 7. Mean percentage of different framing tendencies (satellite- and verb-framing) across monolingual groups in caused motion

To further explore the statistical significance in the framing strategy used by each monolingual group, a mixed-effects logistic model [8] was fitted. As the total amount of verb-framing and satellite-framing strategies was added up to one, the former was used as the dependent variable. The final model was built with the presence or absence of satellite-framing as the binary dependent variable. The fixed effect was participant group. As adding item as one of the random intercept caused the model failed to converge, the final model included and the random effect for participant only. The results showed that Cantonese used less satellite-framing strategies than English ($\beta_{\text{Cantonese-English}}$ = -4.26, SE = 0.25, Wald z =17.27, p < .001), where Japanese was the lowest ($\beta_{\text{Cantonese-Japanese}}$ = 1.69, SE = 0.11, Wald z =-15.11, p < .001).

Table 16. Coefficients of the logistic mixed-effects model for the frequency of satellite-framing in caused motion

| **Fixed effects** | **Estimate** | **SE** | **Wald z** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 0.386 | 0.249 | 1.552 | 0.121 |
| **Participant group** | | | | |
| English | 4.257 | 0.247 | 17.268 | < .001 *** |
| Japanese | -1.689 | 0.112 | -15.107 | < .001 *** |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 1.689 | 1.299 | | |

Note: The intercept represents the condition of Cantonese monolinguals as a baseline for between-group comparisons.

### 7.2.2. Non-linguistic categorization of caused motion

#### 7.2.2.1. Categorical preferences of manner- and path-match alternates in caused motion across three monolingual groups

In the subsequent similarity judgement task, participants' manner- or path-match preferences are presented in Figure 8. According to the data visualization, it was quite clear that participants across each language group had an overall preference for path-match alternates: English monolinguals (M=65.56%; SD=30.93%), Cantonese monolinguals (M=63.89%; SD=28.85%), and Japanese monolinguals (M=70.56%; SD=25.40%) had a consistent pattern for path-match preferences.

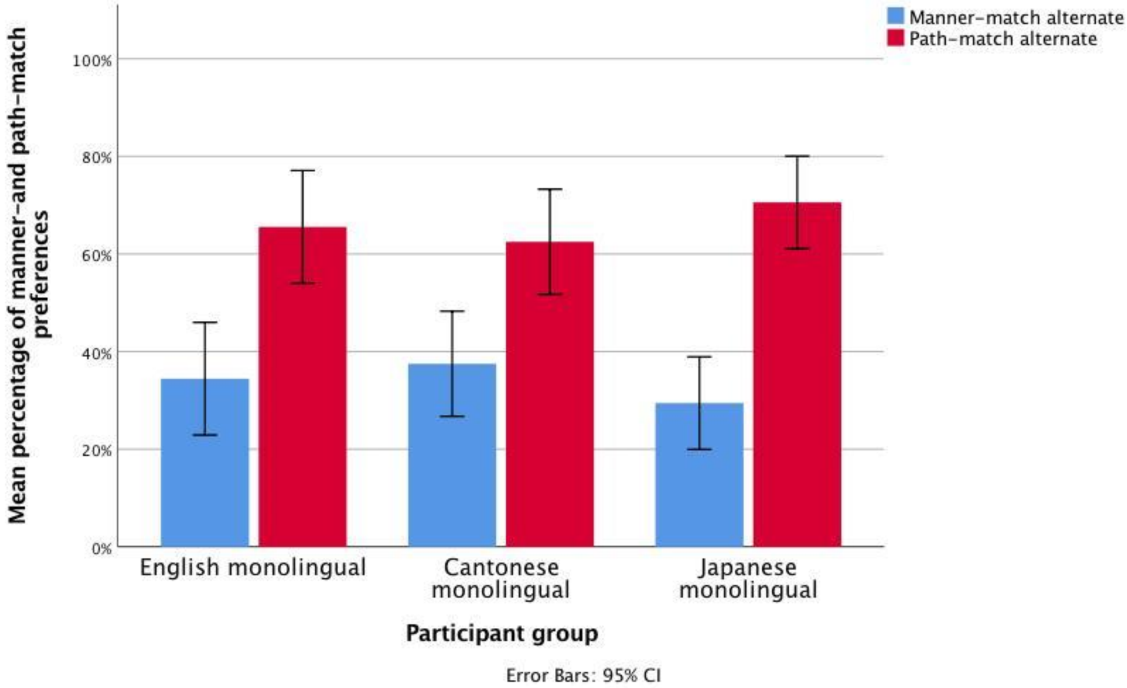



Figure 8. Mean percentage of manner- and path-match preferences across monolingual groups in caused motion

To further support the overserved tendency in the overt selection of caused motion, another mixed-effects logistic model [9] was built to detect whether the observed differences were statistically significant. As the proportion of path-match preferences was 1 minus the proportion of manner-match preferences, the former was used as the dependent variable throughout the statistical analyses. The analysis modelled participants' path-match preferences as a binary dependent variable. The fixed effect was participant group. The random effects were crossed-random intercepts for participant and item. Including participant group as the fixed-effect significantly did not significantly optimize the target model ($\chi^2$ (2) =4.45, p=0.108) compared with the null model, indicating that group was not main effect in participants' similarity judgements of caused motion events. In other words, participants across different language group were more prone to use path-match as their selection criteria when categorizing caused motion events.

#### 7.2.2.2. RT of manner- and path-match alternates in caused motion across three monolingual groups

Although no cross-linguistic differences was examined in the overt selection of manner- and path-match variants, the continuous measurement of RTs demonstrated the processing efficiency in participants' decision-making process. The mean RT to manner- and path-alternate across each participant across the three language groups is presented in Table 17.

Table 17. Mean RT (in millisecond) to manner- and path-match alternate in caused motion

| Participant Group | Mean RT in Manner-match | Mean RT in Path-match |
|---|---|---|
| English | 2048 (SD=683) | 2416 (SD=840) |
| Cantonese | 2165 (SD=810) | 2247 (SD=792) |
| Japanese | 2346 (SD=713) | 2000 (SD=692) |

To further explore the statistical significance, a mixed-effects model was built with RT as the continuous dependent variable [10]. The fixed effects included participant group (three levels: English, Cantonese and Japanese), preference type (two levels: manner- and path-match preference) and their interaction. The random effects included the crossed random intercepts for participant and item. The dependent variable (RT) was log-transformed to meet the assumption of the normality of residuals. Details of fixed-effect parameters are presented in Table 18. The statistical significance in the overall intercept ($\beta_0$ = 7.66, SE = 0.038, t =196.43, p<.001) indicated that participants across language groups had an overall preference for path-match alternates. However, the fixed effects suggested that there was a main effect of participant group and a participant group by preference type effect. This indicated that participants in each group differed in their processing efficiency when making their similarity judgements. In addition, the variance of crossed random intercept and slope further supported that the interaction between preference types and participant group did not appear at chance level and it presented a systematic tendency across individual subject and individual test item. In addition, Japanese monolinguals were significantly quicker in making path-match alternates compared with their monolingual peers in both Cantonese and English.

Table 18. Fixed effects on RT as a function of participant group and preference type in caused motion

| **Fixed effects** | **Estimate** | **SE** | **t value** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 7.655 | 0.038 | 196.428 | < 2e-16 *** |
| **Preference type (Manner)** | -0.044 | 0.035 | -1.237 | 0.2165 |
| **Participant group** | | | | |
| English | 0.058 | 0.029 | 2.022 | 0.043 |
| Japanese | -0.109 | 0.028 | -3.859 | <.001*** |
| **Preference type: Group** | | | | |
| Manner: English | -0.075 | 0.051 | -1.459 | 0.145 |
| Manner: Japanese | 0.200 | 0.051 | 3.937 | <.001*** |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.025 | 0.159 | | |
| Item (Intercept) | 0.003 | 0.054 | | |

Note: The intercept represents the log-transformed RT when the preference type is Path-match alternate and participant group is Cantonese.

To further address the interaction between participant group and preference types, three separate mixed-effects models [11] [12] [13] were built with log-transformed RT as the dependent variable and preference type as the fixed effect to address the within group difference. The random effects included crossed random intercepts for participant and item. The intercept for each model set path-match alternate as the benchmark for comparison. The analysis showed a significant difference in the processing efficiency of manner- and path-match preferences across the monolingual groups. To be more specific, English monolinguals responded more quickly to manner-match than path-match variants in their similarity judgements ($\beta_0 = 0.159$, SE = 0.039, $t = 4.00$, $p < .001$). In contrast, Japanese monolinguals reacted more quickly to path-match than manner-match variants ($\beta_0 = 0.159$, SE = 0.039, $t = 4.00$, $p < .001$). Meanwhile, there was no statistical difference in processing efficiency for Cantonese monolinguals ($\beta_0 = 0.032$, SE = 0.035, $t = 0.90$, $p = 0.369$), as illustrated by Figure 9. The results showed that although participants of different languages had an overall preference for path-match variants, the RTs to manner-match vs. path-match varied significantly. The disparity between the overt selection and RTs indicated the importance to employ different types of measurements when addressing participants' performances in cognitive processing.

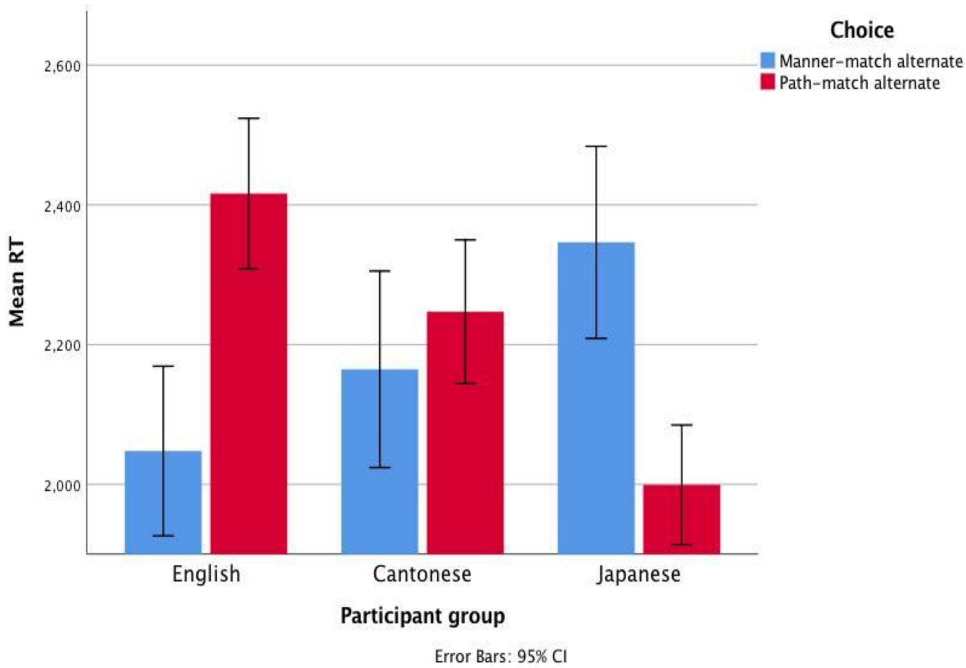



Figure 9. Mean RT to manner- and path-match alternates across monolingual groups in caused motion

# Chapter 8. Experimental results of motion event lexicalization and categorization in Cantonese-English bilingual speakers

Chapter 7 confirms the typological status of Cantonese (E-language), English (S-language) and Japanese (V-language) from a linguistic and cognitive perspective. Based on the monolingual data in both event lexicalization (i.e. event structures and constructions) and categorization (overt preferences and reaction time), Chapter 8 further explores how early Cantonese-English bilinguals lexicalized and categorized motion events in different language contexts compared with monolinguals of each language. Specifically, this chapter addresses how bilinguals in a monolingual (L1) and a bilingual (L1 and L2) context lexicalized and categorized motion compared with monolinguals of each language as a function of recent L2 activation. In addition, it also examines whether the amount of language contact with each language affected bilinguals' performance while controlling for other variables such as age of L2 acquisition and L2 proficiency.

Chapter 8 includes two sections: Section 8.1 discusses about participants' performances in voluntary motion whereas Section 8.2 focuses on the domain of caused motion.

## 8.1. Lexicalization and categorization of voluntary motion in Cantonese, English and early Cantonese-English bilingual speakers

### 8.1.1. Linguistic encoding of voluntary motion

#### 8.1.1.1. Frequency of Manner and Path encoding across Cantonese, English and Cantonese-English bilingual speakers

A total of 4320 target responses (N=36 for each participant) were included for the final analysis. Participants' linguistic encodings were calculated by the frequency of manner and path selection. Participants' responses to each stimulus were transformed into percentages and the mean percentage for each group was compared in terms of participant group and language context. Participants in each group presented a high tendency of path encoding, with a ceiling effect across four language groups (Cantonese: M=93.70%, SD=7.68%; Bilingual in a monolingual context: M=95.18%, SD=7.23%; Bilingual in a bilingual context: M=93.13%, SD=9.24%; English: M=95.56%, SD=4.25%). However, regarding the manner encoding, bilinguals and English monolinguals encoded manner more often than Cantonese monolinguals (Cantonese: M=79.07%, SD=11.31%; Bilingual in a monolingual context: M=95.18%, SD=5.50%; Bilingual in a bilingual context: M=97.41%, SD=4.63%; English: M=98.70%, SD=4.04%), as illustrated in Figure 10.

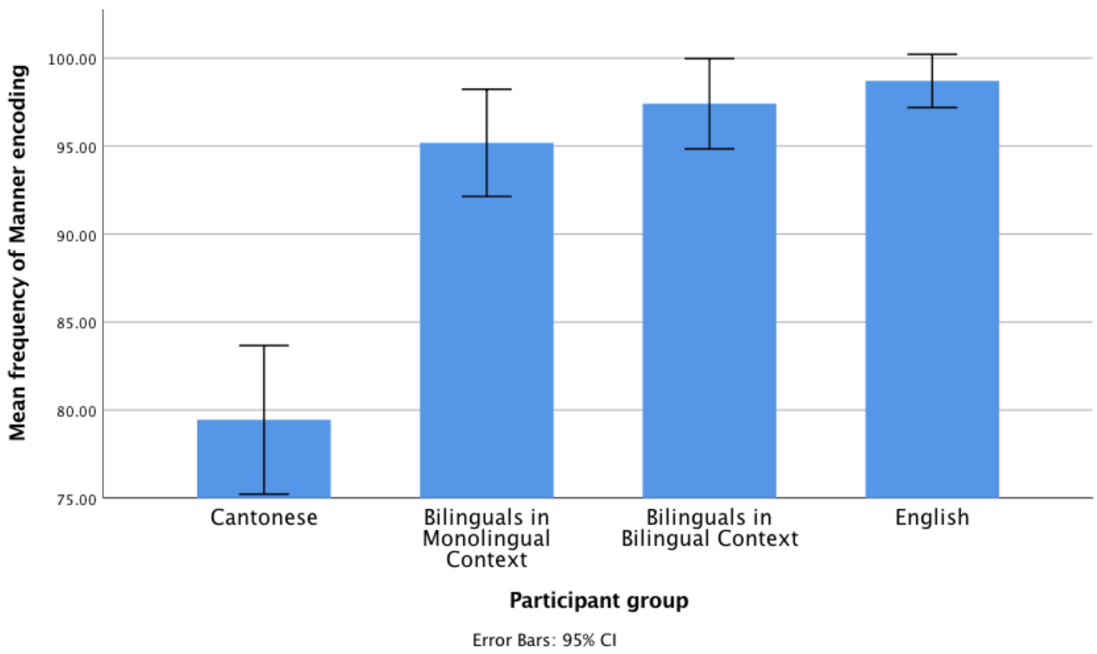



Figure 10. Mean frequency of manner encoding in terms of participant group and language context in voluntary motion

To assess whether speakers from different groups differed in their likelihood of manner and path encoding, two separate logistic mixed-effect models [14] [15] were built with the *lme4* package (Bates et al., 2014) in R (R Development Core Team, 2018). Within each model, the binary dependent variable was whether the target semantic element (e.g. manner and path of motion) was encoded (code=1) or not (code=0). The fixed effect was participant group (four levels: Cantonese monolinguals (N=30) vs. bilinguals in a monolingual context (N=15) vs. bilinguals in a bilingual context (N=15) vs. English monolinguals (N=30)). The random effects were random intercepts for participant and item. For path encoding, results showed that the inclusion of group did not significantly increase the model fit compared with the null model ($\chi^2$ (3) =6.63, p=0.085), indicating that group was not a main effect. In other words, participants across different groups were equally likely to encode path of motion when describing voluntary motion.

However, for the frequency of manner encoding, including participant group as a fixed effect significantly optimized the model compared with the null model ($\chi^2$ (3) =72.29, p<.001), indicating that group was a main effect. Then forward coding was used to compare the likelihood of manner encoding with the next group. As shown in Figure 10, bilinguals in a monolingual context encoded manner more often than Cantonese monolinguals ($\beta_{\text{Cantonese-Bilinguals in monolingual context}}$ = -2.41, SE = 0.55, Wald z =-4.35, p < .001) yet patterned with bilinguals in a bilingual context ($\beta_{\text{Bilinguals in monolingual context--Bilinguals in bilingual context}}$ = -0.96, SE = 0.71, Wald z =-1.36, p = 0.18). Meanwhile, no difference between bilinguals in a bilingual context and English monolinguals was detected ($\beta_{\text{Bilinguals in bilinguals in bilingual context-English monolinguals}}$ = -1.07, SE = 0.71, Wald z =-1.51, p = 0.13). As predicted, Cantonese monolinguals encoded significantly less manner than English monolinguals ($\beta_{\text{Cantonese-English}}$ = -4.44, SE = 0.63, Wald z =-7.10, p < .001). Detailed statistical information is presented in Table 19. The findings indicated that bilinguals demonstrated a cognitive shift towards L2-based encoding patterns regardless of the language context.

Table 19. Coefficients of the logistic mixed-effects model for the frequency of manner encoding in voluntary motion-bilingual speakers

| **Fixed effects** | **Estimate** | **SE** | **Wald z** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 4.037 | 0.713 | 5.662 | < .001 *** |
| **Participant group** | | | | |
| Bilinguals in monolingual context | -0.956 | 0.706 | -1.356 | 0.175 |
| Cantonese | -3.367 | 0.635 | -5.300 | < .001 *** |
| English | 1.072 | 0.711 | 1.508 | 0.131 |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 1.972 | 1.404 | | |
| Item (Intercept) | 2.352 | 1.534 | | |

Note: The intercept represents the condition of Bilinguals in a bilingual context as a baseline for between-group comparisons.

#### 8.1.1.2. Semantic distribution of Manner and Path across Cantonese, English and Cantonese-English bilingual speakers

Based on the frequency of manner and path encoding, I further explored the semantic distribution of manner and path across each utterance. Results indicated that the sematic distributions of manner and path in Cantonese and English reflected the typological status of each language (please refer to the monolingual section for more detail). In addition, the semantic distribution of bilinguals in different language contexts showed a clear shift from L1-Cantonese towards an English-specific way in describing voluntary motion. For example, bilinguals in a monolingual and bilingual context were more prone to encode manner of motion in the main verb (M=69.62%, SD=8.52% for bilinguals in a monolingual context and M=76.30%, SD= 5.89% for bilinguals in a bilingual context), with more details presented in Table 20.

Table 20. Percentages of semantic distribution of manner/path in verb (V) or outside of it (OHT) in voluntary motion.

| | **Cantonese (%)** | | | **Bilinguals in monolingual context (%)** | | | **Bilinguals in bilingual context (%)** | | | **English (%)** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Component | V | OTH | Total | V | OTH | Total | V | OTH | Total | V | OTH | Total |
| Manner | 55.56 | 29.26 | 79.07 | 69.26 | 32.59 | 95.19 | 76.30 | 24.44 | 97.41 | 93.15 | 11.29 | 98.70 |
| Path | 45.74 | 57.71 | 93.70 | 43.33 | 60.00 | 95.19 | 32.22 | 62.86 | 92.96 | 5.56 | 91.85 | 95.56 |

Notes: The sum of the first two columns within each language group doesn't always add up to the total proportion as the manner or path of motion can be double-encoded in V and OTH at the same time (e.g.: The boy is [jumping]verb downstairs [with one leg] OTH).

#### 8.1.1.3. Construction types and framing strategies of voluntary motion event across Cantonese, English and Cantonese-English bilingual speakers

A qualitative analysis was conducted with regard to construction types and framing strategies of each language group. As reviewed in the monolingual section, the typological status of Cantonese and English exhibited typical linguistic features of an equipollent-framed and a satellite-framed language, respectively. With regard to the linguistic data of bilingual speakers in different language contexts, there was a clear shift from the L1- to L2-based patterns in motion event constructions and framing strategies, that is, bilinguals used more 'Manner verb +Path satellite' constructions (satellite-framing strategy) and fewer 'Path only' constructions (verb-framing strategy) compared with speakers of Cantonese. For example, bilingual speakers tended to mention manner of motion more often when describing the same motion scene (i.e. A boy walked into the room), indicating a very clear effect of English learning on the lexicalization of voluntary motion. Examples are given in (67) and (68):

(67) **Verb-framing strategy (type 1): Path verb only**

| | | | |
|---|---|---|---|
| Go3 naam4 zai2 | jap6 | zo2 | gaan1 fong2 |
| A boy | entered | ASP | the room |

‘A boy entered the room.’

(68) **Satellite-framing strategy: Manner verb + Path satellite**

| | | | |
|---|---|---|---|
| Go3 naam4 zai2 | haang6-zo2 | jap6 | gaan1 fong2 |
| A boy | walk-ASP | into | the room |

‘A boy walked into the room.’

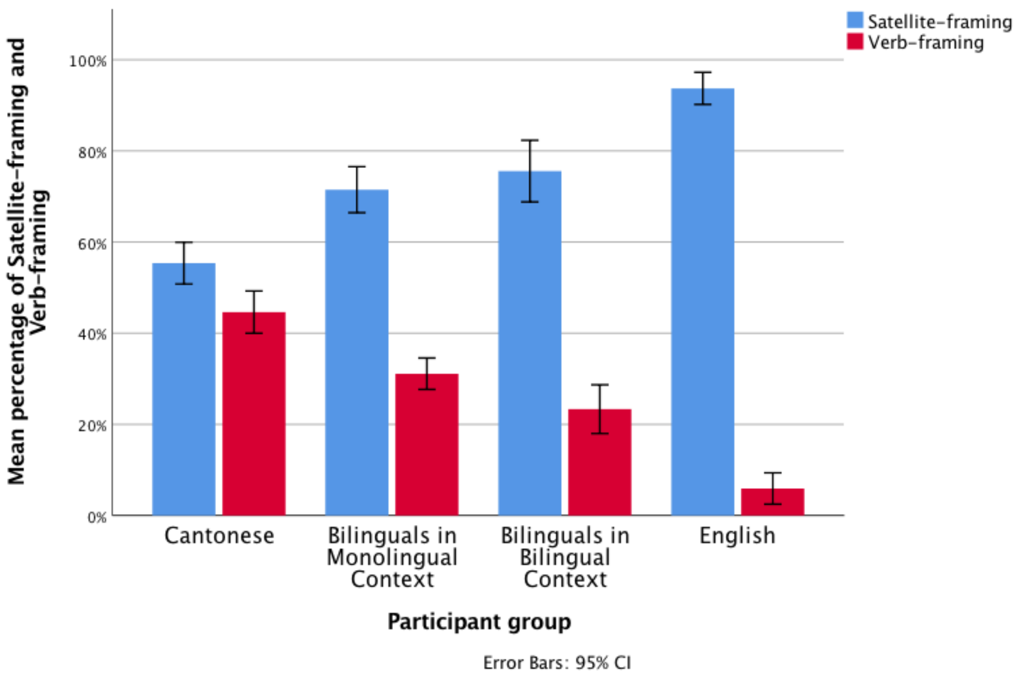



Figure 11. Mean percentage of different framing tendencies (satellite- and verb-framing) in terms of participant group and language context in voluntary motion

In addition, quantitative data further suggested that bilinguals expressed manner more frequently and adopted a predominant satellite-framing strategy compared with Cantonese monolinguals, as illustrated in Figure 11 (Cantonese monolinguals: M=55.37%, SD=12.25%; Bilinguals in a monolingual context: M=71.48%, SD=9.12%;

Bilinguals in a bilingual context: M=75.55%; SD=12.21%; English monolinguals: M=93.70%, SD=9.43%).

Table 21. Coefficients of the logistic mixed-effects model for the frequency satellite-framing in voluntary motion-bilingual speakers

| **Fixed effects** | **Estimate** | **SE** | **Wald z** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 1.539 | 0.315 | 4.889 | < .001 *** |
| **Participant group** | | | | |
| Bilinguals in monolingual context | -0.306 | 0.349 | -0.875 | 0.381 |
| Cantonese | -1.251 | 0.303 | -4.136 | < .001 *** |
| English | 2.156 | 0.334 | 6.344 | < .001 *** |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.723 | 0.850 | | |
| Item (Intercept) | 1.313 | 1.146 | | |

Note: The intercept represents the condition of bilinguals in a bilingual context as a baseline for between-group comparisons.

To further investigate whether there was a statistical difference in framing strategies across different participant group, a mixed-effects model was built. [16] As the proportion of satellite-framing strategy was 1 minus the proportion of verb-framing, the former was used as the binary dependent variable. The fixed effect was group and the random effects were intercepts for participant and item. The results showed that bilinguals in a monolingual context used more satellite-framing strategies than Cantonese monolinguals ($\beta_{\text{Bilinguals in monolingual context-Cantonese}}$ = 0.95, SE = 0.29, Wald z =3.16, p = .003), but patterned with bilinguals in a bilingual context ($\beta_{\text{Bilinguals in monolingual context--Bilinguals in bilingual context}}$ = -0.31, SE = 0.35, Wald z =-0.87, p = 0.38). Meanwhile, bilinguals in a bilingual context used less satellite-framing strategies than English monolinguals ($\beta_{\text{Bilinguals in bilingual context-English monolinguals}}$ = 2.16, SE = 0.34, Wald z =6.34,

$p < .001$), indicating an ongoing cognitive shift towards the target language. Detailed statistical information of framing strategies is presented in Table 21.

#### 8.1.2. Non-linguistic categorization of voluntary motion

##### 8.1.2.1. Categorical preferences of manner- and path-match alternates across Cantonese, English and Cantonese-English bilingual speakers

With regard to participants' manner- or path-match preferences in subsequent categorization (cf. Figure 12), English monolinguals (M=68.89%; SD=19.44%) and bilinguals in different language contexts (M=58.87%; SD=24.52% for monolingual context and M= 60.53%; SD=18.87% for bilingual context) had a manner-match preference compared with Cantonese monolinguals (M=44.44%; SD=20.91%). A mixed-effect logistic model was fitted with participants' manner-match preferences as the binary dependent variable. The fixed effect was participant group. The random effects were crossed-random intercepts for participant and item. Including participant group as a fixed-effect significantly optimized the model ($\chi^2$ (3) =21.55, $p<.001$) compared with the null model, indicating that group was a main effect in participants' similarity judgements. The lack of significance in the overall intercept ($\beta_0$ = 0.64, SE = 0.36, Wald z =1.77, p = 0.07) indicated that participants across language groups did not have an overall preference for either manner- or path-match choices. Then forward difference coding was given to compare the likelihood of manner-match selection in this group with the next group. The results confirmed that bilinguals in a monolingual context preferred more manner-match alternates than Cantonese monolinguals ($\beta_{\text{Cantonese-Bilinguals in monolingual context}}$ = -0.88, SE = 0.34, Wald z =-2.59, p =0.03) but patterned with bilinguals in a bilingual context ($\beta_{\text{Bilinguals in monolingual context-Bilinguals in bilingual context}}$ = -0.07, SE = 0.39, Wald z =-0.19, p = 0.84). Meanwhile, no difference was found between bilinguals in a bilingual context and English monolinguals ($\beta$ Bilinguals in bilinguals in bilingual context-English monolinguals = -0.36, SE = 0.34, Wald z =-1.05, p = 0.60). As expected, Cantonese monolinguals selected less manner-match alternates than English monolinguals ($\beta$ Cantonese-English = -1.31, SE = 0.28, Wald z =-4.67, p < .001). The findings suggested that bilinguals showed a cognitive shift towards L2-based categorization patterns regardless of the language context.

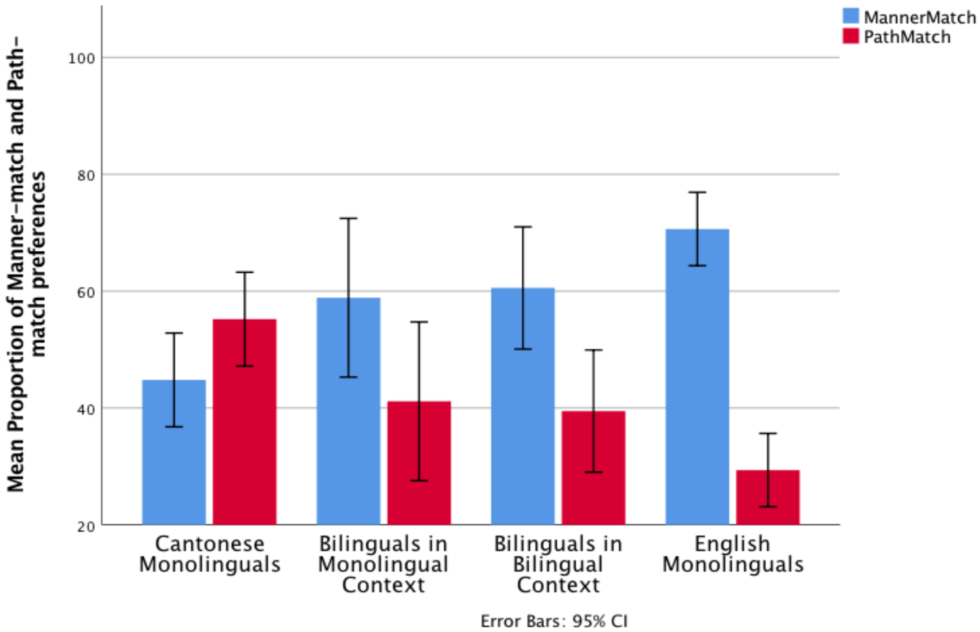



Figure 12. Mean percentage of manner/path preferences in terms of participant group and language context in voluntary motion

A mixed-effect logistic model [17] was fitted to detect whether the observed differences were statistically significant. As the proportion of manner- and path-match preferences was added to 1, the former was used as the dependent variable throughout the statistical analyses. The analysis modelled participants' manner-match preferences as a binary dependent variable. The fixed effect was participant group. The random effects were crossed-random intercepts for participant and item. Including participant group as the fixed-effect significantly optimized the model ($\chi_2$ (3) =21.55, p<.001) compared with the null model, indicating that group was a main effect in participants' similarity judgements.

Table 22. Coefficients of the logistic mixed-effects model for the frequency of manner-match selection in the conceptualization of voluntary motion-bilingual speakers

| **Fixed effects** | **Estimate** | **SE** | **Wald z** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 0.648 | 0.366 | 1.773 | 0.076 |
| **Participant group** | | | | |
| Bilinguals in monolingual context | -0.076 | 0.392 | -0.194 | 0.846 |
| Cantonese | -0.959 | 0.339 | -2.823 | < .001 *** |
| English | 0.360 | 0.342 | 1.054 | 0.292 |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.704 | 0.839 | | |
| Item (Intercept) | 0.681 | 0.825 | | |

Note: The intercept represents the condition of Bilinguals in a bilingual context as a baseline for between-group comparisons.

The lack of significance in the overall intercept ($\beta_0$ = 0.65, SE = 0.36, Wald z =1.77, p = 0.07) indicated that participants across language groups did not have an overall preference for either manner- or path-match alternates. Then forward difference coding was given to compare the likelihood of manner-match selection in this group with the next group. The results confirmed that bilinguals in a monolingual context preferred more manner-match alternates than Cantonese monolinguals ($\beta_{\text{Cantonese-Bilinguals in monolingual context}}$ = -0.88, SE = 0.34, Wald z =-2.59, p =0.03) but patterned with bilinguals in a bilingual context ($\beta_{\text{Bilinguals in monolingual context-Bilinguals in bilingual context}}$ = -0.07, SE = 0.39, Wald z =-0.19, p = 0.84). Meanwhile, no difference was found between bilinguals in a bilingual context and English monolinguals ($\beta_{\text{Bilinguals in bilinguals in bilingual context-English monolinguals}}$ = -0.36, SE = 0.34, Wald z =-1.05, p = 0.61). As expected, Cantonese monolinguals selected less manner-match alternates than English monolinguals ($\beta_{\text{Cantonese-English}}$ = -1.31, SE = 0.28, Wald z =-4.67, p < .001). The findings suggested that bilinguals showed a cognitive shift towards L2-based patterns of categorization regardless of the language context.

#### 8.1.2.2. RT to manner- and path-match alternates across Cantonese, English and Cantonese-English bilinguals

Furthermore, in order to better explore the degree of differences in the decision-making process, participants' RT to manner- and path-match preferences across different language groups was measured as a continuous variable. The mean RT to manner- and path-alternate across each participant group is presented in Table 23.

Table 23. Mean RT (in millisecond) to manner- and path-match alternate across groups in voluntary motion-bilingual speakers

| **Participant Group** | **Mean RT in Manner-match** | **Mean RT in Path-match** |
|---|---|---|
| English | 2187 (735) | 2508 (989) |
| Cantonese | 2393 (956) | 2214 (958) |
| Bilinguals in a monolingual context | 2100 (705) | 2489 (885) |
| Bilinguals in a bilingual context | 2107 (656) | 2251 (773) |

To further explore the statistical significance, a mixed-effects model was built with RT as the continuous dependent variable [18]. The fixed effects were participant group (four levels: Cantonese, bilinguals in a monolingual context and bilinguals in a bilingual context), preference type (two levels: manner- and path-match preference) and their interaction. The random effects included the crossed random intercepts for participant and item. The dependent variable (RT) was log-transformed to meet the assumption of the normality of residuals. Details of the fixed-effect parameters are presented in Table 24.

Table 24. Fixed effects on RT as a function of participant group and preference type in voluntary motion-bilingual speakers

| **Fixed effects** | **Estimate** | **SE** | **t value** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 7.686 | 6.323 | 121.551 | < 2e-16 *** |
| **Preference type** (Manner) | -0.098 | 0.048 | -2.053 | 0.040* |
| **Participant group** | | | | |
| Bilinguals in monolingual context | 0.049 | 0.088 | 0.557 | 0.578 |
| Cantonese | -0.051 | 0.075 | -0.068 | 0.945 |
| English | 0.049 | 0.077 | 0.637 | 0.525 |
| **Preference type: Group** | | | | |
| Manner: monolingual context | -0.027 | 0.068 | -0.402 | 0.688 |
| Manner: Cantonese | -0.012 | 0.051 | -0.058 | 0.041* |
| Manner: English | -0.016 | 0.059 | -0.027 | 0.978 |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.037 | 0.192 | | |
| Item (Intercept) | 0.002 | 0.041 | | |

Note: The intercept represents the log-transformed RT when the preference type is Path-match alternate and participant group is Bilinguals in a bilingual context.

The results suggested that with regard to the fixed effects, there was a main effect of participant group and a participant group by preference type effect. The interaction indicated that participants in each group differed in their processing efficiency when selecting manner- or path-match alternates. In addition, the variance of crossed random intercept and slope provided further evidence that the interaction between preference types and participant group remained consistent and systematic across individual subject and item.

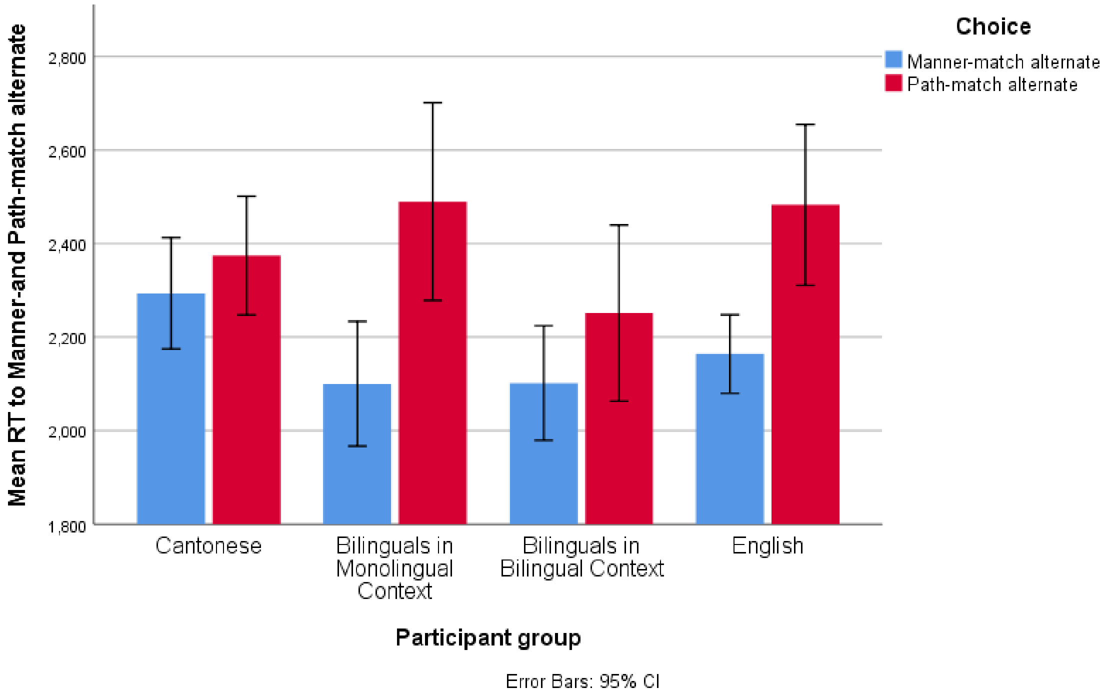



Figure 13. Mean RT to manner- and path-match alternate in terms of language group and language context in voluntary motion

To further address the interaction between participant group and preference types, four separate mixed-effects models [11] [13] [19] [20] were built with log-transformed RT as the dependent variable and preference type as the fixed effect to address the within group difference (four groups: Cantonese, English, Bilinguals in a monolingual and bilingual context). The random effects included crossed random intercepts for participant and item. The intercept for each model set path-match alternate as the benchmark for comparison. For detailed information for the RT of English and Cantonese monolinguals, please refer to the monolingual section.

The results confirmed that for bilinguals in both a bilingual ($\beta_0$ = -0.088, SE = 0.044, t =-1.992, p=0.04) and a monolingual context ($\beta_0$ = -0.123, SE = 0.049, t =-2.506, p = 0.01), their mean RT to manner-match alternate was faster than path-match alternate, as illustrated in Figure 13. The results indicated that first, bilingual speakers patterned with English monolinguals in terms of processing efficiency during the decision-making process, indicating an ongoing process of cognitive restructuring towards the L2-bsaed cognitive patterns. Second, bilingual speakers had the same patterns of RTs regardless of the language context, suggesting that the effect of language learning on cognitive processing was a long process and not mediated by the short-term experimental manipulation.

### 8.1.3. Factors predictive of bilinguals' lexicalization and categorization patterns of voluntary motion

To further explore the modulating effect of other extra-linguistic factors on the degree of cognitive restructuring, mixed-effects models were built with the three target factors: language proficiency, language context (a monolingual vs. a bilingual context) and language contact.

Following Athanasopoulos (2009) and Athanasopoulos et. al (2015), language contact was defined as the amount of use bilingual speakers had with each language and was measured by participants' self-reported scores in doing a series of daily activities, such as watching TV, reading for school and talking with friends etc. (please refer to Appendix I for detailed information). The daily amount of language use (raw data) was converted into percentage scores. As the bilinguals in the current study had no further knowledge about other languages and used Cantonese and English interchangeably in daily interaction, the proportion of English and Cantonese use was added up to 1. Thus, the former (i.e., the amount of English use) was used as the explanatory variable for the statistical analysis. On the whole, Cantonese-English bilinguals used English most of their time (Mean=59.98%, SD=12.80%) in daily activities and therefore, English was the dominant language.

Table 25. Fixed effects on the probability of manner encoding in event lexicalization of voluntary motion

| Fixed effects | Estimate | SE | Wald z | Pr (>\|z\|) |
|---|---|---|---|---|
| Intercept | 0.697 | 1.476 | 0.473 | 0.636 |
| Language Context | -0.936 | 0.581 | -1.611 | 0.107 |
| English proficiency | 5.317 | 1.347 | 3.947 | 0.200 |
| Amount of English use | 6.133 | 2.518 | 2.435 | 0.014 * |

Note: The intercept represents the predicted probability of manner verb encoding (log-odds) when language context is bilingual, the proficiency level is the lowest and the frequency of English use is 0.

First of all, two separate logistic mixed-effects models were built with frequency of manner encoding, and manner-match preferences as separate binary dependent variables. The random effects were random intercepts for participant and item. The main effects were language context, English proficiency and the amount of English use. As the interactions of these three factors were not statistically significant and including the interactions did not significantly optimize the model ($\chi_2$ (3) =2.431, p=0.348 for manner encoding in lexicalization; $\chi_2$ (3) =0.512, p=0.706 for manner-match preferences in categorization), the final models [21] [22] included the main effects (i.e. language context, English proficiency and the amount of English use) only. As shown in Tables 25 and 26, among three different factors under investigation, only language contact surfaced as a significant and unique predictor of bilinguals' performance in both the linguistic descriptions and non-linguistic categorization. That is, the more frequently English was used in participants' daily communication and interaction, the more likely bilingual speakers were to shift from L1-based lexicalization and conceptualization patterns towards L2-based associations.

Table 26. Fixed effects on the probability of manner preferences in categorization of voluntary motion

| Fixed effects | Estimate | SE | Wald z | Pr (>\|z\|) |
|---|---|---|---|---|
| Intercept | -2.452 | 0.857 | -2.864 | 0.004 |
| Language Context | -1.375 | 0.993 | -1.385 | 0.166 |
| English proficiency | 2.215 | 1.215 | 1.823 | 0.360 |
| Amount of English use | 5.291 | 1.341 | 3.947 | <0.001*** |

Note: The intercept represents the predicted probability of manner-match preference (log-odds) when language context is bilingual, the proficiency level is the lowest and the frequency of English use is 0.

In addition, following Ji and Hohenstein (2017, 2018), a multiple linear regression model [23] was fitted with mean differences of RT in manner-match preference minus path-match preference (i.e. the absolute values) as the dependent variable, and the target three factors as explanatory variables. The absolute values of RTs revealed a degree of differences in cognitive restructuring and reflect underlying cognitive mechanisms for motion event processing. To be more specific, positive values of absolute time (i.e. $RT_{manner}$- $RT_{path}$) indicated longer time in making manner-match choices (i.e. a Japanese-way of motion-event processing), while negative values represented longer time in path-match choices (i.e. an English-way of motion-event processing). Thus, if the positive value of absolute time is smaller, it indicates the degree of cognitive restructuring towards the English-based way of processing is greater. This way of conceptualizing RT has been well supported by Ji and Hohenstein (2017, 2018) that given cross-linguistic differences in culture-specific viewpoints and decision-making strategies, speakers of different language may possess different starting points in terms of processing efficiency. Thus, instead of comparing the raw data of RT, absolute values should be used to tackle this cross-linguistic difference. As shown in Table 27, among three different factors under investigation, only language contact surfaced as a significant and unique predictor of bilinguals' processing efficiency in manner- and path-match selection. The results showed that the more frequently English was used in daily communication, the faster participants were to react to manner-match choices in the decision-making process.

Table 27. Multiple regression with mean differences in RT as dependent variable in event categorization of voluntary motion

| **Independent variable** | **Standardized coefficient** | **SE** | **t value** | **p** |
|---|---|---|---|---|
| Intercept | -683.01 | 867.81 | -0.79 | 0.441 |
| Language context | -1302.92 | 477.52 | -2.73 | 0.113 |
| English proficiency | -1204.21 | 396.21 | -3.04 | 0.091 |
| Amount of English use | -1393.36 | 421.20 | -3.31 | 0.004** |

Note: The intercept represents the predicted the mean differences in RT odds) when language context is bilingual, the proficiency level is the lowest and the frequency of English use is 0.

## 8.2. Lexicalization and categorization of caused motion in Cantonese, English and early Cantonese-English bilingual speakers

### 8.2.1. Linguistic encoding of caused motion

#### 8.2.1.1. Frequency of C-Manner and Path encoding in caused motion across Cantonese, English and Cantonese-English bilingual speakers

Altogether 4320 target responses (N=36 for each participant) were included for the final analysis. Participants' linguistic encoding of voluntary notion was first calculated by the frequency of C-manner and path selection. Participants' linguistic responses to each stimulus were transformed into percentages and the mean percentage for each group was compared in terms of participant group. As shown in Figure 14, participants in each group presented a high tendency of path encoding, with a ceiling effect across three groups (English: M=97.31%, SD=3.22%; Cantonese: M=95.83%, SD=5.19%; Bilinguals in a monolingual context: M=97.41%, SD=3.39%; and Bilinguals in a bilingual context: M=96.25 %; SD=4.23%). However, with regard to the C-manner encoding, there was a hierarchical decrease across different language group English: M=98.79%, SD=2.89%; Cantonese: M=86.94%, SD=9.51%; Bilinguals in a monolingual context: M=97.78%, SD=3.35%; and Bilinguals in a bilingual context: M= 97.22%; SD= 3.79%).

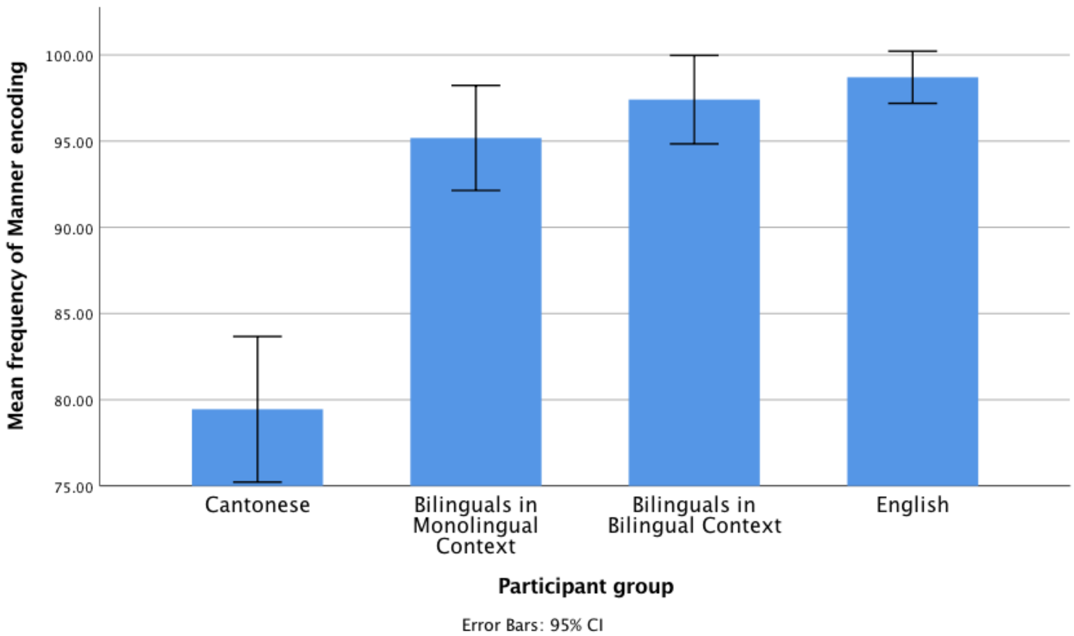



Figure 14. Mean frequency of manner encoding in terms of participant group and language context in caused motion

To assess whether speakers from different groups differed in their likelihood of manner and path encoding, two separate logistic mixed-effect models were built [24] [25]. Within each model, the binary dependent variable was whether the target semantic element (i.e. manner and path of motion) was encoded (code=1) or not (code=0). The fixed effect was participant group. The random effects were random intercepts for participant and item. For path encoding, the results showed that the inclusion of group did not significantly increase the model fit compared with the null model ($\chi^2$ (3) =3.13, p=0.310), indicating that group was not a main effect. In other words, participants across different groups were equally likely to encode path of motion when describing voluntary motion.

However, for the frequency of manner encoding, including participant group as the fixed effect significantly optimized the model compared with the null model ($\chi^2$ (3) =176.18, p<.001), indicating that group was a main effect. Then forward coding was used to compare the likelihood of manner encoding with the next group. As shown in Figure 14, bilinguals in a monolingual context encoded more manner than Cantonese monolinguals ($\beta_{\text{Cantonese-Bilinguals in monolingual context}}$ = -1.97, SE = 0.31, Wald z =-6.33, p < .001) but patterned with bilinguals in a bilingual context ($\beta_{\text{Bilinguals in monolingual context--Bilinguals in bilingual context}}$ = 0.31, SE = 0.41, Wald z =0.74, p = 0.46). Meanwhile, no difference between bilinguals in a bilingual context and English monolinguals was detected ($\beta_{\text{Bilinguals in bilinguals in bilingual context-English monolinguals}}$ = -0.89, SE = 0.39, Wald z =-0.32, p = 0.07). The findings indicated that bilinguals demonstrated a cognitive shift towards L2-based encoding patterns regardless of the language context in the lexicalization of caused motion.

Table 28. Coefficients of the logistic mixed-effects model for the frequency of manner encoding in caused motion-bilingual speakers

| **Fixed effects** | **Estimate** | **SE** | **Wald z** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 4.057 | 0.351 | 11.567 | < .001 *** |
| **Participant group** | | | | |
| Bilinguals in monolingual context | 0.308 | 0.414 | 0.744 | 0.457 |
| Cantonese | -1.742 | 0.284 | -6.121 | < .001 *** |
| English | 0.909 | 0.381 | 2.387 | 0.017* |
| **Random effects** | **Variance** | **Standard deviations** | | |

| | | |
|---|---|---|
| Subject (Intercept) | 0.542 | 0.736 |
| Item (Intercept) | 0.789 | 0.888 |

Note: The intercept represents the condition of Bilinguals in a bilingual context as a baseline for between-group comparisons.

#### 8.2.1.2. Semantic distribution of C-Manner and Path across Cantonese, English and Cantonese-English bilingual speakers

Based on the frequency of manner and path encoding, the semantic distribution of C-manner and path are presented in Table 29. The results suggested that semantic distribution of manner and path in bilinguals in different language contexts showed a clear shift from L1-Cantonese towards an L2 English-specific way in describing voluntary motion. For example, bilinguals in a monolingual and bilingual context were more prone to encode manner of motion in the main verb form (M=69.62%, SD=8.52% for bilinguals in a monolingual context and M=76.30%, SD= 5.89% for bilinguals in a bilingual context), with more details presented in Table 29.

Table 29. Percentages of semantic distribution of C-manner/path in verb (V) or outside of it (OHT) in caused motion.

| | **Cantonese (%)** | | | **Bilinguals in monolingual context (%)** | | | **Bilinguals in bilingual context (%)** | | | **English (%)** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Component | V | OTH | Total | V | OTH | Total | V | OTH | Total | V | OTH | Total |
| C-Manner | 48.33 | 38.05. | 86.94 | 69.26 | 32.59 | 95.19 | 76.30 | 24.44 | 97.41 | 98.06 | 1.85 | 98.79 |
| Path | 42.77 | 54.44 | 95.83 | 43.33 | 60.00 | 95.19 | 32.22 | 62.86 | 92.96 | 1.76 | 95.93 | 97.31 |

Notes: The sum of the first two columns within each language group doesn't always add up to the total proportion as the manner or path of motion can be double-encoded in V and OTH at the same time (e.g.: The boy is [jumping]$_{verb}$ downstairs [with one leg] $_{OTH}$).

#### 8.2.1.3. Construction types and framing strategies of caused motion event across Cantonese, English and Cantonese-English bilingual speakers

A qualitative analysis was conducted with regard to the construction types and framing strategies in each language group. The linguistic data of bilinguals in both a monolingual and a bilingual context exhibited a clear shift from L1- to L2-based patterns in both construction types and framing strategies, that is, bilinguals attempted to more 'C-Manner verb +Path satellite' constructions (satellite-framing strategy) and fewer 'Path only' constructions (verb-framing strategy) compared with Cantonese monolinguals in L1 descriptions. Specific examples are given in (69) and (70):

(69) Verb-framing strategy (type 1): Path verb only

| | | |
|---|---|---|
| Go3 naam4 zai2 | **jap6-zo2** | jat1 go1 bo1 |
| A boy | **entered-ASP** | a ball |

'A boy made a ball enter the cave.'

(70) Satellite-framing strategy: Manner verb + Path satellite

| | | | |
|---|---|---|---|
| Go3 naam4 zai2 | **tek3-zo2** | jap4 | jat1 go1 bo1 |
| A boy | **kick-ASP** | into | a ball |

'A boy kicked a ball into the cave.'

To further explore whether there was a statistical significance in the framing strategies adopted by each language group, a mixed-effect logistic model [26] was fitted. The final model was built with the presence or absence of satellite-framing as the binary dependent variable. The fixed effect was participant group. As adding the random intercept for item caused the model failed to converge, the random effects included random intercept for participant only. Similar to voluntary motion, the results of the framing strategies in caused motion indicated that bilinguals in a monolingual context used more satellite-framing strategies than Cantonese monolinguals ($\beta$ Bilinguals in monolingual context-Cantonese = 0.70, SE = 0.17, Wald z =4.16, p <.001) but patterned with bilinguals in a bilingual context (β Bilinguals in monolingual context--Bilinguals in bilingual context = -0.27, SE = 0.23, Wald z =-1.19, p = 0.24). Meanwhile, bilinguals in a bilingual context used less satellite-framing strategies than English monolinguals (β Bilinguals in bilingual context-English monolinguals = -4.09, SE = 0.29, Wald z =-14.03, p < .001), indicating an ongoing process of conceptualization restructuring towards L2-based patterns in event lexicalization. Detailed information is presented below.

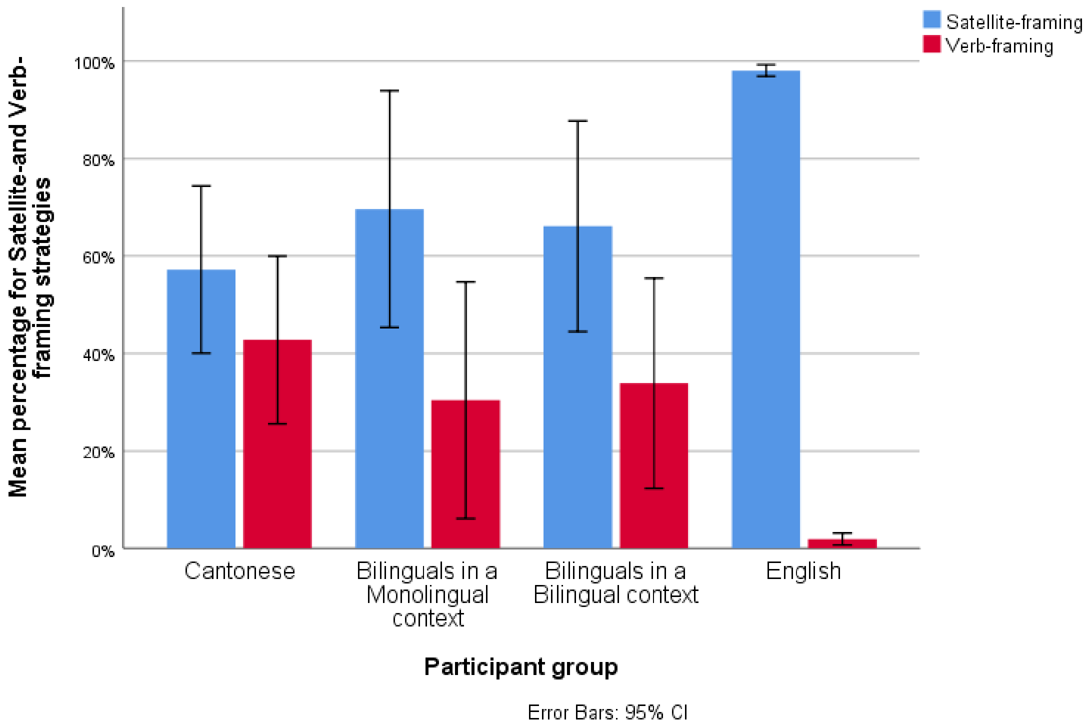



Figure 15. Mean percentage of different framing tendencies (satellite- and verb-framing) in terms of participant group and language context in caused motion

Table 30. Coefficients of the logistic mixed-effects model for the frequency of satellite-framing in caused motion-bilingual speakers

| **Fixed effects** | **Estimate** | **SE** | **Wald z** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 1.990 | 0.558 | 3.568 | < .001 *** |
| **Participant group** | | | | |
| Bilinguals in monolingual context | -0.274 | 0.230 | -1.188 | 0.235 |
| Cantonese | -0.975 | 0.163 | -5.967 | < .001 *** |
| English | 4.089 | 0.292 | 14.027 | < .001 *** |

| Random effects | Variance | Standard deviations |
|---|---|---|
| Subject (Intercept) | 8.199 | 2.863 |

Note: The intercept represents the condition of bilinguals in a bilingual context as a baseline for between-group comparisons.

### 8.2.2. Non-linguistic categorization of caused motion

#### 8.2.2.1. Categorical preferences of manner- and path-match alternates across Cantonese, English and Cantonese-English bilingual speakers

In the conceptualization of caused motion, participants in each group presented an overall preference for path-match alternate: Cantonese monolinguals (M=63.89%, SD=28.85%), bilinguals in a bilingual context (M=62.22%, SD=28.62%), bilinguals in a monolingual context (M=60.11%, SD=38.86%), and English monolinguals (M=65.56%; SD=30.93%), as shown in Figure 16.

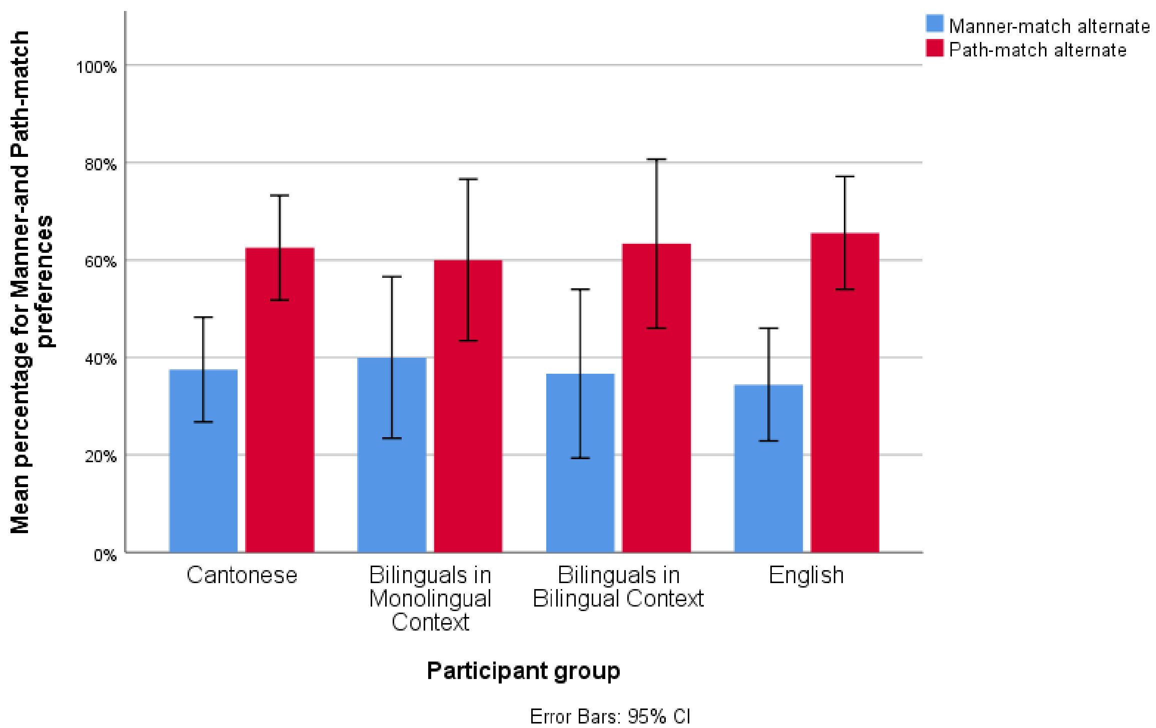



Figure 16. Mean percentage of manner/path preferences in terms of participant group and language context in caused motion

A mixed-effects logistic model [27] was fitted to detect whether the observed differences were statistically significant. The final analysis modelled participants' manner-match preferences as a binary dependent variable. The fixed effect was participant group. The random effects were crossed-random intercepts for participant and item. Including participant group as the fixed-effect significantly optimized the model ($\chi^2$ (3) =1.52, p=0.677) compared with the null model, indicating that group was a main effect in participants' similarity judgements. The significance in the overall intercept ($\beta_0$ = -0.59, SE = 0.22, Wald z =-2.71, p = 0.006) indicated that participants across language groups had an overall preference for path-match alternates. In other words, participants across different language groups were more prone to use path-match as their selection criteria when categorizing caused motion events.

Table 31. Coefficients of the logistic mixed-effects model for the frequency of manner selection in the conceptualization of caused motion-bilingual speakers

| Fixed effects | Estimate | SE | t value | Pr (>\|t\|) |
|---|---|---|---|---|
| **Intercept** | -0.592 | 0.218 | -2.716 | 0.006 ** |
| **Participant group** | | | | |
| Bilinguals in monolingual context | 0.019 | 0.211 | 0.091 | 0.927 |
| Cantonese | 0.062 | 0.213 | 0.293 | 0.769 |
| English | -0.152 | 0.165 | -0.919 | 0.358 |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.526 | 0.725 | | |
| Item (Intercept) | 0.196 | 0.648 | | |

Note: The intercept represents the condition of Bilinguals in a bilingual context as a baseline for between-group comparisons.

#### 8.2.2.2. RT to manner- and path-match alternates across Cantonese, English and Cantonese-English bilinguals

Furthermore, in order to better explore the degree of differences in the decision-making process, participants' RT to manner- and path-match preferences was measured as a continuous variable and used to indicate participants' efficiency in cognitive processing. The mean RT to manner- and path-alternate across each participant group is presented in Table 32.

Table 32. Mean RT (in millisecond) to manner- and path-match alternate across groups in caused motion-bilingual speakers

| Participant Group | Mean RT in Manner-match | Mean RT in Path-match |
|---|---|---|
| English | 2048 (SD=683) | 2416 (SD=840) |
| Cantonese | 2165 (SD=810) | 2247 (SD=792) |
| Bilinguals in a monolingual context | 2310 (SD=601) | 2624 (SD=920) |
| Bilinguals in a bilingual context | 2240 (SD=554) | 2489 (SD=766) |

To further explore the statistical significance, a mixed-effects model [28] was built with RT as the continuous dependent variable. The fixed effects included participant group (four levels: Cantonese, bilinguals in a bilingual context, bilinguals in a monolingual context and English monolinguals), preference type (two levels: manner- and path-match preference) and their interaction. The random effects included the crossed random intercepts for participant and item. The dependent variable (RT) was log-transformed in order to meet the assumption of the normality of residuals. Details of fixed-effect parameters are presented in Table 33.

The results suggested that with regard to the fixed effects, there was a main effect of participant group and a participant group by preference type interaction effect. This indicated that participants in each group differed in their processing efficiency when selecting manner- or path-match alternates. In addition, the variance of crossed random intercept and slope provided further evidence that the interaction between preference types and participant group remains consistent and systematic across individual subject and item.

Table 33. Fixed effects on RT as a function of participant group and preference type in caused motion-bilingual speakers

| Fixed effects | Estimate | SE | t value | Pr (>\|t\|) |
|---|---|---|---|---|
| **Intercept** | 7.634 | 0.040 | 188.711 | < 2e-16 *** |
| **Preference type** (Manner) | 0.014 | 0.037 | 0.372 | 0.710 |
| **Participant group** | | | | |
| Bilinguals in monolingual context | 0.155 | 0.070 | 2.207 | 0.029 * |
| Cantonese | 0.185 | 0.070 | 2.629 | 0.009 ** |
| English | 0.083 | 0.057 | 1.461 | 0.146 |
| **Preference type: Group** | | | | |
| Manner: monolingual context | -0.148 | 0.064 | -2.300 | 0.021 * |
| Manner: Cantonese | -0.143 | 0.064 | -2.225 | 0.026 * |
| Manner: English | -0.142 | 0.053 | -2.659 | 0.008 ** |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.036 | 0.191 | | |
| Item (Intercept) | 0.001 | 0.021 | | |

Note: The intercept represents the log-transformed RT when the preference type is Path-match alternate and participant group is Bilinguals in a bilingual context.

To further address the interaction between participant group and preference types, four separate mixed-effects [29] [30] models were built with log-transformed RT as the dependent variable and preference type as the fixed effect to address the within group difference (for the RT for Cantonese and English monolinguals, please refer to the monolingual section). Random effects included crossed random intercepts for participant and item. The intercept for each model set path-match alternate as the benchmark for comparison.

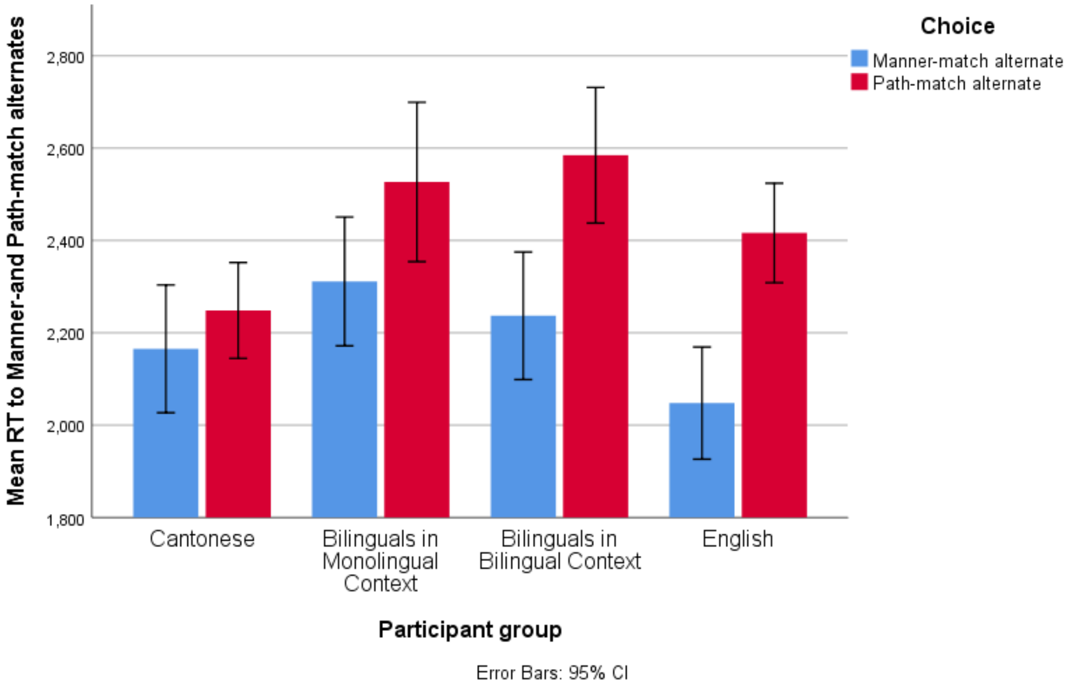



Figure 17. Mean RT to manner- and path-match alternate in terms of participant group and language context in caused motion

In line with voluntary motion, the results of caused motion further confirmed that for bilinguals in both a bilingual ($\beta_0$ = -0.11, SE = 0.047, t =-2.265, p=0.02) and a monolingual context ($\beta_0$ = -0.132, SE = 0.056, t =-2.401, p = 0.01), their RTs to manner-match alternate was much faster than path-match alternate, as shown in Figure 17. The results indicated an ongoing process of cognitive restructuring from L1-based processing patterns towards L2-based ones. And the restructuring process was not subject to short-term experimental manipulation.

### 8.2.3. Factors modulating the degree of cognitive restructuring in bilinguals' lexicalization and categorization of caused motion

Like voluntary motion, this section further explored how the three target factors (i.e. language context, the amount of English use and English proficiency) tended to modulate the degree of cognitive restructuring within the bilingual mind.

As indicated by bilinguals' performance in overt selection, participants across different language groups had an overall preference for path-match alternate. Thus, this factor was not computed as a dependent variable for the current analysis. With regard to the linguistic encoding, a logistic mixed-effects model [31] was built with frequency of manner encoding as the binary dependent variable. The random effects were random intercepts for participant and item. The main effects were language context, English proficiency and the amount of English use. As the interactions of these three factors were not statistically significant and including the interactions did not significantly optimize the model ($\chi^2(1) = 2.036$, $p = 0.758$ for manner encoding in lexicalization), the final model included the main effects (i.e. language context, English proficiency and language contact) only. As shown in Table 34, among the three different factors under investigation, only language contact surfaced as a significant and unique predictor of bilinguals' lexicalization of manner in the linguistic descriptions. That is, the more frequently English was used in participants' daily communication and interactions, the more likely bilingual speakers were to shift from the L1-based patterns of lexicalization and conceptualization towards the L2-based ones.

Table 34. Fixed effects on the probability of manner encoding in lexicalization of caused motion

| **Fixed effects** | **Estimate** | **SE** | **Wald z** | **Pr(>\|z\|)** |
|---|---|---|---|---|
| Intercept | 0.981 | 2.346 | 0.418 | 0.987 |
| Language Context | -0.334 | 0.876 | -0.381 | 0.812 |
| English proficiency | -2.315 | 1.205 | -1.922 | 0.101 |
| Amount of English use | 7.231 | 1.926 | 3.754 | <.001*** |

Note: The intercept represents the predicted probability of manner verb encoding (log-odds) when language context is bilingual, the proficiency level is the lowest and the frequency of English use is 0.

Then a multiple linear regression [32] was used with mean differences of RT in manner-match preference minus path-match preference as dependent variable, and the target three factors as explanatory variables. Positive values of RT indicated longer response latency in making manner-match choices whereas negative values represented longer time in path-match choices. As shown in Table 35, among three different factors under investigation, only language contact surfaced as a significant and unique predictor of bilinguals' processing efficiency. The results showed that the more frequently English was used in daily communication, the faster participants were to react to manner-match choices in the response latency of caused motion.

Table 35. Multiple regression with mean differences in RT the as dependent variable in the categorization of caused motion

| **Independent variable** | **Standardized coefficient** | **SE** | **t value** | **p** |
|---|---|---|---|---|
| Intercept | -876.01 | 617.12 | -1.419 | 0.201 |
| Language context | -1219.36 | 296.23 | -4.12 | 0.098 |
| English proficiency | 1013.34 | 517.13 | 1.96 | 0.104 |
| Amount of English use | -958.91 | 203.21 | -4.718 | .003** |

Note: The intercept represents the predicted the mean differences in RT odds) when language context is bilingual, the proficiency level is the lowest and the frequency of English use is 0.

# Chapter 9. Experimental results of motion event lexicalization and categorization in Cantonese-English-Japanese multilingual speakers

The results of monolingual and bilingual speakers in Chapters 7 and 8 suggest that monolingual controls of Cantonese, English and Japanese display language-specific patterns in both event lexicalization and conceptualization. In addition, the bilingual data further confirms that the conceptual change (cognitive restructuring) within the bilingual mind is a long-term process and not subject to short-term experimental manipulation (i.e. language context).

Given the preliminary findings, this section further explores how Cantonese-English-Japanese multilinguals in different language contexts lexicalize and conceptualize motion events compared with the monolingual and bilingual controls of each language. Specifically, this section addresses how multilinguals in a monolingual (L1) and a bilingual (L1 and L3) context tend to behave as a function of recent L3 activation. In addition, it also further explores how the other extra-linguistic factors (i.e., language contact, and language proficiency in the L2 and L3) modulate the degree of cognitive restructuring within the multilingual mind.

## 9.1. Lexicalization and categorization of voluntary motion in monolingual, bilingual and Cantonese-English-Japanese multilingual speakers

### 9.1.1. Linguistic encoding of voluntary motion

#### 9.1.1.1. Frequency of Manner and Path encoding across monolinguals, bilinguals and Cantonese-English-Japanese multilinguals

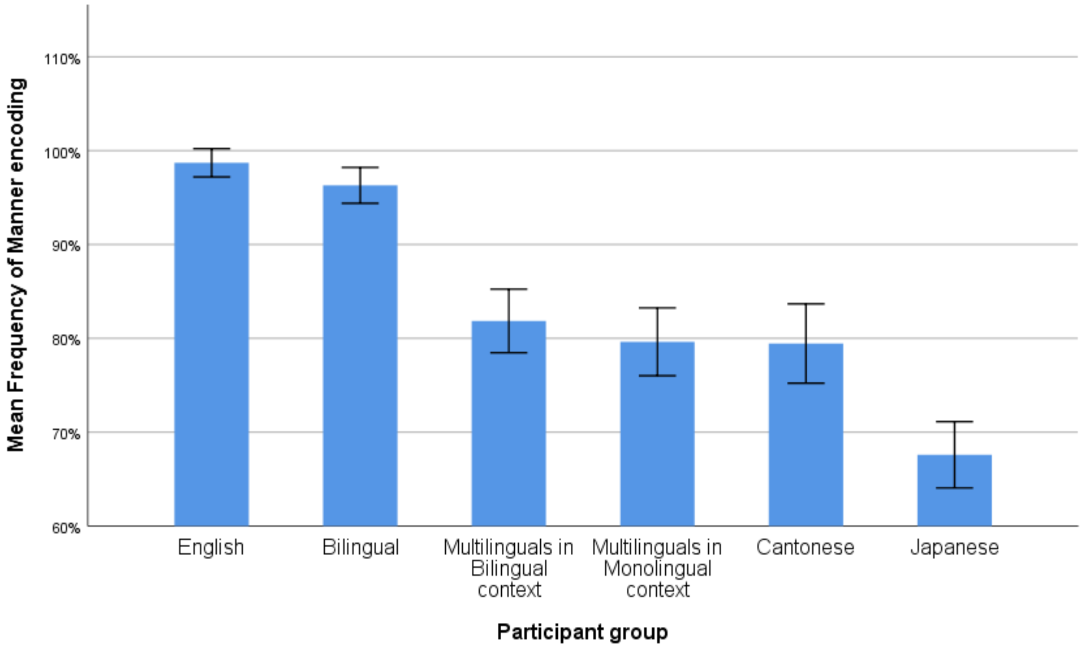



Figure 18. Mean percentage of manner encoding in terms of participant group and language context in voluntary motion

A total of 5400 target responses were included for the final analysis. Participants' linguistic encodings were calculated by the frequency of manner and path selection. Participants' linguistic responses to each stimulus were transformed into percentages and the mean percentage for each group was compared in terms of the participant group and language context. Given that bilinguals in a monolingual and a bilingual context exhibited the same patterns for event lexicalization and conceptualization, I computed the bilingual group as a whole in this analysis. Results showed that participants in each group presented a high tendency of path encoding, with a ceiling effect across four groups. However, with regard to manner encoding, multilinguals encoded manner more often than Japanese monolinguals yet patterned with Cantonese monolinguals (Bilinguals: M=96.30%, SD=5.12%; Multilinguals in a bilingual context: M=83.18%, SD=8.88%; Multilinguals in a monolingual context:

M=81.41%, SD=9.26%, and Cantonese: M=79.07%, SD=11.32%), as illustrated in Figure 18 (See Table 9 for encoding patterns of other language groups).

To assess whether speakers from different groups differed in their likelihood of manner and path encoding, two separate logistic mixed-effect models [33] [34] were built with the *lme4* package. Within each model, the binary dependent variable was whether the target semantic element (e.g. manner and path of motion) was encoded (code=1) or not (code=0). The fixed effect was participant group (five levels: English monolinguals (N=30) *vs*. Bilinguals (N=30) vs. Multilinguals in a monolingual context (N=15) vs. Multilinguals in a bilingual context (N=15) vs. Japanese monolinguals (N=30)). The random effects were random intercepts for participant and item. For path encoding, the results showed that the inclusion of group did not significantly increase the model fit compared with the null model ($\chi^2$ (5) =9.63, p=0.185), showing that group was not a main effect. In other words, participants across different groups were equally likely to encode path of motion when describing voluntary motion.

However, for the frequency of manner encoding, including participant group as the fixed effect significantly optimized the model compared with the null model ($\chi^2$ (5) =89.29, p<.001), indicating that group was a main effect. Then forward coding was used to compare the likelihood of manner encoding with the next group. As shown in Figure 18, multilinguals in a bilingual context encoded manner less often than bilingual speakers ($\beta_{\text{Bilinguals-Multilinguals in bilingual context}}$ = -2.51, SE = 0.36, Wald z =-6.84, p < .001) but patterned with multilinguals in a monolingual context ($\beta_{\text{Multilinguals in monolingual context--Multilinguals in bilingual context}}$ = -0.96, SE = 0.71, Wald z =-1.36, p = 0.18). Meanwhile, no difference was detected between multilinguals in a monolingual context and Cantonese monolinguals ($\beta_{\text{Multilinguals in bilinguals in monolingual context-Cantonese monolinguals}}$ = -1.07, SE = 0.71, Wald z =-1.51, p = 0.13). Detailed statistical information is presented in

Table 36. The findings indicated that multilinguals demonstrated a cognitive shift from the L1- and L2-based encoding patterns towards the L3-based ones regardless of the language context. That is, the acquisition of an L2 (S-language) and an L3 (V-language) of contrastive linguistic features seemed to counterbalance the potential effects of language learning and restored the original lexicalization patterns established in the L1.

Table 36. Coefficients for the logistic mixed-effects model for the frequency of manner encoding in voluntary motion-mutilingual speakers

| **Fixed effects** | **Estimate** | **SE** | **Wald z** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 4.519 | 0.377 | 11.971 | <.001*** |
| **Participant group** | | | | |
| Cantonese | -2.501 | 0.313 | -7.975 | <.001*** |
| English | 1.154 | 0.424 | 2.721 | 0.769 |
| Japanese | -3.269 | 0.294 | -11.109 | <.001*** |
| Multilingual B | -2.336 | 0.368 | -6.338 | <.001*** |
| Multilngual M | -2.515 | 0.367 | -6.846 | <.001*** |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.692 | 0.831 | | |
| Item (Intercept) | 0.196 | 0.648 | | |

Note: The intercept represents the condition of Bilinguals as a baseline for between-group comparisons.

#### 9.1.1.2. Semantic distribution of Manner and Path across monolinguals, bilinguals and Cantonese-English-Japanese multilinguals

Based on the frequency of manner and path encoding, I further explored their semantic distributions within each utterance, as illustrated in Table 37. The results indicated that the semantic distribution in multilingual speakers showed a clear effect of L3 learning. For example, multilinguals in a monolingual and bilingual context used less manner verbs but more path verbs compared with bilinguals yet patterned with Cantonese monolinguals, indicating that the learning of an L3-Japanese (V-language) seemed to counterbalance the potential effects of L2-English (V-language) on the L1 in the semantic distribution of manner and path of motion. Detailed information on the linguistic encoding of manner and path is presented in Table 37.

Table 37. Percentages of semantic distribution of manner/path in verb (V) or outside of it (OHT) in voluntary motion.

| | **Cantonese (%)** | | | **Multilinguals in monolingual context (%)** | | | **Multilinguals in bilingual context (%)** | | | **Bilinguals (%)** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Component | V | OTH | Total | V | OTH | Total | V | OTH | Total | V | OTH | Total |
| Manner | 55.56 | 29.26 | 79.07 | 57.24 | 32.59 | 83.41 | 60.30 | 24.44 | 85.18 | 72.78 | 28.51 | 96.30 |
| Path | 45.74 | 57.71 | 93.70 | 54.33 | 60.00 | 95.12 | 49. 22 | 42.86 | 93.01 | 37.78 | 61.43 | 96.83 |

Notes: The sum of the first two columns within each language group doesn't always add up to the total proportion as the manner or path of motion can be double-encoded in V and OTH at the same time (e.g.: The boy is [jumping]$_{verb}$ downstairs [with one leg] $_{OTH}$).

#### 9.1.1.3. Construction type and framing strategies of voluntary motion event across monolinguals, bilinguals and Cantonese-English-Japanese multilinguals

A qualitative analysis was conducted with regard to construction types and framing strategies of each language group. As reviewed in the monolingual section, the typological status of Cantonese, English and Japanese exhibited language-specific characteristics (cf. Figure 3). With regard to the linguistic encoding of multilingual speakers in different language contexts, they demonstrated an ongoing shift from the L1- and L2-based patterns towards the L3-based ones in terms of the semantic distribution of manner and path, as well as the framing strategies. A qualitative analysis showed that multilinguals tended to use more 'Path only' constructions (verb-framing strategy) and to omit manner of motion more frequently compared with bilingual speakers as a principle of processing costs and benefits. For example, multilingual speakers tended to mention manner of motion less often compared with bilingual speakers but patterned with Cantonese monolinguals when describing the same motion scene, suggesting an ongoing effect of L3 learning. Specific examples of Cantonese are given in (71) and (72).

(71) **Satellite-framing strategy: Manner verb + Path satellite**

| | | | |
|---|---|---|---|
| Go3 naam4 zai2 | haang6-zo2 | jap6 | gaan1 fong2 |
| A boy | walk-ASP | into | the room |

'A boy walked into the room.'

(72) **Verb-framing strategy: Path verb only**

| | | | |
|---|---|---|---|
| Go3 naam4 zai2 | jap6 | zo2 | gaan1 fong2 |
| A boy | entered | ASP | the room |

'A boy entered the room.'

Quantitative data further confirmed that multilinguals in different language contexts adopted less satellite-framing strategy compared with their bilingual counterparts yet displayed L1-based patterns, as illustrated in Figure 19 (Bilinguals: M=67.23%, SD=8.79%; Cantonese monolinguals: M=55.37%, SD=12.25%; Multilinguals in a monolingual context: M=56.74%, SD=7.85%; Multilinguals in a bilingual context: M=58.28%; SD=8.35%). A mixed-effects logistic model [35] was fitted with the presence or absence of satellite-framing as the binary dependent variable. The fixed effect was group and the random effects were intercepts for participant and item. The results showed that multilinguals in a monolingual context used less satellite-framing strategies than bilinguals ($\beta_{\text{Multilinguals in bilingual context-Bilinguals}} = -0.92$, SE = 0.25, Wald z =-3.27, p<.001; $\beta_{\text{Multilinguals in monolingual context-Bilinguals}} = -1.12$, SE = 0.27, Wald z =-3.76, p =.001), but patterned with Cantonese monolingual regardless of the language context ($\beta_{\text{Multilinguals in monolingual context-Cantonese}} = -0.31$, SE = 0.35, Wald z =-0.87, p = 0.38).

Meanwhile, Japanese monolingual still demonstrated the lowest frequency of satellite-framing ($\beta$ Multilinguals in bilingual context-Japanese monolinguals = -2.16, SE = 0.34, Wald z =-6.34, p < .001). The results indicate an effect of L3 learning on motion event lexicalization in multilinguals with two typologically different languages. Detailed statistical information of framing strategies is presented in Table 38.

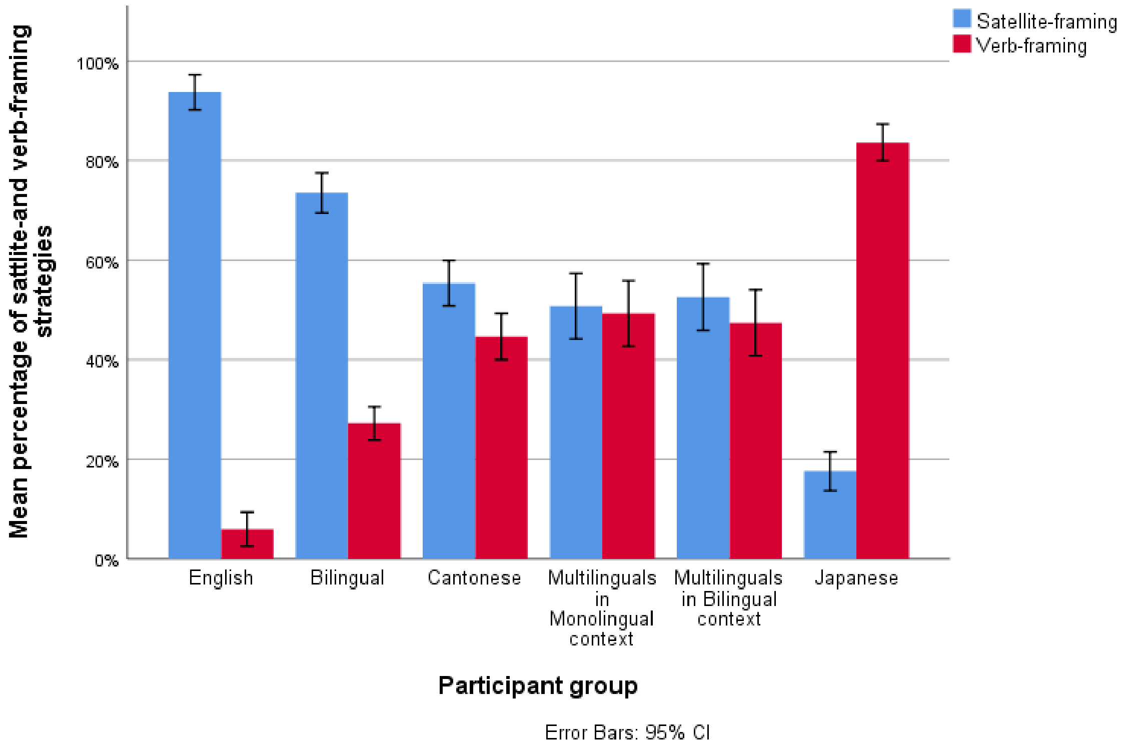



Figure 19. Mean percentage of different framing tendencies (satellite- and verb-framing) in terms of participant group and language context in voluntary motion

Table 38. Coefficients of the logistic mixed-effects model of satellite- and verb-framing strategy in voluntary motion-multilingual speakers

| **Fixed effects** | **Estimate** | **SE** | **Wald z** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 1.275 | 0.211 | 6.043 | <.001*** |
| **Participant group** | | | | |
| Cantonese | -1.006 | 0.198 | -5.063 | <.001*** |
| English | 2.091 | 0.239 | 8.731 | 0.769 |
| Japanese | -3.192 | 0.210 | -15.196 | <.001*** |
| Multilingual B | -0.804 | 0.246 | -3.270 | .0017** |

| | | | | |
|---|---|---|---|---|
| Multilngual M | -0.924 | 0.245 | -3.764 | .001*** |

| Random effects | Variance | Standard deviations |
|---|---|---|
| Subject (Intercept) | 0.432 | 0.656 |
| Item (Intercept) | 0.863 | 0.929 |

Note: The intercept represents the condition of bilinguals as a baseline for between-group comparisons.

## 9.1.2. Non-linguistic categorization of voluntary motion

### 9.1.2.1. Categorical preferences of manner- and path-match alternates across monolinguals, bilinguals and Cantonese-English-Japanese multilinguals

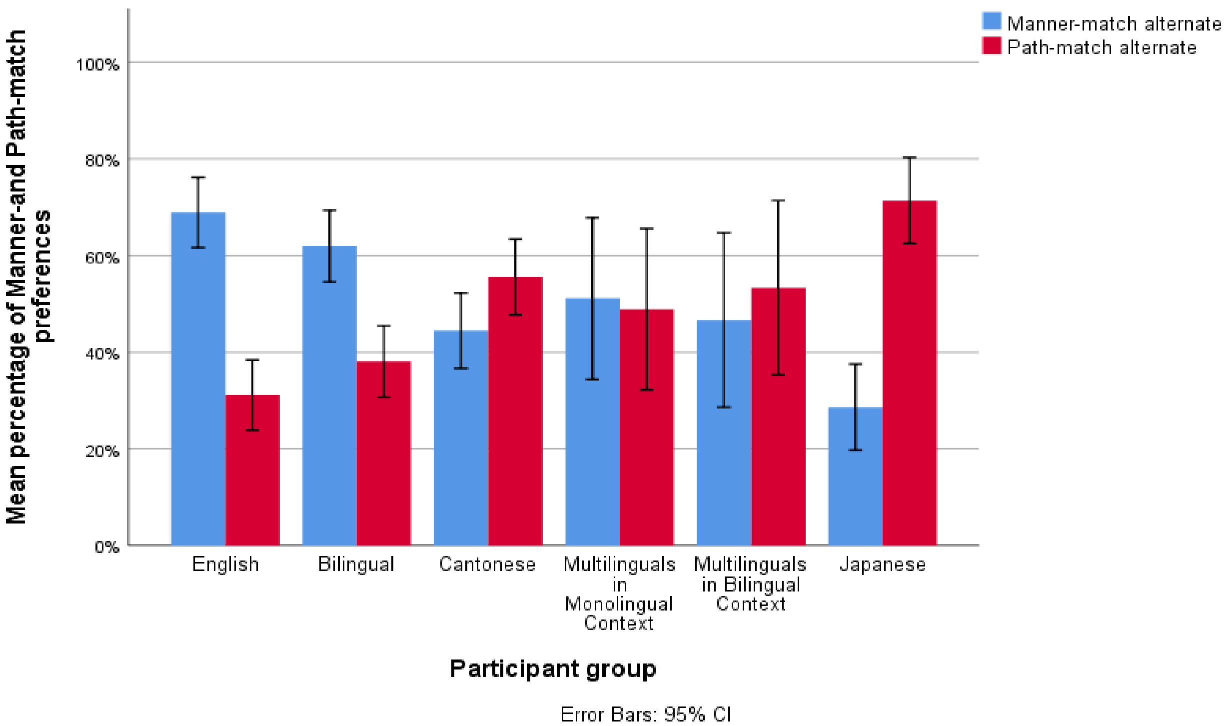



Figure 20. Mean percentage of manner/path preferences in terms of participant group and language context in voluntary motion

Regarding manner- or path-match preferences in subsequent categorization, English monolinguals (M=68.89%; SD=19.44%) and bilinguals (M=59.87%; SD=24.52%) had a manner-match preference compared with Cantonese monolinguals (M=44.44%; SD=20.49%), while multilinguals presented a similar pattern with Cantonese monolinguals regardless of the language context (M=51.11%; SD=30.19% for multilinguals in a monolingual context and M=46.67%; SD=32.55% for multilinguals in a bilingual context). Japanese monolinguals demonstrated the lowest level of manner-match preference (M=28.61%; SD=23.84%), as shown in Figure 20.

A mixed-effects logistic model [36] was fitted to detect whether the observed differences were statistically significant. The analysis modelled participants' manner-match preferences as a binary dependent variable. The fixed effect was participant group. The random effects were crossed-random intercepts for participant and item. Including participant group as the fixed-effect significantly optimized the model ($\chi^2$ (5) =21.55, p<.001) compared with the null model, indicating that group is a main effect in participants' similarity judgements. The lack of significance in the overall intercept ($\beta_0$ = 0.62, SE = 0.30, Wald z =1.77, p = 0.07) indicated that participants across language groups did not have an overall preference for either manner- or path-match alternates. Then forward difference coding was given to compare the likelihood of manner-match selection in this group with the next. The results confirmed that multilinguals in both a monolingual and bilingual context patterned with Cantonese monolinguals in the selection of manner-match preferences. Meanwhile, no difference was found between multilinguals in a monolingual and bilingual context ($\beta_{\text{Multilinguals in bilingual in bilingual context-Multilinguals in monolingual context}}$ = -0.36, SE = 0.34, Wald z =-1.05, p = 0.61). The findings suggested an effect of L3 learning on L1- and L2-based patterns in event categorization. In addition, the conceptual changes in the multilingual mind were not subject to short-term task manipulation of language context. Detailed information of the coefficients of each parameter is presented in Table 39.

Table 39. Coefficients of the logistic mixed-effects model for the frequency of manner-match selection in the conceptualization of voluntary motion-multilingual speakers

| Fixed effects | Estimate | SE | Wald \|z\| | Pr (>\|t\|) |
|---|---|---|---|---|
| **Intercept** | 0.617 | 0.302 | 2.038 | 0.042 * |
| **Participant group** | | | | |
| Cantonese | -0.929 | 0.328 | -2.829 | 0.004 ** |
| English | 0.402 | 0.332 | 1.213 | 0.225 |
| Japanese | -1.819 | 0.336 | -5.411 | <.001 |
| Multilingual B | -0.353 | 0.408 | -0.866 | 0.387 |
| Multilngual M | -1.007 | 0.405 | -2.487 | 0.01* |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 1.167 | 1.081 | | |
| Item (Intercept) | 0.448 | 0.669 | | |

Note: The intercept represents the condition of bilinguals a baseline for between-group comparisons.

#### 9.1.2.2. RT to manner- and path-match alternates across monolinguals, bilinguals and Cantonese-English-Japanese multilinguals

Furthermore, to better explore the degree of differences in the decision-making process, participants' RTs to manner- and path-match preferences across different language groups was measured as a continuous variable. Their mean RT to manner- and path-alternate across each participant group is presented in Table 40.

To further explore the statistical significance, a mixed-effects model [37] was built with RT as the continuous dependent variable. The fixed effects included participant group (six levels: Cantonese, English, Japanese, Bilinguals, multilinguals in a monolingual context and in a bilingual context), preference type (two levels: manner- and path-match preference) and their interaction. The random effects included the crossed random intercepts for participant and item. The dependent variable (RT) was log-transformed to meet the assumption of the normality of residuals. Details of fixed-effect parameters are presented in Table 41.

Table 40. Mean RT (in millisecond) to manner- and path-match alternate across groups in voluntary motion-multilingual speakers

| **Participant Group** | **Mean RT in Manner-match** | **Mean RT in Path-match** |
|---|---|---|
| English | 2187 (735) | 2508 (989) |
| Cantonese | 2393 (956) | 2214 (958) |
| Japanese | 2349 (1207) | 2156 (853) |
| Bilinguals | 2100 (679) | 2373 (837) |
| Multilinguals in monolingual context | 2106 (740) | 2284 (794) |
| Multilinguals in bilingual context | 2271 (634) | 2231 (715) |

The results suggested a main effect of participant group and a participant group by interaction effect. The interaction indicated that participants in each group differed in their processing efficiency when selecting manner- or path-match alternates. In addition, the variance of crossed random intercept and slope provided further evidence that the interaction between preference types and participant group remained consistent and systematic across individual subjects and items.

To further address the interaction between participant group and preference types, four separate mixed-effects models [38] [39] were built with log-transformed RT as the dependent variable and preference type as the fixed effect to address the within group difference (six groups: Cantonese, English, Japanese Bilinguals, Multilinguals in a bilingual and in a monolingual context). The random effects included crossed random intercepts for participant and item. The intercept for each model set path-match alternate as the benchmark for comparison.

Table 41. Fixed effects on RT as a function of participant group and preference type in voluntary motion-multilingual speakers

| **Fixed effects** | **Estimate** | **SE** | **t value** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 7.706 | 0.044 | 173.580 | < 2e-16 ** |
| **Preference type** (Manner) | -0.105 | 0.035 | -3.034 | 0.002 ** |
| **Participant group** | | | | |
| Cantonese | -0.025 | 0.060 | -0.415 | 0.679 |
| English | 0.025 | 0.064 | 0.387 | 0.699 |
| Japanese | -0.130 | 0.059 | -2.187 | 0.029 * |
| Multilingual in bilingual context | -0.023 | 0.076 | -0.295 | 0.768 |
| Multilingual in monolingual context | -0.085 | 0.074 | -1.154 | 0.249 |
| **Preference type: Group** | | | | |
| Manner: Cantonese | 0.122 | 0.048 | 2.519 | 0.012* |
| Manner: English | 0.011 | 0.050 | 0.225 | 0.822 |
| Manner: Japanese | 0.265 | 0.053 | 5.075 | 4.3e-07 *** |
| Manner: Multilingual B | 0.064 | 0.065 | 0.983 | 0.326 |
| Manner: Multilingual M | 0.144 | 0.062 | 2.329 | 0.02 * |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.0004 | 0.017 | | |
| Item (Intercept) | 0.087 | 0.295 | | |

Note: The intercept represents the log-transformed RT when the preference type is Path-match alternate and participant group is Bilinguals.

The results confirmed that for multilinguals in both a bilingual ($\beta_0$ = -0.04, SE = 0.05, t =-0.791, p=0.43) and monolingual context ($\beta_0$ = 0.034, SE = 0.051, t =0.663, p = 0.508), their RT to manner-match alternates was as fast as path-match alternates, as illustrated in Figure 21. The results indicated that multilinguals patterned with Cantonese monolinguals in terms of processing efficiency when making manner- and path-match decisions regardless of the language context, indicating a backward influence from L3-based conceptualization patterns on the L1.

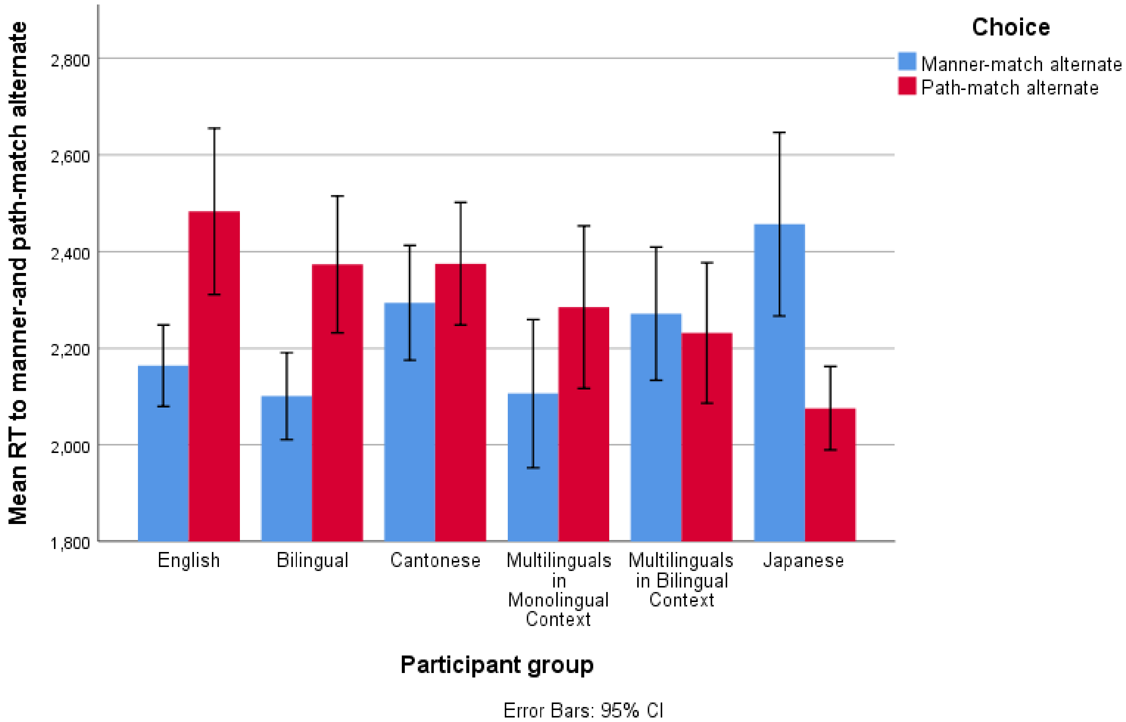



Figure 21 Mean RT to manner- and path-match alternate in terms of language group and language context in voluntary motion

### 9.1.3. Factors modulating the degree of cognitive restructuring in multilinguals' lexicalization and categorization of voluntary motion

To further explore the modulating effect of extra-linguistic factors in the process of cognitive restructuring, mixed-effects models were built with the three target factors under investigation: language proficiency, language context (a monolingual vs. a bilingual context) and language contact.

Following Athanasopoulos (2009) and Athanasopoulos et. al (2015), language contact here was defined as the amount of use multilinguals have with each of their languages and measured by participants' self-reported scores in a series of daily activities, such as watching TV, reading for school and talking with friends etc. The daily amount of language use (raw data) was converted into percentage scores. As multilingual speakers in the current study used Cantonese, English and Japanese interchangeably in daily interaction, the proportion of English, Cantonese and Japanese use was added up to 1. On the whole, CEJ-multilinguals used Japanese's 57.34% (SD=19.01%) of the time and therefore, Japanese was the dominant language in daily activities.

First of all, two separate logistic mixed-effects models [40] [41] were built with frequency of manner encoding, and manner-match preferences as separate binary dependent variables. The random effects were random intercepts for participant and item. The main effects were language context, English proficiency, Japanese proficiency and the amount of English use and Japanese use.

As the interactions of the above-mentioned factors were not statistically significant and including these interactions did not significantly optimize the model, the final models included the main effects (i.e. language context, English and Japanese proficiency and the amount of Japanese use) only. As shown in Table 42 and 43, among different factors under investigation, only language contact surfaced as a significant and unique predictor of multilinguals' processing efficiency. That is, the more frequently Japanese was used in participants' daily communication and interaction, the more likely multilingual speakers were to shift from the L1- and L2-based lexicalization and conceptualization patterns towards the L3-based associations. In contrast, as the amount of English was positively correlated with the frequency of manner encoding, it showed that the more frequently English was used, the more likely participants were to keep L2-based patterns in motion event lexicalization. Similar patterns were examined for motion event categorization.

Table 42. Fixed effects on the probability of manner encoding in event lexicalization of voluntary motion

| Fixed effects | Estimate | SE | Wald z | Pr(>\|z\|) |
|---|---|---|---|---|
| Intercept | 0.928 | 0.328 | -2.829 | 0.046 * |
| Language context | 0.402 | 0.332 | 1.213 | 0.786 |
| English proficiency | 1.819 | 0.336 | 5.411 | 0.157 |
| Japanese proficiency | -0.353 | 0.407 | -0.866 | 0.786 |
| Amount of English use | 2.221 | 0.340 | 6.532 | <.001 *** |
| Amount of Japanese use | -1.409 | 0.408 | -3.456 | 0.006 ** |

Note: The intercept represents the predicted probability of manner verb encoding (log-odds) when language context is bilingual, the proficiency level is the lowest and the frequency of English use is 0.

Table 43. Fixed effects on the probability of manner-match preferences in event categorization of voluntary motion

| Fixed effects | Estimate | SE | Wald z | Pr (>\|z\|) |
|---|---|---|---|---|
| Intercept | -2.452 | 0.857 | -2.864 | 0.004 |
| Language Context | -0.266 | 0.327 | -0.813 | 0.415 |
| English proficiency | 2.215 | 1.215 | 1.823 | 0.360 |
| Japanese proficiency | 0.812 | 0.408 | 1.987 | 0.329 |
| Amount of English use | 2.221 | 0.340 | 6.532 | <.001 *** |
| Amount of Japanese use | -1.466 | 0.413 | -3.547 | 0.005 ** |

Note: The intercept represents the predicted probability of manner-match preference (log-odds) when language context is bilingual, the proficiency level is the lowest and the frequency of English use is 0.

In addition, following Ji and Hohenstein (2017, 2018), a multiple linear regression[42] was fitted with mean differences of RT in manner-match preference minus path-match preference as dependent variable, and the target three factors as explanatory variables. Positive values of RT indicated longer response latency in making manner-match choices whereas negative values represented longer time in path-match choices. As shown in Table 44, among three different factors under investigation, only language contact surfaced as a significant and unique predictor of multilinguals' processing efficiency. The results showed that the more frequently Japanese was used in daily communication, the faster participants were to react to path-match choices in the decision-making process.

Table 44. Multiple regression with mean differences in RT as dependent variable in event categorization of voluntary motion

| **Independent variable** | **Standardized coefficient** | **SE** | **t value** | **p** |
|---|---|---|---|---|
| Intercept | -472.01 | 617.12 | -1.419 | 0.201 |
| Language context | -2219.06 | 266.13 | -8.34 | 0.098 |
| English proficiency | 113.34 | 517.13 | 0.22 | 0.104 |
| Amount of English use | -395.23 | 209.4 | -1.88 | 0.312 |
| Amount of Japanese use | 1018.11 | 280.11 | 3.635 | .031* |

Note: The intercept represents the predicted the mean differences in RT odds) when language context is bilingual, the proficiency level is the lowest and the frequency of English use is 0.

## 9.2. Lexicalization and categorization of caused motion in monolingual, bilingual and Cantonese-English-Japanese multilingual speakers

### 9.2.1. Linguistic encoding of caused motion

#### 9.2.1.1. Frequency of C-Manner and Path encoding in caused motion across monolinguals, bilinguals and Cantonese-English-Japanese multilinguals

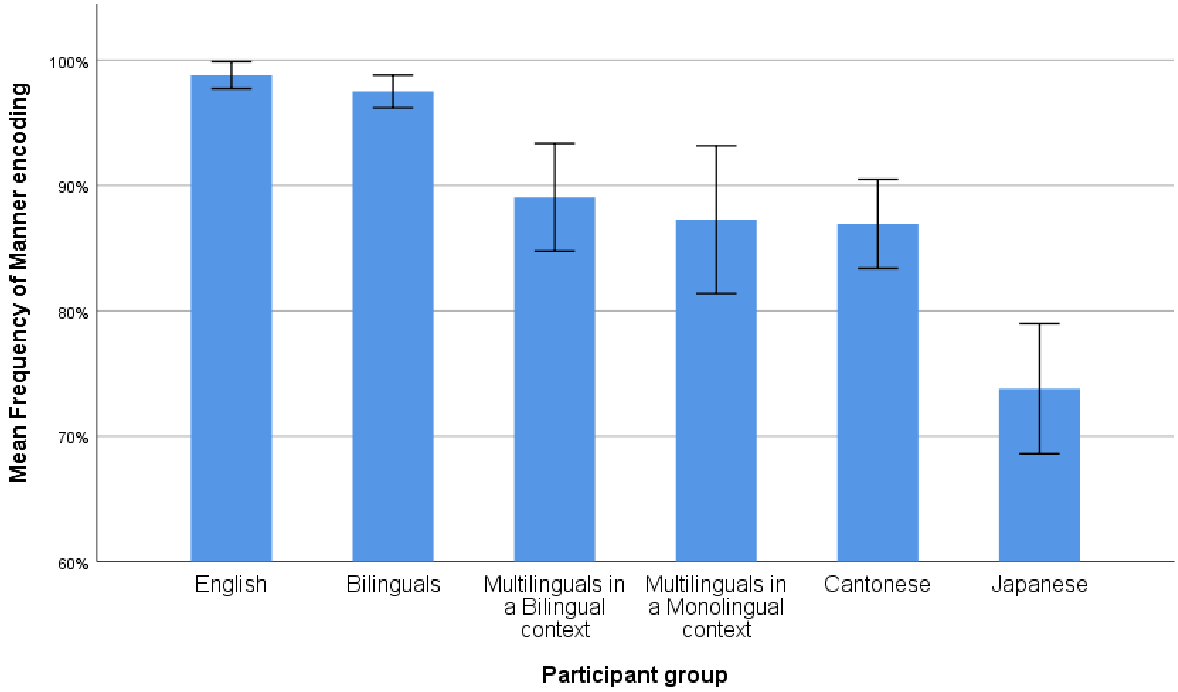



Figure 22. Mean frequency of manner encoding in terms of participant group and language context in caused motion

A total of 5400 target responses were included for the final analysis. Participants' linguistic encodings were calculated by the frequency of C-manner and path selection. Participants' responses to each stimulus were transformed into percentages and the mean percentage for each group was compared in terms of participant group and language context. Participants in each group presented a high tendency of path encoding, with a ceiling effect across four language groups. However, regarding C-manner encoding, multilinguals encoded manner more often than Japanese monolinguals but patterned with Cantonese monolinguals, as illustrated in Figure 22.

To assess whether speakers from different groups differed in their likelihood of manner and path encoding, two separate logistic mixed-effects models were built. Within each model, the binary dependent variable was whether the target semantic element (e.g. manner and path of motion) was encoded (code=1) or not (code=0). The fixed effect was participant group (five levels: English monolinguals *vs*. Bilinguals vs. Multilinguals in a monolingual context vs. Multilinguals in a bilingual context vs. Japanese monolinguals). The random effects were random intercepts for participant and item. For path encoding, the results showed that the inclusion of group did not significantly increase the model fit compared with the null model ($\chi_2$ (5) =9.42, p=1.241), indicating that group was not a main effect. In other words, participants across different groups were equally likely to encode path of motion when describing voluntary motion.

Table 45. Coefficients of the logistic mixed-effects model for the frequency of manner encoding in caused motion-multilingual speakers

| **Fixed effects** | | **Estimate** | **SE** | **Wald z** | **Pr (>\|t\|)** |
|---|---|---|---|---|---|
| **Intercept** | | 3.262 | 0.214 | 15.228 | <.001*** |
| **Participant group** | | | | | |
| Cantonese | | -1.104 | 0.168 | -6.576 | <.001*** |
| English | | 1.490 | 0.309 | 4.811 | 0.069 |
| Japanese | | -2.053 | 0.159 | -12.916 | <.001*** |
| Multilingual B | | -0.766 | 0.206 | -3.704 | <.001*** |
| Multilngual M | | -1.097 | 0.199 | -5.502 | <.001*** |
| **Random effects** | **Variance** | **Standard deviations** | | | |
| Subject (Intercept) | 0.384 | 0.619 | | | |
| Item (Intercept) | 0.421 | 0.649 | | | |

Note: The intercept represents the condition of Bilinguals as a baseline for between-group comparisons.

However, with regard to C-manner encoding, multilinguals tended to mention manner of motion more often compared with Japanese monolinguals yet patterned with Cantonese monolinguals (Multilingual in a monolingual context: M=87.40%, SD=8.88%; Multilingual in a bilingual context: M=89.00%, SD=9.26%; Cantonese: M=86.94%, SD=9.51%; Japanese: M=73.79%, SD=10.89%), as illustrated in Figure 22 (Please refer to Table 15 for the linguistic encoding of caused motion in other language groups). Then forward coding was used to compare the likelihood of manner encoding with the next group. As shown in Figure 22, multilinguals in a bilingual context encoded manner less often than bilingual speakers ($\beta_{\text{Bilinguals-Multilinguals in bilingual context}}$ = -0.766, SE = 0.21, Wald z =-3.71, $p < .001$) but patterned with multilinguals in a bilingual context ($\beta_{\text{Multilinguals in monolingual context--Multilinguals in bilingual context}}$ = -0.33, SE = 0.22, Wald z =-1.51, p = 0.26). Meanwhile, no difference between multilinguals in a monolingual context and Cantonese monolinguals was detected ($\beta_{\text{Multilinguals in bilinguals in monolingual context-Cantonese monolinguals}}$ = 0.1, SE = 0.17, Wald z =0.039, p = 0.97). Detailed statistical information is presented in Table 45. The findings indicated that multilinguals demonstrated a cognitive shift towards L3-based lexicalization patterns irrespective of language contexts.

#### 9.2.1.2. Semantic distribution of C-Manner and Path across monolinguals, bilinguals and Cantonese-English-Japanese multilinguals

Based on the frequency of C-manner and path encoding, the semantic distribution of manner and path was further explored within each utterance. As mentioned in the coding outline, there are two possible loci for C-manner and path: either in the form of the main verb, or outside of the main verb in a satellite. As shown in Table 46, speakers in the three language groups differed from each other in the information locus of manner and path. The results indicated that the semantic distribution in multilingual speakers showed a clear L3 influence. For example, multilinguals in both monolingual and bilingual contexts demonstrated a shift from the L2-based linguistic descriptions but patterned with Cantonese monolingual in encoding C-manner in both main verbs (M=53.33%, SD=9.52% for multilinguals in a monolingual context and M=46.85%, SD= 9.89% for multilinguals in a bilingual context) and subordinate forms. Detailed information is presented in Table 46.

Table 46. Percentages of semantic distribution of manner/path in verb (V) or outside of it (OHT) in caused motion.

| | **Cantonese (%)** | | | **Multilinguals in monolingual context (%)** | | | **Multilinguals in bilingual context (%)** | | | **Bilinguals (%)** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Component | V | OTH | Total | V | OTH | Total | V | OTH | Total | V | OTH | Total |
| C-Manner | 48.33 | 38.05 | 86.94 | 53.33 | 34.70 | 87.47 | 46.85 | 41.85 | 89.00 | 64.17 | 30.09 | 95.00 |
| Path | 42.77 | 54.44 | 95.83 | 38. 15 | 58.52 | 93.70 | 45.56 | 50.00 | 95.90 | 32.41 | 62.48 | 96.22 |

Notes: The sum of the first two columns within each language group doesn't always add up to the total proportion as the manner or path of motion can be double-encoded in V and OTH at the same time (e.g.: The boy is [jumping]$_{verb}$ downstairs [with one leg] $_{OTH}$).

#### 9.2.1.3. Construction type and framing strategies of caused motion event across monolinguals, bilinguals and Cantonese-English-Japanese multilinguals

To further explore the effects of L3 learning on motion event encoding in multilingual speakers, a qualitative analysis was conducted regarding construction types and framing strategies in each language. As reviewed in the monolingual section, the typological status of Cantonese, English and Japanese exhibited language-specific characteristics (cf. Figure 7). For the linguistic data of multilingual speakers in different language contexts, multilingual speakers demonstrated an ongoing shift from the L1- and L2-based patterns towards the L3-based ones in terms of the semantic distribution of C-manner and path, as well as framing strategies. A qualitative analysis showed that multilinguals tended to use 'Path only+ Causation' constructions (verb-framing strategy) and omit C-manner of motion more frequently compared with bilinguals. For example, multilinguals tended to mention manner of motion less often compared with bilinguals yet patterned with Cantonese monolinguals in describing the same motion scene, suggesting that an ongoing effect of L3 learning on counterbalancing language-specific effects from L2 and restoring the original L1-based linguistic patterns. Specific examples are given in (73) and (74).

(73) **Verb-framing strategy (type 1): Path verb + Causation of motion**

Go3 naam4 zai2 **jap6-zo2** jat1 go1 bo1

A boy **entered-ASP** a ball

'A boy made a ball enter the cave.'

(74) **Satellite-framing strategy: Manner verb + Path satellite**

Go3 naam4 zai2 **tek3-zo2** jap4 jat1 go1 bo1

A boy **kick-ASP** into a ball

'A boy kicked a ball into the cave.'

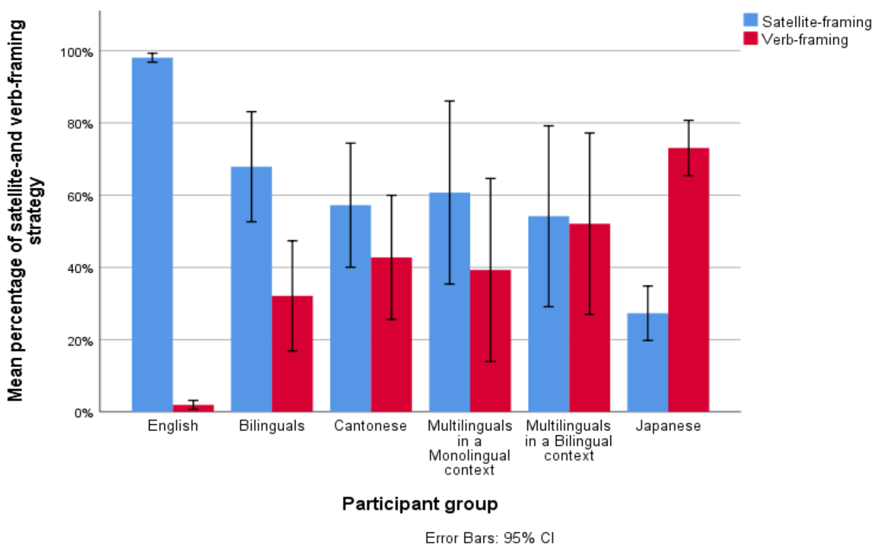



Figure 23. Mean percentage of different framing tendencies (satellite- and verb-framing) in terms of participant group and language context in caused motion

Quantitative data further confirmed that multilinguals in different language contexts adopted less satellite-framing strategy compared with their bilingual counterparts, as illustrated by Figure 23 (Bilinguals: M=68.87%, SD=10.79%; Cantonese monolinguals: M=57.22%, SD=46.00%; Multilinguals in a monolingual context: M=59.85%, SD=32.85%; Multilinguals in a bilingual context: M=55.44%; SD=32.04%).

Table 47. Coefficients of the logistic mixed-effects model for the frequency of satellite-framing strategies in caused motion-multilingual speakers

| **Fixed effects** | **Estimate** | **SE** | **Wald z** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | 1.065 | 0.290 | 3.671 | <.001*** |
| **Participant group** | | | | |
| Cantonese | -0.676 | 0.109 | -6.181 | <.001*** |
| English | 3.886 | 0.250 | 15.548 | 0.101 |
| Japanese | -2.487 | 0.119 | -20.816 | <.001*** |
| Multilingual B | -0.701 | 0.136 | -5.125 | <.001*** |
| Multilngual M | -5.299 | 0.144 | -3.692 | <.001*** |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.445 | 0.667 | | |
| Item (Intercept) | 0.800 | 0.895 | | |

Note: The intercept represents the condition of bilinguals as a baseline for between-group comparisons.

A mixed-effects logistic model was fitted with the presence or absence of satellite-framing as the binary dependent variable. The fixed effect was group and the random effects were intercepts for participant and item. The results showed that multilinguals in a monolingual context used less satellite-framing strategies than bilinguals ($\beta_{\text{Multilinguals in bilingual context-Bilinguals}}$ = -5.29, SE = 0.14, Wald z =-3.69, p<.001), but patterned with multilinguals in a bilingual context ($\beta_{\text{Multilinguals in monolingual context-Multilinguals in a bilingual context}}$ = 0.172, SE = 0.17, Wald z =0.98, p = 0.89). Meanwhile, Japanese monolinguals still demonstrated the lowest frequency of satellite-framing ($\beta_{\text{Multilinguals in bilingual context-}}$

$_{\text{Japanese monolinguals}}$ = 1.79, SE = 0.14, Wald z =12.69, $p < .001$). The results indicated an ongoing effect of L3 learning on motion event lexicalization in multilinguals with two typologically different languages. Detailed statistical information of the framing strategies is presented in Table 47.

### 9.2.2. Non-linguistic categorization of caused motion

#### 9.2.2.1. Categorical preferences of manner- and path-match alternates across monolinguals, bilinguals and Cantonese-English-Japanese multilinguals

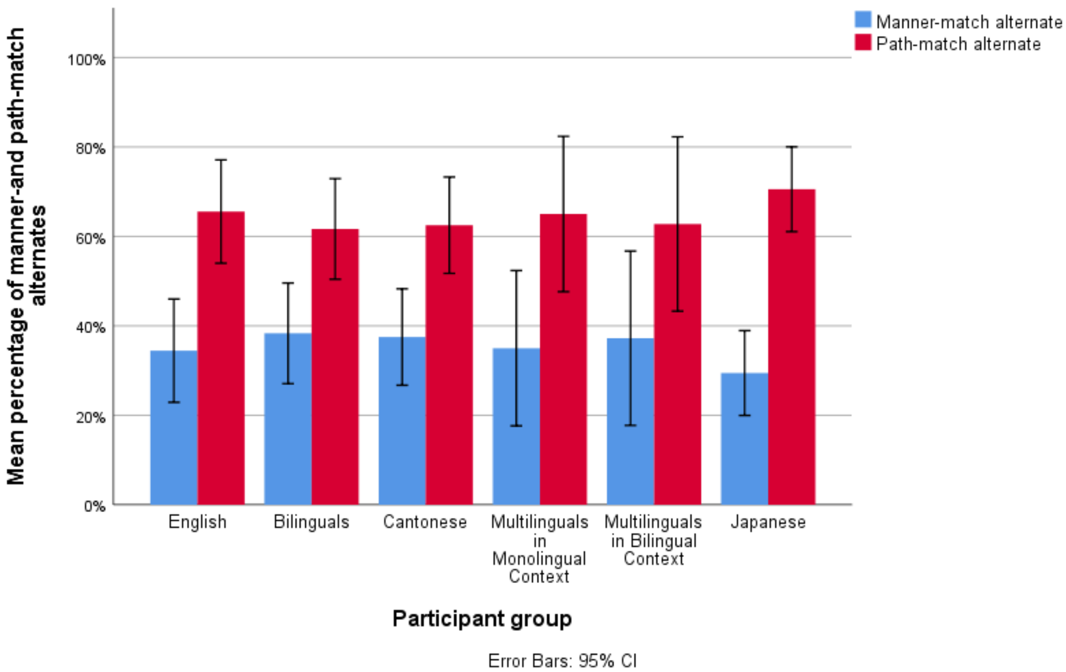



Figure 24. Mean percentage of manner/path preferences in terms of participant group and language context in caused motion

In conceptualizing caused motion, participants in each group presented an overall preference for path-match alternate: Cantonese monolinguals (M=63.89%, SD=28.85%), bilinguals (M=61.67 %, SD=29,13%), multilinguals in a bilingual context (M=62.78%, SD=35.20%), multilinguals in a monolingual context (M=65.00%, SD=31.37%), and Japanese monolinguals (M=70.56%; SD=25.40%), as shown in Figure 24.

A mixed-effects logistic model was fitted to detect whether the observed differences were statistically significant. The final analysis modelled participants' manner-match preferences as a binary dependent variable. The fixed effect was participant group. The random effects were crossed-random intercepts for participant and item. Including participant group as the fixed-effect significantly optimized the model ($\chi_2$ (5) =9.48, p=0.09) compared with the null model, indicating that group was a main effect in participants' similarity judgements. The significance in the overall intercept ($\beta_0$ = -0.53, SE = 0.21, Wald z =-2.61, p = 0.009) indicated that participants across language groups had an overall preference for path-match alternates. In other words, participants across different language groups were more prone to use 'path-match' as their selection criteria when categorizing caused motion.

Table 48. Coefficients of the logistic mixed-effects model for the frequency of manner selection in the conceptualization of caused motion-multilingual speakers

| **Fixed effects** | **Estimate** | **SE** | **t value** | **Pr (>\|t\|)** |
|---|---|---|---|---|
| **Intercept** | -0.534 | 0.205 | -2.6.1 | 0.009 ** |
| **Participant group** | | | | |
| Cantonese | -0.039 | 0.162 | -0.245 | 0.806 |
| English | -0.188 | 0.164 | -1.115 | 0.250 |
| Japanese | -0.445 | 0.167 | -2.662 | 0.358 |
| Multilingual B | -0.205 | 0.206 | -0.966 | 0.319 |
| Multilngual M | -0.004 | 0.207 | -0.022 | 0.982 |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.368 | 0.606 | | |
| Item (Intercept) | 0.199 | 0.447 | | |

Note: The intercept represents the condition of bilinguals as a baseline for between-group comparisons.

#### 9.2.2.2. RT to manner- and path-match alternates across monolinguals, bilinguals and Cantonese-English-Japanese multilinguals

Furthermore, in order to better explore the degree of differences in the decision-making process, participants' RTs to manner- and path-match preferences were measured as a continuous variable and used to indicate participants' efficiency in cognitive processing. The mean RT to manner- and path-alternate across each participant group is presented in Table 49.

Table 49. Mean RT (in millisecond) to manner- and path-match alternate across groups in caused motion-multilingual speakers

| Participant Group | Mean RT in Manner-match | Mean RT in Path-match |
|---|---|---|
| English | 2048 (683) | 2416 (840) |
| Cantonese | 2165 (810) | 2247 (792) |
| Japanese | 2346 (713) | 2000 (692) |
| Bilinguals | 2275 (577) | 2556 (847) |
| Multilinguals in monolingual context | 2605 (740) | 2395 (516) |
| Multilinguals in bilingual context | 2492 (839) | 2433 (536) |

To further explore the statistical significance, a mixed-effects model was built with RT as the continuous dependent variable. The fixed effects included participant group (five levels: Cantonese, Japanese, English monolinguals and multilinguals in a bilingual context and a monolingual context), preference type (two levels: manner- and path-match preference) and their interaction. The random effects included the crossed random intercepts for participant and item. The dependent variable (RT) was log-transformed in order to meet the assumption of the normality of residuals. Details of fixed-effect parameters are presented in Table 50.

Table 50. Fixed effects on RT as a function of participant group and preference type in caused motion-multilingual speakers

| Fixed effects | Estimate | SE | t value | Pr (>\|t\|) |
|---|---|---|---|---|
| **Intercept** | 7.795 | 0.031 | 249.03 | < 2e-16 *** |

| | | | | |
|---|---|---|---|---|
| **Preference type** (Manner) | -0.108 | 0.034 | -3.175 | <.001** |
| **Participant group** | | | | |
| Cantonese | -0.139 | 0.292 | -4.764 | <.001** |
| English | -0.076 | 0.029 | -2.658 | <.001** |
| Japanese | -0.250 | 0.028 | -8.831 | <.001** |
| Multilingual in bilingual context | -0.019 | 0.037 | -0.530 | 0.596 |
| Multilingual in monolingual context | -0.042 | 0.036 | -1.183 | 0.237 |
| **Preference type: Group** | | | | |
| Manner: Cantonese | 0.064 | 0.048 | 1.324 | 0.186 |
| Manner: English | -0.024 | 0.048 | -0.500 | 0.617 |
| Manner: Japanese | 0.270 | 0.050 | 5.363 | <.001** |
| Manner: Multilingual B | 0.036 | 0.049 | 0.618 | 0.537 |
| Manner: Multilingua lM | 0.213 | 0.060 | 3.543 | <.001** |
| **Random effects** | **Variance** | **Standard deviations** | | |
| Subject (Intercept) | 0.037 | 0.192 | | |
| Item (Intercept) | 0.002 | 0.041 | | |

Note: The intercept represents the log-transformed RT when the preference type is Path-match alternate and participant group is Bilinguals.

With regard to the fixed effects, there was a main effect of participant group and a participant group by preference type interaction effect. This indicated that participants in each group differed in their processing efficiency when selecting manner- or path-match alternates. In addition, the variance of crossed random intercept and slope provided further evidence that the interaction between preference type and participant group remains consistent and systematic across individual subject and item.

To further address the interaction between participant group and preference types, four separate mixed-effects models were built with log-transformed RT as the dependent variable and preference type as the fixed effect to address the within group difference (for the RT for Cantonese and English monolinguals, please refer to the monolingual section). The random effects included crossed random intercepts for participant and item. The intercept for each model set path-match alternate as the benchmark for comparison.

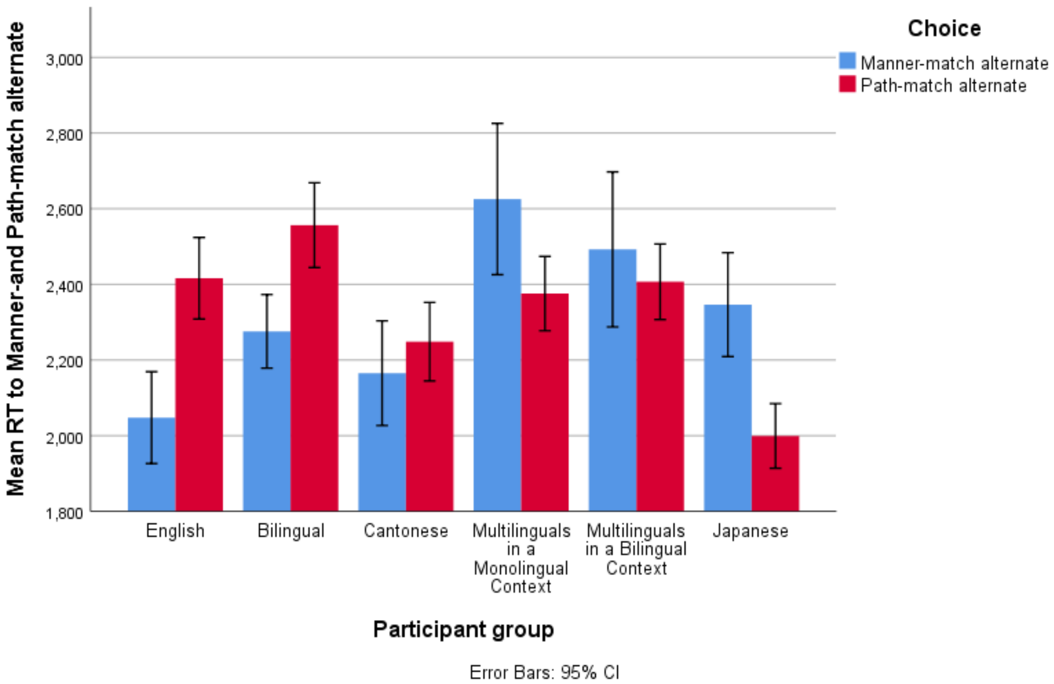



Figure 25. Mean RT to manner- and path-match alternate as a function of language group and language context in caused motion

Similar to of voluntary motion, the results of caused motion further confirmed that for multilinguals in both a bilingual ($\beta_0$ = 0.006, SE = 0.04, t =-0.153, p=0.87) and monolingual context ($\beta_0$ = 0.058, SE = 0.038, t =1.542, p = 0.125), their mean RT to manner- and path-match alternate was the same (Figure 25). The results indicated an ongoing process of cognitive restructuring from L1-based processing efficiency towards L3-based patterns. And the restructuring process was not subject to short-term experimental manipulation.

#### 9.2.3. Factors modulating the degree of cognitive restructuring in multilinguals' lexicalization and categorization of caused motion

Like voluntary motion, this section further explored how the target factors (i.e. language context, English proficiency, Japanese proficiency, the amount of English use and the amount of Japanese use) tended to modulate the degree of cognitive restructuring within the multilingual mind. As indicated by multilinguals' performance in overt selection, participants across different language groups had an overall preference for path-match alternate. Thus, this factor was not computed as a dependent variable for the current analysis.

With regard to the linguistic encoding, one logistic mixed-effects model was built with frequency of manner encoding as the binary dependent variable. The random effects were random intercepts for participant and item. The main effects were language context, English proficiency, Japanese proficiency and the amount of English and Japanese use. As the interaction of these factors were not statistically significant and including the interaction did not significantly optimize the model ($\chi_2$ (1) =2.036, p=0.758 for manner encoding in lexicalization), the final model included the main effects (i.e. language context, English and Japanese proficiency and language contact with English and Japanese) only. As shown in Table 51, among the four different factors under investigation, only language contact was significantly predicted bilinguals' frequency of manner encoding in the linguistic descriptions. That is, the more frequently Japanese was used in participants' daily communication and interaction, the more likely bilingual speakers were to shift from the L1-based lexicalization patterns towards the L3-based associations.

Table 51. Fixed effects on the probability of manner encoding in lexicalization of caused motion

| **Fixed effects** | **Estimate** | **SE** | **Wald z** | **Pr(>\|z\|)** |
|---|---|---|---|---|
| Intercept | 0.706 | 0.868 | 0.814 | 0.415 |

| | | | | |
|---|---|---|---|---|
| Language context | -2.826 | 0.535 | -5.276 | 0.215 |
| English proficiency | 0.439 | 0.299 | 1.464 | 0.286 |
| Japanese proficiency | -0.654 | 0.471 | -1.389 | 0.786 |
| Amount of English use | -1.887 | 0.553 | -3.412 | 0.132 |
| Amount of Japanese use | -5.005 | 0.263 | -19.03 | 0.021* |

Note: The intercept represents the predicted probability of manner verb encoding (log-odds) when language context is bilingual, the proficiency level is the lowest and the frequency of English use is 0.

Then a multiple linear regression was fitted with the mean differences of RT in manner-match preference minus path-match preference as dependent variable, and the target three factors as explanatory variables. Positive values indicated longer RT in making manner-match choices whereas negative values represented longer time in path-match choices. As shown in Table 52, among three different factors under investigation, only language contact surfaced as a significant and unique predictor of bilinguals' processing efficiency. The results showed that the more frequently Japanese was used in daily communication, the faster the participants were to react to path-match choices in the response latency of caused motion.

Table 52. Multiple regression with mean differences in RT as dependent variable in categorization of caused motion

| **Independent variable** | **Standardized coefficient** | **SE** | **t value** | **p** |
|---|---|---|---|---|
| Intercept | 205.45 | 88.8 | 2.313 | < 2e-16 *** |
| Language context | -1023.11 | 192.9 | -5.304 | 0.271 |
| English proficiency | 1051.20 | 275.41 | 3.686 | 0.251 |
| Japanese proficiency | -1023.39 | 456.35 | -2.242 | 0.147 |

| | | | | |
|---|---|---|---|---|
| Amount of English use | -987.79 | 290.23 | -3.403 | 0.981 |
| Amount of Japanese use | 1787.67 | 523.13 | 3.414 | 0.020 * |

Note: The intercept represents the predicted the mean differences in RT odds) when language context is bilingual, the proficiency level is the lowest and the frequency of English use is 0.

# Chapter 10. General discussion and conclusion

The current study goes beyond the bipartite classification of motion events with multilingual speakers of three typologically different languages (Cantonese, English, and Japanese). It aims to expand the sphere of the effect of language learning on cognition from bilingualism to multilingualism by adopting a lexical perspective of manner versus path. Specifically, it examines the questions of how, and to what extent, the acquisition of an L2-English (satellite-framed) and an L3-Japanese (verb-framed) restructure lexicalization (i.e. event structures and semantic distributions) and conceptualization patterns (i.e. categorical preferences and RT) associated with the L1-Cantonese (equipollently-framed).

In addition, the current study also addresses how the long-term effect of language learning and short-term effect of language manipulation interact with each other in modulating the conceptual changes within the bi- and multilingual mind. That is, how Cantonese-English bilinguals and Cantonese-English-Japanese multilinguals lexicalize and conceptualize motion events in different language contexts. According to the experimental design, the language context in the current study is manipulated with a monolingual (L1) and a bilingual (L1 and L2/L3) context as a function of recent activation of the L2 or L3. In addition, it also explores how other long-term extra-linguistic factors such as L2/L3 proficiency and language use with the L2/L3 affect bilinguals' performances in both event lexicalization and conceptualization.

In this chapter, the main research findings are summarized and discussed in terms of four research questions. At the end of this section, the theoretical and methodological innovations, pedagogical implications, limitations, as well as recommendations for future research will be presented.

### 10.1. Research Question #1: Motion event lexicalization and categorization in monolinguals of Cantonese, English and Japanese

The first research question examined how monolingual controls of Cantonese, English and Japanese lexicalize and conceptualize motion events, and whether a higher level of manner salience in the linguistic encoding is associated with more attention to manner in categorization when the access to language is not blocked in the decision-making process (i.e. the thinking-for-speaking and thinking-with-language effects). As the typological status of Cantonese is still under great debate (i.e. whether it is a satellite- or equipollent-framed language), the first research question can provide some insights on this issue by offering some evidence from both a linguistic and cognitive perspective.

#### 10.1.1. Lexicalization and categorization of voluntary motion in monolinguals

##### 10.1.1.1. Lexicalization of voluntary motion

The lexicalization patterns of voluntary motion were analysed from three perspectives, 1) the overall frequency of manner and path selection, 2) the semantic distributions of manner and path of motion, and 3) language-specific framing strategies for motion event constructions.

The monolingual data confirmed the typological status of each language in how speakers selected and structured information when talking about voluntary motion. First, speakers of each language differed from each other in ‘what to say’ in terms of the macroplanning, that is, which elements they opted to encode in motion event expressions (i.e. information segmentation and selection). The results showed that on the one hand, monolinguals of each group were equally likely to express path with a high frequency regardless of the path types, indicating that path is a core and obligatory element in motion constructions (Slobin, 2006; Talmy, 1985, 2000).

However, for manner encoding, the frequency of manner selection clearly contrasted across each language. As a typical S-language, speakers of English expressed manner of motion more frequently than Cantonese (E-language), while Japanese (V-language) exhibited the lowest frequency. The differences in information selection of manner of motion can be attributed to the language-specific conflation patterns in voluntary motion (Talmy, 1985, 2000). As a typical S-language, manner of motion in English is predominantly encoded in the verb root whereas path in the satellite. The marked grammatical category of manner verbs makes English speakers less likely to omit manner during their oral descriptions (i.e. manner salience). Although English has a set of path verbs (i.e. ascend, descend), most of them are Latinate borrowings and less colloquial in style. Thus, path verbs in English are seldom used in oral expressions (Slobin, 1996a). In contrast, as a typical V-language, Japanese conflates path of motion in the main verb, leaving manner not expressed, or in subordinate forms (i.e. via the -te form). As there is no obligatory syntactic slot for the manner encoding, and therefore, manner of motion can be easily added or dropped in speakers' oral expressions (i.e. path salience). Thus, for the ease of information encoding and processing, Japanese speakers were more prone to opt for leaving out manner in their oral descriptions unless they would like to inform the interlocutor in what specific ways a voluntary motion is performed.

Qualitative analysis further confirmed that speakers of English and Japanese differed significantly in terms of motion event constructions (i.e. Manner verb + Path satellite or Path verb-only) and the framing strategies (i.e. satellite- or verb-framing), that is, speakers differed in 'how to say it' (microplanning of structuring and linearization). To be more specific, English-speaking participants used semantically and syntactically dense constructions and structures of ‘manner verb+ path satellite’, a typical satellite-framing strategy throughout their oral descriptions, even for the type of manner with the lowest level of salience (i.e. walk, run). However, Japanese speakers most often adopted ‘Path-only’ construction, a typical verb-framing strategy, and omitted manner information especially when manner is less salient or can be easily inferred from the motion scenes. The results of the current studies echoed with the previous literature that speakers of English and Japanese differed from each other not only in ‘what to say’, but also ‘how to say it’. These typological differences are well-reflected in English and other S-languages in general (Hickmann & Hendriks, 2010; Larrañaga et al., 2012; Papafragou & Selimis, 2010; Park, 2019; Pavlenko & Volynsky, 2015a; Slobin, 1996a, 2004, 2006; Talmy, 1985, 2000, 2009, 2012), as well as Japanese and other V-languages (Brown & Gullberg, 2008, 2010, 2011; Hendriks et al., 2008; Hickmann, Hendriks, et al., 2009; Inagaki, 2001, 2002; Soroli, 2012; Soroli et al., 2012).

Following Slobin (2006) and Özçalışkan and Slobin (2003), the typological status of a certain language in motion encoding can be assessed from two perspectives: (1) the frequent use of manner expressions across different language modalities (i.e. oral narrative, news reporting, etc.) and (2) the lexical size and diversity of manner expressions made available in a particular language. Based on the typological features of English and Japanese, and other S-languages and V-languages in general, the results from Cantonese-speaking participants confirmed that Cantonese integrates the typological features of both S- and V-languages and consistently posts itself between English and Japanese in terms of the frequency of manner selection, motion event constructions, and the framing strategies. Regarding the overall frequency of manner selection, Cantonese speakers encoded manner less frequently than English speakers but more frequently than Japanese speakers. These differences can be attributed to the typological status and different degrees of manner and path salience embodied by the three languages. Due to the availability of path verbs (directional verbs) in Cantonese, manner of motion can be easily added or dropped in linguistic descriptions. This allows Cantonese speakers to opt freely for different constructions (i.e. Path-only construction or Manner-Path construction) and framing strategies of motion, depending on the various perceived degrees of manner salience. Thus, the frequency of manner encoding in Cantonese is lower than that in English, where path verbs are seldom used in oral expressions. Meanwhile, although Cantonese allows manner to be easily added or dropped and follows a typical verb-framing conflation pattern, this construction is not used as frequently as in Japanese (Yiu, 2013, 2014). Qualitative data regarding construction types and framing strategies further confirmed that in Cantonese, both satellite-framing and verb-framing strategies were used equally frequently (cf. Figure 3) in speakers' oral descriptions. And the most prototypical constructions of motion were 'Manner verb +Path satellite' (satellite-framing), 'Manner adjunct + Path verb' and 'Path-only constructions' (verb-framing). Given the most prototypical patterns observed in speakers' oral description, Cantonese is language that incorporates both satellite-framing and verb-framing strategies, and both strategies are used systematically and consistently throughout the encoding of voluntary motion.

The typological status of Cantonese in the current study support recent findings that typological distinctions between the satellite- and verb-framed dichotomy are not absolute but as a matter of varying degrees of manner and path salience (Ibarretxe-Antuñano, 2009; Ji et al., 2011a, 2011b; Ji & Hohenstein, 2018; Loucks & Pederson, 2011; Montero-Melis et al., 2017; Pavlenko & Volynsky, 2015a; Slobin, 2004, 2006;

Yiu, 2013, 2014). And the major difference between an S- and V-framed language lies in the number of available manner verbs and path verbs in a certain language (Özçalışkan & Slobin, 2003; Slobin, 2004). In fact, this interpretation well reflects Slobin's typological argument that 'manner expressions are better dealt with by placing languages on a cline of manner salience, rather than placing them into dichotomized or tri-chotomized typologies' (Slobin, 2004, p228). In addition, the current findings also lend support to the hypothesis of manner salience (Slobin, 1996a, 2000, 2006) that cross-linguistic differences in event lexicalization are only overserved in the likelihood of manner selection, but not in path. And the degree of difference is modulated by perceived salience of manner by speakers of different languages (Slobin, 2004, 2006; Soroli & Hickmann, 2010).

The overall findings on the lexicalization of voluntary motion suggest that language-specific constraints modulate the process of information selection, organization, and structuring during the process of speech planning (Carroll et al., 2011; Daller et al., 2011; Flecken, Gerwien, et al., 2015; Von Stutterheim & Nuse, 2003). And speakers of different languages structure and package information differently based on the linguistic resources made available in their language. The current findings lend support to Levelt's (1989) speech production model and Slobin's thinking-for-speaking hypothsis (Slobin, 1996a) that language-specific knowledge plays an important role both at the global and macro-structural level, as well as the local and micro-structural level when speakers are engaged in language driven activities, such as language production and comprehension.

#### 10.1.1.2. Categorization of voluntary motion in monolinguals

Based on language-specific lexicalization patterns in motion event encoding, the first research question further examined whether cross-linguistic differences can go beyond the linguistic level and affect a deeper level of cognition, and whether the codability of manner in event lexicalization is associated with the accessibility to this information in event categorization. A similarity judgment task was manipulated with two types of measurements: a categorical preference of participants' overt selection (i.e. manner- and path-match variants), and a continuous variable of reaction time (RT) in the decision-making process. The results showed that in both the overt selection and processing efficiency in making manner and path decisions, speakers of each language exhibited language-specific patterns in accordance with the respective languages, indicating a clear 'thinking-for-speaking' or 'thinking-with-language' effect of the language-specific properties on categorization.

In terms of the overt selection of manner and path preferences, results provided clear evidence that as a typical S-language, English monolinguals reached a ceiling level of manner encoding and had a manner-match preference in categorization. In contrast, Japanese monolinguals had the lowest level of manner encoding and had an overall preference for path-match variants. Meanwhile, Cantonese, right located in the middle of the S- and V-language continuum, demonstrated an equal preference for manner-match and path-match variants in event categorization. The results suggest that speakers' attention during perceptual or cognitive tasks are likely to be guided by the linguistic components that are naturally highlighted by language (Boroditsky, 2001; Bylund & Jarvis, 2011; Langacker, 2008; Lucy, 1992a, 2014; Slobin, 2000, 2004, 2006; Von Stutterheim et al., 2012). Thus, the manner-path asymmetry in linguistic encoding and different degrees of manner and path salience across languages may directly prompt speakers to draw on different criteria when making their similarity judgements. Consequently, speakers of English (manner-salience language) displayed a categorization preference for manner-match variant, while speakers of Japanese (path-salience language) exhibited a preference for path-match variant. However, speakers of Cantonese (manner-salience and path-salience) used manner- and path-based criteria equally frequently in their decision-making process, indicating that manner and path in Cantonese are of equal amount of salience (comparable proportions of manner-match and path-match alternates, cf. Figure 4).

Cross-linguistic differences observed in categorization echoed with previous findings that participants' lexicalization patterns are predictive of their conceptual preferences when the access to language is not blocked in the decision-making process (Athanasopoulos, Bylund, et al., 2015; Engemann, Hendriks, Hickmann, Soroli, & Vincent, 2015; Lai et al., 2014; Montero-Melis & Bylund, 2017; Papafragou & Selimis, 2010; Soroli & Hickmann, 2010). For example, Soroli and Hickmann (2010) investigated whether cross-linguistic differences in motion event lexicalization modulated English and French speakers' categorical preferences by using a triad-matching paradigm. Results suggested that in line with language-specific lexicalization patterns, French speakers showed a preference for path-match variants under both a verbal and non-verbal condition, whereas English speakers presented a significant manner preference only in the verbal condition where linguistic encoding was provided prior to categorization. Similarly, Montero-Melis and Bylund (2017) examined the language effects on cognition with Swedish (S-language) and Spanish (V-language) native speakers under different conditions. Results suggested that monolinguals of Swedish were more likely to use 'same-manner' criteria for the similarity assessment as long as they can access the language during their decision-making process. These results echoed with the findings of Gennari and others that Spanish-speaking participants were more likely to use ‘path similarity’ rather than ‘manner similarity’ as the criteria to group motion events when they were asked to verbalize the target motion scenes before categorization (Gennari et al., 2002).

For these converging findings on the effects of language on cognition, one possible explanation is that during language production, speakers of different languages may have different mental representations of motion events. When participants are asked to do a categorization task which involves higher-level processing within a limited time and when the task lacks a correct answer, they may search for and depend on all available resources, including recent linguistic experience as a strategy to solve the cognitive task (Bylund & Athanasopoulos, 2014b). In this case, as proposed by the ‘thinking-for-speaking’ and ‘thinking-with-language’ accounts, the role of language-specific features during cognitive processing can be viewed as a meddler, or augmenter, to facilitate the categorical perceptions (Lupyan, 2012; Slobin, 1996a; Wolff & Holmes, 2011). In the current case, as participants not only watched, but also described all dynamic motion stimuli immediately prior to the subsequent categorization task, cross-linguistic differences in their linguistic encoding (i.e. different degrees of manner salience or codability of manner across each language) may exert an online effect on the accessibility to the corresponding conceptual representations of manner or path during cognitive processing. These converging findings are in line with the ‘thinking for speaking’ and ‘thinking with language’ account that the effects of language on thought are the strongest when the tasks require an overt verbalization of the stimuli (Gennari et al., 2002; Montero-Melis & Bylund, 2017; Papafragou & Selimis, 2010; Soroli & Hickmann, 2010), or recognition memory (Filipović, 2011; Kersten et al., 2010).

Another possible and very much related explanation is that from the label feedback account that language-specific labelling may become activated in tandem with non-linguistic processes and form an automatic association between linguistic labelling and other mental processes, such as perception and visual representations (Lupyan, 2012; Lupyan et al., 2007; Vanek, 2019). Based on the 'label-feedback' hypothesis (Lupyan, 2012), language can produce transient modulation and affect ongoing perceptual and higher-level processing, such as categorization and object perception. When speakers are primed with language-specific labels before presenting with the visual stimuli, the co-activation of linguistic and non-linguistic representations tend to guide participants to group events in accordance with the linguistic categories they are primed with. This process of co-activation will become automatic and conventionalized when participants are given sufficient exposure to the target linguistic forms. Therefore, the linguistic encoding right before event categorization can enhance participants' sensitivity to language-specific properties and boost their L1-based associations when they describe all events before the non-linguistic categorization. As a result, the different weight or salience that speakers attach to each semantic element in linguistic encoding tends to trigger a language-specific way of thinking during their similarity judgements.

In addition to the language-specific application of manner- and path-match criteria in event categorization, the RT to manner- and path-match selection also demonstrated clear language-specific patterns: English monolinguals reacted more quickly in making manner-match choices than path-match choices. In contrast, Japanese monolinguals reacted more quickly in making path-match decisions than manner-match decisions. Meanwhile, Cantonese monolinguals had equal processing efficiency in making either manner- or path-match choices. And the results of by-item analysis further confirmed that cross-linguistic differences in processing efficiency were, to a large degree, consistent and systematic across each item, suggesting that the differences in processing efficiency was not a superficial or accidental effect, but as a result of language-specific patterns in motion event encoding. This result is backed up by a number of previous studies that language effects are most likely to appear when the stimuli are complex, or when the task involves higher-level processing and with a time limitation (Athanasopoulos & Bylund, 2013a; Filipovic, 2018; Trueswell & Papafragou, 2010). Thus, the different processing efficiency that participants had in similarity judgements can be interpreted as a consequence of language mediation.

Given the language-specific properties, in English (manner-salient language), as manner is expressed in the verb root/finite verbs (a marked grammatical category) and used with high frequency, the high manner codabiltiy may contribute to a higher cognitive salience in mental representations which increases its accessibility during cognitive processing (Slobin, 2000, 2004, 2006). According to the cognitive grammar (Langacker, 2008), language-specific ways of selecting and organizing information are directly related to how conceptualizations represented in cognition. As a result, speakers of different languages construe the same event in conceptually different ways depending on the grammatical devices made available in their language (Athanasopoulos, Bylund, et al., 2015; Bylund & Jarvis, 2011; Flecken et al., 2015; Von Stutterheim et al., 2012; Von Stutterheim & Nuse, 2003). As a consequence, linguistic structures highlighted by grammar (i.e. number, gender, aspect making, and finite verbs) tend to be placed with greater prominence in speakers' mental representations. For example, Von Stutterheim et al. (2012) examined how speakers of seven languages conceptualized motion events, with a particular focus on the effects of grammatical aspects on motion event construal. Using a triad-matching paradigm, participants verbalized all stimuli and before performed a non-linguistic categorization based on event trajectories or endpoints. The results suggested that speakers of languages with obligatory aspect marking tended to categorize motion based on the criteria of 'on-goingness' (trajectories), while speakers of languages that lack an aspect system categorized motion based on 'completion' (endpoints). The findings can be interpreted by a psycholinguistic model that 'conceptual categories encoded in the grammar of a language play an active role in the cognitive filer set up during attention allocation and information selection when taking about motion' (Von Stutterheim et al., 2012, p.863). In other words, language-specific ways of conceptualizing motion were closely related to the core grammatical categories made available during cognitive processing.

The direct link between grammatical knowledge and conceptual representations is also well reflected by the account of associative and attentional learning that attention is drawn towards form-meaning associations and form-meaning pairs that are highlighted by grammar (Smith & Samuelson, 2006). In other words, participants are more likely To assess the highlighted linguistic elements when perceiving and retrieving relevant information from their memory. Therefore, RT has been successfully used to indicate how fast a certain piece of information is retrieved from participants' memory (Collins & Quillian, 1969). With regard to the speed of information processing, English monolinguals may have attended to manner of motion at the first instance due to its higher salience. Although participants finally opted for path-match alternate, their RT to manner was much more quickly. In contrast, as Japanese (path-salient language) typically encodes path in the main verb whereas manner in subornation form such as gerunds, coverbs or adverbial expressions with relatively low codability, the easy access to path directed speakers' attention to path at the first instance. This may facilitate the information retrieval of path and processing efficiency in making path-match choices, but not that efficient for the processing speed for manner-match choices. As for Cantonese (manner and path salient language), where manner and path are typically expressed with equal salience, it is plausible to assume that manner and path were retrieved ‘in a parallel fashion’ with equal amount of attention being paid to both elements simultaneously in terms of overt selection and processing efficiency (Ji, 2019; Ji & Hohenstein, 2017, 2018).

The cross-linguistic differences observed in processing efficiency echoed with previous findings that the different grammatical salience may direct speakers’ attention to manner and/or path of motion when processing various types of motion events (Ji, 2017, 2019; Ji & Hohenstein, 2017, 2018). For example, in a very recent study, Ji and Hohenstein (2018) examined how children (3-year-olds and 8-year-olds) and adult speakers of Mandarin Chinese and English responded to visual stimuli of caused motion under a condition with verbal interference. Results suggested that participants’ RT to manner- and path-match choices presented clear cross-linguistic variations in their spatial cognition: English speakers reacted faster to manner-match choices than path-match choices whereas Chinese speakers reacted equally quickly to manner- and path-match choices. The authors attributed cross-linguistic differences in processing efficiency to the language-specific grammatical status of manner and path. For example, in English, manner is expressed in the marked verb forms (finite verbs) whereas path in satellite with less prominent grammatical status. Thus, speakers of English may habitually direct their attention to manner at the first instance. However in Mandarin Chinese, both manner and path are marked with equal grammatical status. As a result, Mandarin Chinese speakers can access manner and path in a ‘parallel fashion’. The findings were in line with the associative/attentional learning account that there is an association between the marked grammatical category (i.e. manner and path finite verbs) and non-linguistic recognition and perception. Thus, the frequent use of manner and path simultaneously can reinforce an overarching ‘presence’ of manner and path in one’s mental processing (Hohenstein, 2005; Smith & Samuelson, 2006). However, although the study of Ji and Hohenstein (2018) had a similar findings with the current study in terms of participants’ RT during the decision-making process, with regard to the experimental design, the former used verbal interference as a way to block participants’ access to language, either overtly or covertly, whereas the current study maximally boosted the active involvement of language during the decision making process. The current study tries to examine the cognitive processes taking place during speech production.

Similar findings were reported by other studies on the cognitive processing of various grammatical features such as word order, case marking, aspect marking, and animacy of subject or object (Baten et al., 2011; Boutonnet et al., 2012; Kempe, 1999; Rousselet, Fabre-Thorpe, & Thorpe, 2002; Tokowicz & MacWhinney, 2005). For example, Kempe (1999) examined cross-linguistic differences in morphological cues of case marking in Russian and German during online processing. The results suggested that although Russian and German speakers did not differ from each other in locating the correct agent in sentence interpretation, processing efficiency (as indicated by the RT) of Russian speakers was actually much faster than that of the German counterparts. The results indicate that grammatical categories across languages play an important role in modulating the magnitude of participants’ processing efficiency during the decision-making process.

### 10.1.2. Lexicalization and categorization of caused motion in monolinguals

#### 10.1.2.1. Lexicalization of caused motion

With regard to caused motion, participants' responses in event lexicalization were analysed in terms of the frequency of manner and path selection, semantic distribution (i.e. in the main verb or satellite) and framing strategies for event constructions. The results from the monolingual data confirmed the typological constraints of each language from different perspectives. For example, in terms of 'what to say' in information selection, all three groups reached a ceiling level path encoding, indicating that path is a core and central element in motion events (Talmy, 1985, 2000). However, with regard to the overall frequency of C-manner encoding, speakers of English (S-language) encoded manner of cause more frequently than Cantonese (E-language), while Japanese (V-language) presented the lowest frequency of specifying manner of cause in caused motion.

The cross-linguistic differences in information selection of caused motion can be attributed to the language-specific conflation patterns and availability of C-manner expressions in caused motion (Talmy, 2000). As a typical S-language, manner of cause (i.e. drag and kick), like manner (i.e. the manner in voluntary motion such as walk, run and jump), is most prototypically conflated in the verb root whereas path in the satellite (i.e. A boy pushed a box up hill). In addition, English also has a large set of vocabularies to express the specific manner of causation (Slobin, 2006; Talmy, 2000). Thus in English, manner verbs and path satellites exhibit a trend of complementary distribution. However, in Japanese, as cause of motion is frequently conflated with path in verb root, there is no obligatory syntactic slot for the encoding of manner of cause. Consequently, manner of cause can be easily added or dropped from the description. In addition, Japanese has a limited set of lexical devices for manner expressions (Matsumoto, 1996, 2017, 2018), such that speakers opted for the use of more general expressions to encode pure causation (i.e. carry, take, and move) rather specifying the exact manner of different causations (i.e. push, drag).

Qualitative analysis further confirmed that speakers of English and Japanese differed significantly in event constructions and framing strategies of caused motion. To be more specific, English-speaking participants consistently adopted a syntactically more compact structure of 'C-Manner verb+ Path satellite' (CMP), a typical satellite-framing strategy with higher level of information density, while the semantic distribution of manner of motion in Japanese was more diverse (cf. Table 15). For example, in contrast to English, speakers of Japanese used fewer CM verbs (16.39 %) but more subordination forms as gerund, subordination and adverbial clauses (57.59%) . In other words, encoding C-manner in subordination rather than the verb root allows Japanese to encode more components of caused motion (i.e. a dense or compact syntactic structure). These typical patterns of caused motion in English and Japanese are in line with previous literature that speakers of English and Japanese differed not only in 'what to say', but also 'how to say it', which reflect typological characteristics of English and other S-languages (Allen et al., 2007; Hendriks et al., 2008; Hickmann & Hendriks, 2006; Hickmann et al., 2018; Larrañaga et al., 2012; Montero-Melis & Bylund, 2017; Ochsenbauer & Engemann, 2011) as well as Japanese and other V-languages (Brown & Chen, 2013; Choi, 2011b, 2017; Choi & Bowerman, 1991; Park, 2019).

Incorporating language-specific properties of both English and Japanese, Cantonese has roughly equal amount of expressions following both satellite- and verb-framed conflation patterns. Qualitative data further confirmed that Cantonese speakers normally recruited the disposal '*zoeng construction'* (zoeng marker + CM verb + Path satellite), a typical satellite-framing strategy in caused motion, as illustrated in Example (32). Meanwhile, Cantonese speakers also frequently used 'zyu construction' of C-manner adjunct +Path verb, and 'Path-verb' construction, which are sub-categories of typical verb-framing strategies (cf. Figure 7). The typological status of Cantonese supports current findings that Cantonese displays properties of both satellite-framed and verb-framed languages in motion event descriptions (Yiu, 2013, 2014). The typological status of Cantonese can be attributed to the fact that Classical Chinese is undergoing typological transformations from being a V-language to an S-language (Lamarre, 2008; T. Li, 2013; Peyraube, 2006; Xu, 2006) and such typological changes in Cantonese have not been finished yet (Yiu, 2013, 2014). Such going transformations are well reflected in the description of caused motion. According to Yiu (2014), Cantonese to a certain degree, exhibits more verb-framed properties compared with Mandarin Chinese, as the former allows the conflation of causation with path in the main verb, which is regarded as ungrammatical in Mandarin Chinese. In addition, as illustrated in the voluntary motion, the degree of integration of manner and path is higher in Mandarin Chinese (**zou3jin4le**-jiao4shi4, **walk-enter-ASP**-classroom) compared with Cantonese (**haang4 zo2 jap6**-haak2 si1, **walk-ASP-into**-classroom). In other words, the grammaticalization of path of motion in Mandarin Chinese progresses further compared with Cantonese, as the degree of intergradation between manner and path is more profound in Mandarin Chinese.

In addition, like the linguistic findings reviewed in voluntary motion, the observed typological properties of caused motion lent further support for the Manner Salience Hypothesis that cross-linguistic differences in event lexicalization were not only found in the likelihood of manner selection (i.e. manner of the agent), but also applied to the likelihood of manner of causation in caused motion (Slobin, 2000, 2004, 2006).

#### 10.1.2.2. Categorization of caused motion

The second research question further probed how bilingual speakers in two language contexts conceptualized caused motion in comparison with monolingual controls for each language. The results suggested that on the one hand, participants had an overall preference for path-match alternates irrespective of language background in overt selection. However, the RT to manner- and path-match alternate was closely associated with language-specific lexicalization patterns, demonstrating a 'thinking-for-speaking' effect of language-specific features on event conceptualization.

With regard to the lack of cross-linguistic influence of different languages on overt selection, one possible explanation is that path is the core element of motion (i.e. the 'Basic Scheme for Motion') (Talmy, 1985, 2000, 2012). Previous studies reported that children demonstrated a universal cognitive salience towards path before fully acquiring the language-specific lexicalization patterns for motion events (Allen et al., 2007; Choi & Bowerman, 1991; Hohenstein, 2005; Ji & Hohenstein, 2018).

Another possible explanation for the lack of discrepancy might be attributed to the fact that inter-typological distinctions across languages are cline rather than categorical, such that cross-linguistic differences in lexicalization might not be clear-cut enough to cause absolute distinctions in non-linguistic categorization (Finkbeiner et al., 2002; Ji & Hohenstein, 2017; Loucks & Pederson, 2011). This view is well supported by the associative learning account in that 'the similarities that categorize any category are statistical, not necessary and sufficient' (Bylund & Athanasopoulos, 2014b; Smith & Samuelson, 2006).

In the current case, the caused motion stimuli used in the study involved three types of specific manners: agent manner (i.e. walking throughout all stimuli), object manner (i.e. either rolling or sliding), and manner of cause (i.e. push, pull, drag and kick). On the contrary, there were four types of path contrasts (i.e. up and down, into and out of, towards and away from, across and along). In addition, the trajectory of the agent and object remained the same across all stimuli. Thus, the strong path contrasts and congruent path between and agent and object may serve as a more convenient dimension for speakers to adopt the path-match dimensions when making their categorical selections.

The inconsistent findings between the categorical preferences in voluntary and caused motion are in line with studies of other cognitive domains such as objects and materials that cross-linguistic differences in cognition is a matter of relative degree, but not an absolute bias (Lucy, 1992a, 2014). As a consequence, preferences or statistical tendency towards one specific cognitive pattern never means total exclusion of the other (Bylund & Athanasopoulos, 2014b). And the degree of divergence or convergence towards one or more cognitive patterns depends on the degree of variations in learners' language learning trajectories.

Despite the lack of cross-linguistic differences in overt selection, the RT to manner- and path-match alternate displayed clear language-specific patterns: English monolinguals reacted more quickly in making C-manner-match choices than path-match choices. However, Japanese monolinguals reacted more quickly in making path-match choices than C-manner-match choices. Meanwhile, Cantonese monolinguals had equal efficiency in making either C-manner- or path-match choices. The results are line with previous findings that language-specific regularities in linguistic encoding tend to mediate participants' performances of processing efficiency in non-linguistic categorization (Gennari et al., 2002; Montero-Melis & Bylund, 2017; Wolff & Holmes, 2011).

In line with the ‘thinking for speaking’ and ‘thinking with language’ accounts (Slobin, 1996a) that the effects of language on cognition are most prominent when the experimental stimuli are complex or when the task has a time limitation (Filipovic, 2018; Trueswell & Papafragou, 2010). Thus, the different processing efficiency participants had in similarity judgements can be interpreted as a consequence of language mediation. In English, as C-manner is encoded in the main verb and used with high frequency, the high C-manner codabiltiy may contribute to a higher cognitive salience in mental representations and increases the accessibility to this element in cognitive processing (Slobin, 2004, 2006). Based on the theory of memory and attention and associative learning (Smith & Samuelson, 2006), attention is drawn towards form-meaning associations that was highlighted by grammar. Participants are more likely To assess the highlighted linguistic elements when perceiving and retrieving relevant information from memory. Thus, English monolinguals may have attended to manner of motion at the first instance due to its higher salience. Although participants finally opted for path-match alternate, their RT to manner was much more quickly. In contrast, as Japanese typically encodes path in the main verb whereas manner in subornation with relatively low codability, the easy access to path directed speakers’ attention to path at the first instance. This may facilitate the retrieval of path and processing efficiency in making path-match choices. As manner and path are typically expressed in a verb compound with equal salience in Cantonese, Cantonese speakers retrieved manner and path ‘in a parallel fashion’ with equal amount of attention being paid to both elements simultaneously (Ji & Hohenstein, 2017, 2018).

Given the discrepancies in the result of overt categorization and processing efficiency, it is important to adopt different types of measurements (overt selection and RT) when examining the dynamic relationship in linguistic relativity research.

## 10.2. Research Question # 2: Motion event lexicalization and categorization in monolinguals of Cantonese, English and early Cantonese-English bilingual speakers

The second research question further investigates whether the language used at the time of speaking had a ‘thinking-for-speaking’ effect on subsequent categorization for Cantonese-English bilinguals in different language contexts. I tested early Cantonese-English bilinguals in a monolingual (L1) and a bilingual context (L1 and L2) to explore whether recent L2 activation has immediate consequences for event categorization and the processing efficiency in making manner and path decisions. Based on the monolingual baselines in Research Question #1, it has been confirmed that Cantonese differs from English in both linguistic encoding (i.e. frequency of manner selection and event structures) and non-linguistic categorization (i.e. categorical preferences and RT). Instead of being a pure satellite- or verb-framed language, Cantonese is an equipollent-framed language with both satellite- and verb-framed properties. Given the typological differences between Cantonese and English, this section further explores whether the acquisition of an L2 English with a partially overlapping language system with the L1 brings about conceptual changes in the bilingual mind, and whether they are subject to short-term experimental manipulation (i.e. different language contexts). In addition, this section also addresses the modulating effects of other extra-linguistic variables, such as language contact and language proficiency, on different degrees of cognitive restructuring.

### 10.2.1. Lexicalization and categorization of voluntary motion in bilingual speakers

#### 10.2.1.1. Lexicalization of voluntary motion for bilingual speakers in different language contexts

The first sub-research question examined how Cantonese-English bilinguals in different language contexts lexicalized voluntary motion compared with monolinguals of each language (please refer to Research Question #1 for a detailed discussion of the monolingual data). Bilinguals' linguistic responses to voluntary motion in a monolingual and bilingual context were compared.

Statistical analyses showed that on the one hand bilinguals' overall performance differed significantly from Cantonese monolinguals but patterned with English monolinguals in 1) encoding manner of motion more frequently; 2) opting for a 'Manner verb +Path satellite' construction more frequently; and 3) using the satellite-framing strategies more frequently. On the other hand, bilinguals' performance patterned with each other irrespective of the language context. Their attested behaviours indicated a cognitive shift from the L1-based towards the L2-based patterns for lexicalization. Qualitative analysis of the linguistic descriptions further indicated that bilinguals chose to use more manner verbs with finer details and opted for less verb-framing strategies compared with Cantonese monolinguals.

Findings of the current study are in line with previous studies that cross-linguistic difference, or conceptualization transfer in event lexicalization is very likely to take place in L2 learners or bilingual speakers of two typologically different languages (Brown & Gullberg, 2008, 2011; Cadierno, 2008, 2010; Hohenstein et al., 2006; Larrañaga et al., 2012; Ruiz, 2006). In addition, the directionality of conceptualization transfer could be bi-directional, that is, it can take place not only from the L1 to the L2, but also from an opposite direction (backward transfer from the L2 to the L1). The directionality of conceptualization transfer depends on various factors in speaker's language learning trajectories, such as frequency of L2 use, age of L2 acquisition, and L2 proficiency (Athanasopoulos, Damjanovic, et al., 2015b; Brown & Gullberg, 2013; Jarvis & Pavlenko, 2008; Pavlenko, 2011). For example, L1 speakers of V-languages (e.g. Japanese or Spanish) learning an L2 S-language (e.g. English and Swedish) tend to encode manner more frequently in their L1 expressions as a result of an L2 learning, even for speakers of a moderate level of L2 proficiency (Bylund, 2009a; Ji, 2017, 2019; Park, 2019). These findings suggest that the learning of an L2 S-language may increase the salience of manner for L1 V-language speakers. The current findings are supported by Levelt's speech production model that conceptualization transfer could place at a macro-level of information selection and structuring (Levelt, 1989).

However, I didn't detect any effect of recent L2-activation on event lexicalization. There are two reasons that may account for this result. One possible explanation is that for early Cantonese-English bilinguals, it may be more efficient to rely on a common linguistic pattern (i.e. Manner verb + Path satellite) that works well for both languages. According to Talmy (2000), the most characteristic way of expressing a voluntary motion in English is to encode manner in the main verb whereas path in the satellite due to the grammatical constraint that a path particle cannot stand alone as an independent element. For example, sentence like '*A boy up the hill*' is regarded as ungrammatical in English. Although there is a set of path verbs in English, they are not often used in daily communication  (Slobin, 1996a). However, in Cantonese, as path verbs can be used independently, the encoding of manner is optional and depends on whether speakers opt for it or not. In addition, Cantonese speakers may rely on a wider range of event constructions to describe voluntary motion. For example, descriptions like '*A boy ascends the hill'* (Path-only type) and '*A boy walks up a hill*' (Manner + Path type) in Cantonese are both grammatically correct and frequently used in daily communication. Among different patterns of event encoding available for Cantonese but not for English, it may be more efficient for early bilinguals to draw on a single linguistic pattern which is workable in both languages and generalizable to a wide range of communication contexts.

In addition, previous findings show that language-specific linguistic patterns and concepts are established in early childhood (Hickmann & Hendriks, 2006; Ji & Hohenstein, 2018). And children of S-languages and V-languages start to display language-specific patterns of event lexicalization as early as 17 months to three years old (Allen et al., 2007; Choi, 2011b; Choi & Bowerman, 1991; Daller et al,. 2010; Hickmann, Hendriks, et al., 2009; Hohenstein et al., 2006; Ji & Hohenstein, 2018; Maguire et al., 2010). This can be attributed to the fact that early exposure to a second language may facilitate bilinguals to be more sensitive to the routinized patterns of linguistic descriptions in both languages. In the current case, both Cantonese and English are official languages in HK. Participants usually acquire an L2 at an average age of three years old and use two languages actively and simultaneously both at school, and in daily interaction. Thus, early exposure to and active use of these two languages may lead to a convergence of different linguistic categories into one 'work-for-all' linguistic pattern, which is consistent and will not be affected by short-term experimental manipulation (i.e. language context in the current case).

#### 10.2.1.2. Categorization of voluntary motion for bilingual speakers in different language contexts

The second sub-research question examined how bilinguals in different language contexts conceptualized motion events compared with the monolingual baselines. Two types of measurements were used: a categorical measurement of overt selection, and a continuous measurement of RTs. Bilinguals' responses in a monolingual and bilingual context were compared in terms of overt selection and the processing efficiency in making manner and path decisions.

With regard to the overt selection, results suggested that on the one hand, bilinguals' overall performances were significantly different from Cantonese monolinguals but patterned with English monolinguals in showing manner-match preferences in overt categorization. On the other hand, bilinguals' performances patterned with each other irrespective of the language context. In addition, as a co-verbal measurement of participants' processing efficiency, the RTs of bilinguals in language contexts showed that bilinguals patterned with English monolinguals in processing more quickly in making manner-match decisions than path-match decisions regardless of different language contexts. Their attested behaviours in both event categorization and processing efficiency indicated a cognitive shift from Cantonese-based towards the English-based patterns in event conceptualization.

The current findings are in line with the Manner Salience Hypothesis (Slobin, 1996a, 2000, 2004, 2006) that the higher codability of manner in lexicalization is associated with the accessibility to this information in cognition. Thus, the higher manner salience in linguistic encoding tends to modulate speakers' reliance on this component in the overt selection and the processing efficiency of manner in subsequent similarity judgements. The cross-linguistic differences observed in event categorization echo with previous findings that participants' lexicalization patterns are predictive of their conceptual preferences when the access to language is not blocked in the decision-making process (Lai et al., 2014; Montero-Melis & Bylund, 2017; Papafragou & Selimis, 2010; Soroli & Hickmann, 2010). One possible explanation is that when participants are asked to do a categorization task which lacks a correct answer, they might depend on all available resources, including recent linguistic experience to solve the task. In this case, as proposed by the 'thinking-for-speaking' or 'thinking-with-language' hypothesis, the involvement of language as a strategy for categorical perception exhibit language-specific features (Lupyan, 2012; Wolff & Holmes, 2011). Therefore, the different weight or salience speakers attach to each semantic element in linguistic encoding tends to trigger a language-specific way of thinking during the similarity judgment. Similar results are reported in previous findings that language-specific labels tend to modulate categorization when a certain linguistic element is made salient during or prior to event categorization (Lupyan, 2012; Montero-Melis et al., 2016; Papafragou & Selimis, 2010). Furthermore, it is suggested that the active learning and using of two languages in daily life give rise to the readjustment of original categories developed in the L1 and provides positive evidence for cognitive restructuring (Jarvis & Pavlenko, 2008).

However, I didn't detect any effect of recent L2-activation on event categorization. One possible explanation is that as explained in the linguistic  section, early Cantonese-English bilinguals have already established a common linguistic pattern (i.e. Manner verb + Path satellite) that works well for both languages. Given the speakers' language learning trajectories, this pattern may be stable and resistant to change in accordance with different language contexts. It is plausible to draw a conclusion that early exposure to and active use of two languages may lead to a convergence of different conceptual categories, which are not easily affected by short-term language manipulation.

Similar findings are reported in other cognitive processes such as event classification (Kersten et al., 2010; Lai et al., 2014) and memory recognition (Filipović, 2011). For example, Lai et al. (2014) reported that early English-Spanish bilinguals patterned with

L2-Spanish in event categorization regardless of the language in use. It is suggested that the effects of speaking on thinking are not limited to the languages at immediate operation, but also applicable to a common pattern of 'thinking-for-speaking' developed through a whole lifetime of experience. Similar results were reported by Filipović (2011) that early English-Spanish bilinguals showed the L2-based conceptualization patterns irrespective of the test language. It was indicated that bilinguals tended to opt for a 'whatever-works-in-both' approach in cognitive processing in terms of the processing costs and benefits. It was concluded that 'bilinguals do not seem to have two separate processing systems but rather an intertwined use of the two' (p. 481).

I am aware that not all studies reported the same results. For example, Berthele and Stocker (2017) examined the effect of language mode on event lexicalization with German-French bilinguals. Results showed that participants in a bilingual mode were more likely to converge towards a French pattern when both languages were kept active. One possible reason for the discrepancy may be attributed to the typological status of languages under investigation. In fact, the degree of discrepancy between Cantonese and English (and in general, between E- and S-languages) is not as contrastive as that between S-languages and V-languages in terms of manner salience. Strictly speaking, learning an L2-English is more about the activation of an already-existed pattern in Cantonese rather than the internalization of a brand-new linguistic frame. Under this view, the partial overlap between Cantonese and English reinforces the degree of convergence between these two languages compared with other language pairs under investigation.

### 10.2.2. Lexicalization and categorization of caused motion in bilingual speakers

#### 10.2.2.1. Lexicalization of caused motion with bilingual speakers in different language contexts

With respect to caused motion, results suggested that on the one hand, bilinguals' overall performance was significantly different from Cantonese monolinguals but patterned with English monolinguals in terms of both the frequency of C-manner selection (i.e. with high frequency) and the semantic distribution of C-manner and path of motion (C-manner verb and path satellite). Qualitative data further confirmed that bilingual speakers were more prone to adopt satellite-framing strategies and encoded manner of cause with more fine-grained details in their descriptions. One most typical feature in bilingual's linguistic descriptions was that there was a statistical significant decrease in using the Path-only construction (Causation is conflated with path whereas manner of cause is omitted), a typical characteristic of V-language speakers, indicating a cognitive shift from the L1-based towards the L2-based patterns of motion event lexicalization. On the other hand, similar to the results reported in voluntary motion, bilinguals' lexicalization of caused motion in a monolingual and bilingual context patterned with each other.

The consistent findings of voluntary motion and caused motion suggest that bilinguals of advanced L2 proficiency have fully acquired L2-based patterns of lexicalization due to early exposure to and active use of the L2 in daily communication and social interactions (Aveledo & Athanasopoulos, 2016; Bylund & Athanasopoulos, 2014a; Bylund et al., 2013). And their L2-based lexicalization patterns continue to exert an ongoing backward influence on their L1-based description of motion events. One possible reason for the backward transfer may be that given the typological differences between Cantonese and English in caused motion, there are a wide range of structures that are available to speakers of Cantonese, but not for English. For example, in Cantonese, descriptions like ‘A boy pushed a box uphill’ (satellite-framing) and ‘A boy *ascended a box uphill’ (verb-framing) are grammatical and both constructions are used frequently in daily communication. However, it is completely ungrammatical in English to conflate cause of motion with path in the verb root as English lacks a set of equivalent causation-path verbs (Tamly, 2000). As suggested in the discussion of voluntary motion, early exposure to and active use of English and Cantonese may lead to a convergence of different linguistic categories into one ‘work-for-all’ linguistic pattern, which is systematic across utterances and will not be affected by short-term experimental manipulation.

#### 10.2.2.2. Categorization of caused motion with bilingual speakers in different language contexts

With regard to the conceptualization of caused motion, the results from the bilinguals suggested that in terms of overt selection, participants in both a monolingual and a bilingual context had a path-match preference. In fact, as reviewed in the monolingual section, both monolinguals of Cantonese and English and bilingual speakers had an overall preference for path-match choices regardless of the language group. In contrast with the overt selection, processing efficiency in selecting manner- and path-match alternates did show a language-specific effect, that is, bilinguals patterned with English monolinguals yet differed from Cantonese monolinguals in reacting more quickly to manner-match alternate than path-match alternate irrespective of the language context, indicating an ongoing process of cognitive restructuring in the conceptual categories of caused motion.

The results indicate that early exposure to an L2 not only gave rise to the internalization of novel linguistic frames, but also an L2-specific way of thinking in processing efficiency of caused motion. In fact, bilingual speakers were able to reconstruct their conceptualization patterns towards the target language, even in the subtle form of cognitive processing, when provided with sufficient examples of language-specific form-meaning parings and event constructions (i.e. event structures and semantic representations). In the current case, although participants across different language groups had an overall preference of path-match alternate, their processing efficiency of manner- and path-match variants measured by the RT did show a language-specific effect. As mentioned in the monolingual section, manner and path of motion in Cantonese receives equal salience. As a result, Cantonese monolinguals were likely to put equal weight on manner and path in their recognition memory and followed a 'parallel processing' mode during event processing (Ji, 2019; Ji & Hohenstein, 2017, 2018; Rousselet et al., 2002) Thus, the RT to manner- and path-match alternates remained the same. In contrast, as English monolinguals attached more importance to manner of motion due to its prominently marked grammatical status (i.e. in the verb root), they attached more salience to manner of motion during cognitive processing (Langacker, 2000, 2008). Even though they finally opted for path-match alternate, their RT to manner was much more quickly. As for bilingual speakers, due to the frequent use and high proficiency, they can restructure their L1-based patterns towards the language during online processing where language is actively involved during the decision-making process. In line with the 'thinking 'for-speaking' and 'thinking-with-language' account, the convergent patterns of event lexicalization may serve as a meddler or strategy to aid information retrieval and processing during the categorization process. Thus, bilingual speakers not only have a convergent pattern for lexicalization, but also for event conceptualization when engaged in online language-recruiting activities (Lupyan, 2012; Slobin, 1996b, 2006; Wolff & Holmes, 2011).

The current findings are in line with previous studies that on the one hand, non-linguistic representations tend to be modulated by language-specific properties when the access to the target language is not blocked during or prior to event categorization (Athanasopoulos, Bylund, et al., 2015; Kersten et al., 2010; Lai et al., 2014; Montero-Melis et al., 2017); on the other, bi- and multilingual's conceptualization patterns are consistent and may not be susceptible to change with the language at operation (Filipović, 2011; Lai et al., 2014; Wang & Li, 2019). Given that conceptual representations within the bilingual mind are multi-modal and highly interactive, discrepancies in the current research findings can be attributed to a wide range of linguistic and extra-linguistic factors such as language pairs under investigation, different conceptual domains, particular experimental contexts, and various individual differences of the participants, which will be discussed in more detail in the next section.

### 10.3. Research Question # 3: Motion event lexicalization and categorization in monolinguals of Cantonese, English and Japanese, and Cantonese-English-Japanese multilingual speakers

Based on previous findings that early Cantonese-English bilinguals have developed a convergent mode of 'thinking-for-speaking' during event lexicalization and conceptualization due to early exposure to and frequent use of an L2, the third research question further examines whether the language(s) used at the time of speaking have an impact on subsequent categorization for Cantonese-English-

Japanese multilingual speakers in different language contexts as a consequence of L3 learning.

Following the same reasoning, Cantonese-English-Japanese multilinguals were randomly assigned to a monolingual (L1) and a bilingual context (L1 and L3) as a function of recent L3 activation. The task manipulation examines whether recent L3 activation has immediate consequences for event categorization and processing efficiency during the decision-making process. Given the typological differences between Cantonese and Japanese, it is interesting to explore further whether the acquisition of L3 Japanese in adulthood with a partial overlap language system with the L1-Cantonese brings about conceptual changes in the multilingual mind, and whether the process of cognitive restructuring is subject to short-term experimental manipulation (i.e. different language contexts). In addition, this section also addresses whether other extra-linguistic variables, such as the frequency of L2 and L3 use, as well as L2 and L3 proficiency, tend to affect the restructuring process.

### 10.3.1. Lexicalization and categorization of voluntary motion in multilingual speakers

#### 10.3.1.1. Lexicalization of voluntary motion for multilingual speakers in different language contexts

The first research question examined how multilinguals lexicalize voluntary motion in comparison with monolingual and bilingual controls and whether different short-term task manipulation (i.e. language context) affect the restructuring process. Linguistic patterns of motion event encoding were analysed in terms of the frequency of manner and path encoding, and the framing strategies used for event construction.

Monolingual data confirmed the typological status of each language and how speakers across different languages select and structure information in voluntary motion. For path encoding, participants across each group expressed path with a high frequency regardless of path types (trajectory vs. boundary-crossing), indicating that path is a core and obligatory element of event construction (Slobin, 2006; Talmy, 1985, 2000). However, the frequency of manner encoding certainly contrasted across each language. As a typical S-language, English expressed manner of motion more frequently than Cantonese (E-language), whereas Japanese (V-language) exhibited the lowest frequency. On the one hand, the differences in information selection can be attributed to language-specific conflation patterns in voluntary motion (Talmy, 2000). In Japanese, as the path of motion is conflated with the main verb, there is no obligatory syntactic slot for manner encoding. As a result, manner of motion can be easily added to or dropped into an expression in contrast to English. Although Cantonese follows a typical verb-framing conflation pattern, this construction is not used as frequently as in Japanese (Yiu, 2013, 2014). Qualitative analysis further confirmed that for events with various types of manner, English-speaking participants used semantically and syntactically dense structures with manner verbs and path satellites, even for manners with the lowest salience (i.e. walk). However, Japanese-speaking participants most often omitted this information when manner was less salient or could be easily inferred from the motion scenes. The results resonate with the Manner Salience Hypothesis, in that cross-linguistic differences in motion event lexicalization are only observed in the likelihood of manner encoding, and degrees of differences may stem from how salient the manners are perceived by different groups of speakers (Slobin, 2004, 2006; Soroli & Hickmann, 2010). Turning to bilingual speakers, the results suggested that Cantonese-English bilinguals followed English monolinguals in the frequency of manner selection and the framing strategies for event construction regardless of manner type. These results confirm the previous findings that early exposure to and active use of an L2 that has a partial overlap system with the L1 motivates learners to converge to a single lexicalization pattern that is compatible with both languages, indicating reverse conceptualization transfer from the L2-based concepts back to the L1 (Filipović, 2011; Lai et al., 2014; Wang & Li, 2019).

As a case in point, the results for Cantonese-English-Japanese multilinguals exhibited instances of reverse conceptual transfer in the L1 descriptions, both L2- and L3-based patterns of information selection and event construction. As manner of motion was highly encoded in the descriptions of Cantonese-English bilinguals, multilinguals with high proficiency in L3 presented a tendency to encode manner less frequently, indicating a reverse transfer from L3-based patterns to the L1, as the frequent omission of manner information is a typical feature for V-language speakers. In addition, as reviewed in section 8.2.1.1., Cantonese-English bilinguals have already established an L2-based pattern for event lexicalization. Thus, multilinguals with an L3 did not totally shake off their L2-based encoding patterns as they used slightly more manner verbs compared with Cantonese monolinguals, suggesting an in-between performance under the influence of both English and Japanese (cf. Figure 22).

There are two possible reasons that account for this. First, as mentioned in Chapter 4 and also the monolingual results, Cantonese is an E-language with properties of both S- and V-languages (Yiu, 2013, 2014). The partial overlap between the L1 and L2, and the L1 and L3, facilitates a simultaneous influence from both languages back on to the L1. Similar findings are shown by other studies, i.e. that mutual influence and interaction across different languages is a common linguistic practice in bilingual speakers or various types of L2 learners, especially for languages sharing a common set of linguistic properties (Daller et al., 2011; Hohenstein et al., 2006; Ji et al., 2011b; Ji & Hohenstein, 2014). For example, Ji and Hohenstein (2014) examined how English speakers of varying levels of lexicalized motion events in L2-Chinese. By using a video description task, results suggested that English learners were more prone to use the compact structure of ‘BA-construction’ when describing voluntary motion with a higher rate of accuracy compared with the loose structure of ‘Zhe-construction’. The authors attributed the high rate of accuracy to the fact that the compact BA construction in Chinese actually shares more similarities with English. Thus, the partial overlap between these two languages facilitated learners’ learning process in the L2. Following the same reasoning, the partial overlap between different languages also promotes the mutual interaction across different languages in the context of bilingualism and multilingualism, which may lead to a transfer phenomenon in either a forward or backward direction (Jarvis & Pavlenko, 2008; Pavlenko, 2011).

In addition, in the current case, as multilinguals have already achieved an advanced level in L3 and use Japanese frequently as one of the predominant languages in their daily interaction (cf. Table 3), the sufficient exposure to and active use of the L3 may accelerate the process of reverse restructuring back to the L1 (Bylund & Athanasopoulos, 2014a; Bylund et al., 2013; Hohenstein et al., 2006; Park & Ziegler, 2014).

However, similar to bilingual speakers, the results for multilingual speakers did not indicate a significant role being played by the language context in the lexicalization of voluntary motion. There are two possible reasons that account for it. One possible explanation for the results is that multilinguals’ lexicalization patterns are consistent and may not be susceptible to change due to short-term task manipulation, such as language in operation (Athanasopoulos, Bylund, et al., 2015; Filipović, 2011; Kersten et al., 2010; Lai et al., 2014). As reviewed in the bilingual section, it is suggested that the effects of speaking on thinking are not limited to the languages at immediate operation, they are also applicable to a common pattern of 'thinking-for-speaking' developed through a whole lifetime of experience. Thus, cognitive restructuring, by its very nature, is a long-term process and not subject to short-term task manipulation (Jarvis, 2011, 2016; Jarvis & Pavlenko, 2008).

Another possible explanation concerns processing efficiency. In the current case, Japanese and Cantonese share a partial overlap system in motion event encoding, that is, the path element can be used as an independent element without mentioning the manner of motion. Thus, in terms of processing costs and benefits, multilingual speakers may tend to opt for it as strategy so as not to mention manner of motion as a principle of economy. Thus, instead of keeping three linguistic systems with overlapping typological features, it is more economical and efficient for multilinguals to converge to simple patterns for thinking and speaking.

#### 10.3.1.2. Categorization of voluntary motion for multilingual speakers in different language contexts

The second sub-research question probed how multilinguals in different language contexts made similarity judgements of voluntary motion in comparison with bilingual and monolingual controls for each language. The results suggested that with regard to the overt section, monolinguals demonstrated a hierarchal decrease in manner-match preferences: English monolinguals and Cantonese-English bilinguals selected more manner-match variants than Cantonese monolinguals. Japanese monolinguals showed the lowest level of manner preferences. Multilinguals patterned with their Cantonese counterparts in the overt selection regardless of the language context, demonstrating a cognitive restructuring of thinking-for-speaking within the multilingual mind as a consequence of L3 learning.

The current findings show that a higher frequency of manner encoding is closely related to a higher level of cognitive salience for manner. As a consequence, different degrees of manner salience in event lexicalization tend to modulate subsequent categorization in a consistent and predictable manner. That is, a high degree of manner codability may direct speakers' attention to a more prominent element when perceiving and retrieving relevant information from memory, whereas a lower level of manner salience may reduce the association between language and categorization (Lai et al., 2014; Slobin, 2004, 2006). Similar findings have been reported by previous studies in that language-specific conventions made available in the linguistic task tend to mediate participants' immediate performance in a subsequent non-linguistic task (Gennari et al., 2002; Kersten et al., 2010; Montero-Melis & Bylund, 2017; Soroli & Hickmann, 2010; Wolff & Holmes, 2011).

One novel finding of the current study was that similar to the findings in event lexicalization, multilinguals' categorization patterns demonstrated traces of both L2- and L3-based patterns: they selected manner-match variants less frequently than bilinguals, but more frequently than Japanese monolinguals. Their in-between performance indicated an ongoing process of cognitive restructuring in their L1-based patterns of event conceptualization. It is suggested that multilinguals' conceptual representations are flexible and dynamic, thus supporting earlier work that bi- and multilingual learners are able to reconstruct their L1-based conceptualization patterns towards L2- or L3-based patterns when provided with sufficient instances of the target language (Bylund & Athanasopoulos, 2014a; Bylund et al., 2013).

However, another interesting point worth mentioning is that, unlike some of the previous studies (Athanasopoulos, Bylund, et al., 2015; Kersten et al., 2010; Lai et al., 2014) which demonstrated that bilinguals' cognitive behaviour can be manipulated by short-term language activation, (i.e. language in operation), that is, they may display L1-based patterns in L1 instructions, but L2-based ones in L2 instructions, multilingual speakers in the current study patterned with their Cantonese counterparts regardless of the language in operation (L1 condition vs. L1 and L3 condition). As a consequence, the various degrees of manner salience in lexicalization may be well reflected in participants' mental representations.

One possible explanation of the discrepancies in results might be attributed to the target languages under investigation. As Cantonese is an E-language with properties of both S- and V-languages (Yiu, 2013, 2014), the partial overlap of the linguistic system across three languages may facilitate a converged mode of thinking-for-speaking in the multilingual mind, which might be resistant to change as a function of short-term task manipulation (Filipović, 2011; Wang & Li, 2019).

In addition to the overt selection reviewed above, the RT of multilinguals' processing efficiency also indicated a process of restructuring during the process of cognitive processing. That is, bilinguals patterned with English monolinguals in reacting more quickly to manner-match alternate than path-match alternate, whereas Japanese monolinguals reacted more quickly to path-match alternates than manner-match ones. Meanwhile, multilinguals patterned with their Cantonese counterparts in reacting similarly quickly when making manner- and path-match decisions regardless of different language contexts, indicating an ongoing process of cognitive restructuring in conceptual categories as a consequence of L3 learning.

As mentioned in the monolingual section, for Cantonese monolinguals, given that the manner and path of motion receive equal salience based on the linguistic structures, Cantonese monolinguals were likely to put equal weight on manner and path in their recognition memory and to follow a ‘parallel processing’ mode during event processing (Ji, 2019; Ji & Hohenstein, 2017, 2018; Rousselet et al., 2002) Thus, the RTs to manner- and path-match alternates remained the same. In contrast, as Japanese monolinguals attached more importance to path of motion due to its prominently marked grammatical status (i.e. in the verb root), they attached more salience to manner of motion and followed a way of ‘sequential processing’ (Langacker, 2000, 2008; Rousselet et al., 2002). As for multilingual speakers, due to their frequent use of and high proficiency in L3-Japanese, they can restructure their L1-based patterns towards the language during online processing when the target language is actively involved during the decision-making process. In line with the ‘thinking ‘for-speaking’ and ‘thinking-with-language’ account, the convergent patterns of event lexicalization may serve as a meddler or strategy to aid information retrieval and processing during the categorization process. Thus, multilingual speakers have a convergent pattern not only for lexicalization, but also for event conceptualization when engaged in online language-recruiting activities or when their access to language is not blocked during the decision-making process (Frank, Everett, Fedorenko, & Gibson, 2008; Lupyan, 2012; Montero-Melis & Bylund, 2017; Montero-Melis et al., 2016; Papafragou & Selimis, 2010; Slobin, 1996b, 2006; Wolff & Holmes, 2011).

### 10.3.2. Lexicalization and categorization of caused motion with multilingual speakers

#### 10.3.2.1. Lexicalization of caused motion with multilingual speakers in different language contexts

With the linguistic encoding of caused motion, the first sub-research question examined how multilingual speakers in different language contexts lexicalized caused motion in comparison with bilingual and monolingual controls for each language. The results from the monolingual data confirmed the typological constraints of each language. For manner encoding, English (S-language) expressed manner of cause more frequently than Cantonese (E-language). Meanwhile, Japanese (V-language) presented the lowest frequency of manner encoding.

As reviewed in the previous section, results for monolingual speakers were in line with the typological status of each language. Turning to bilingual speakers, results suggested that bilinguals largely patterned with English monolinguals in both manner selection (i.e. with high frequency) and semantic distribution of manner and path of motion (i.e. manner in verb + path in satellite). This suggests that bilingual speakers had fully acquired the L2-based lexicalization patterns due to early exposure to and active use of the L2 in daily communication (Aveledo & Athanasopoulos, 2016; Bylund & Athanasopoulos, 2014a; Bylund et al., 2013).

For multilingual speakers in different language contexts, results suggested an ongoing process of restructuring from L1- and L2-based patterns of event lexicalization towards L3-based patterns as multilinguals with high proficiency in Japanese presented a tendency to encode manner less frequently, a typical characteristic of V-language speakers. I have ruled out the possibility that the lower frequency of manner encoding in L3 learners might be due to incomplete acquisition of the target vocabulary or the use of avoidance as a communication strategy, because they had already mastered all target manner expressions in their descriptions of the control items.

In addition, multilingual learners presented a divergence from L1- and L2-based patterns towards target L3-based ones in using the construction of ‘path verbs +C-manner subordinate’ when describing a boundary-crossing event, a construction that poses difficulties for learners with contrastive linguistic features (Daller et al., 2011).

There are two reasons that may account for this. First, as mentioned in Chapter 4, Cantonese is an E-language with properties of both S- and V-languages (Yiu, 2013). The most conventional way in Cantonese is a serial-verb construction (i.e. the zoeng-construction) to encode manner in a subordinate form, whereas path in the main verb is also frequently used in oral description. Therefore, the partial overlap between the L1 and L3 facilitates learners' acquisition of the target forms and also the mutual interaction across different languages (Ji et al., 2011b; Ji & Hohenstein, 2014, 2017, 2018). In addition, as multilinguals have already achieved an advanced level in the L3 and use Japanese frequently in their daily communication (cf. Table 1), the active use of language in daily interaction can accelerate the restructuring process towards the target linguistic forms (Bylund & Athanasopoulos, 2014a; Bylund & Jarvis, 2011; Carroll et al., 2011; Park & Ziegler, 2014). In terms of processing costs and benefits, instead of keeping three linguistic-systems with overlapping typological features, it is more economical and efficient for multilinguals to converge to a common pattern that is compatible with both languages (both the L1 and the L3) in their descriptions of caused motion.

#### 10.3.2.2. Categorization of caused motion with multilingual speakers in different language contexts

The second sub-research question probed how multilingual speakers in different language contexts conceptualized caused motion in comparison with bilingual and monolingual controls for each language. Two types of measurement were used: a categorical preference and RT. The results suggested that, on the one hand, participants preferred a path-match alternate irrespective of their language background and the language context in event categorization. However, the RT to manner- and path-match alternate was closely associated with language-specific lexicalization patterns, demonstrating a 'thinking-for-speaking' or 'thinking-with-language' effect. One possible explanation for the lack of language-specific properties in the overt selection might be that path is the core element in motion events (Tamly, 1985, 2000). Previous studies have reported that children demonstrated a cognitive salience towards path in non-verbal behaviours before fully acquiring language-specific patterns for motion event descriptions (Allen et al., 2007; Ji & Hohenstein, 2018). A second possible reason is that the inter-typological distinctions across languages are cline rather than categorical, such that cross-linguistic differences in lexicalization might not be clear-cut enough for absolute distinctions in non-linguistic categorization (Ji & Hohenstein, 2017; Loucks & Pederson, 2011; Pavlenko & Volynsky, 2015).

In contrast, the RTs of manner- and path-match selection presented clear language-specific patterns: English monolinguals reacted more quickly in making manner-match choices than path-match choices. However, Japanese monolinguals reacted more quickly in making path-match choices than manner-match choices. Meanwhile, Cantonese monolinguals had equal efficiency in making either manner- or path-match choices. In line with the 'thinking for speaking' and 'thinking with language' accounts (Slobin, 1996a), language-specific regularities made available in a linguistic task tend to mediate participants' performances in a subsequent non-linguistic task (Gennari et al., 2002; Montero-Melis & Bylund, 2017; Wolff & Holmes, 2011).

For bilingual speakers, the results suggested that bilinguals patterned with English monolinguals in reacting more quickly to manner-match alternates than path-match alternates, indicating that early exposure to an L2 not only gave rise to the internalization of novel linguistic frames, but also an L2-specific way of 'thinking for speaking' in event perception. Turning to multilingual speakers, results showed that proficient multilinguals shifted away from bilinguals in reacting more quickly when making manner-match decisions but patterned with Cantonese monolinguals in reacting equally quickly to both path- and manner-match alternates regardless of the language context, indicating an ongoing process of cognitive restructuring as a consequence of L3 learning.

Current findings suggest that bi- and multilingual learners are able to reconstruct their conceptualization patterns towards the target language when provided with sufficient examples of language-specific constructions (i.e. event structures and semantic representations). The findings are in line with previous studies that on the one hand, non-linguistic representations tend to be modulated by language-specific properties when the access to the target language is not blocked during or prior to event categorization; on the other hand, bi- and multilinguals' conceptualization patterns are consistent and may not be subject to short-term experimental manipulation (Filipović, 2011; Kersten et al., 2010; Lai et al., 2014; Lucy, 2016; Wang & Li, 2019).

### 10.4. Research Question #4: Factors that modulate cognitive restructuring in motion event lexicalization and categorization in the bi- and multilingual mind

The final research question further examined whether and how participants' individual differences in long-term effects of L2 and L3 learning interact with short-term language manipulation in modulating the process of cognitive restructuring in both verbal and non-verbal behaviours established in the L1. The target long-term biographical factors under investigations are L2/L3 proficiency and the amount of language contact bi- and multilinguals had with each of the languages they speak. The short-term effect is the language context (monolingual vs. bilingual context) as a function of recent activation of L2 and L3. Mixed-effects modelling was used to establish the correlations between various factors in learner's language learning trajectories with motion event lexicalization (i.e. the frequency of manner encoding and semantic distribution of manner verbs), as well as motion event conceptualization (i.e. the categorical choices of manner- and path-preferences and processing efficiency as measured by the RT).

### 10.4.1. Factors that modulate cognitive restructuring in motion event lexicalization and categorization in the bilingual mind

Results for bilingual speakers in both voluntary and caused motion suggest that, on the one hand, short-term experimental manipulation (i.e. language context) did not play a significant role in modulating bilinguals' conceptual restructuring in either event lexicalization or conceptualization. One the other hand, individual differences related to the effects of long-term language learning played a significant role in their cognitive restructuring process. That is, the amount of language contact bilinguals had with L2-English per day was positively associated with their language-specific performances and consistently related to bilinguals' restructuring processes. To be more specific, the more frequently English was used in daily communication, the more likely the speaker was to encode manner information in lexicalization and to show a manner-match preference in subsequent categorization. In the current case, given the partial overlap between L1-Cantonese (being both satellite-framing and verb-framing) and L2-English (satellite-framing) in linguistic encoding of motion, the frequent use of English tends to activate the satellite-framing properties of Cantonese, which may accelerate the restructuring process towards L2-based patterns in terms of both event lexicalization and conceptualisation. In this regard, the partial overlap between Cantonese and English motivated bilingual speakers to establish a convergent mode of thinking-for-speaking that is compatible with both languages, rather than retain a separate mode of thinking.

With regard to the long-term effects of L2 learning, the current findings suggest that early exposure to and active use of an L2 present learners with sufficient instances of language-specific patterns. For instance, with sufficient L2 exposure, bilinguals may understand that English attaches great importance to manner in linguistic encoding. The reinforcement of language-specific encoding patterns may strengthen the associations of language-related conceptual representations (Bylund & Jarvis, 2011; Jarvis, 2011). Given the theory of 'associative learning', mental representations of different conceptual categories in the bilingual mind are subject to continuous readjustment as a function of continuous language exposure and language use. In this view, the frequent use of language may strengthen established language-specific associations whereas inadequate language use may weaken them. Consequently, in the current study, participants who used English more frequently tended to exhibit categorization preferences based on manner-match variants as well as in their processing efficiency when making manner-match decisions. The results are in line with previous findings that 'frequent use of the forms directs attention to their functions, perhaps even making these functions especially salient on the conceptual level. That is, by accessing a form frequently, one is also directed to the conceptual content expressed by that form' (Berman & Slobin, 1994, p.640).

Similar findings were reported by previous studies, that language contact plays an important role in cognitive restructuring. That is, the more frequently learners use an L2, the more likely is their cognitive behaviour to shift away from the L1 and pattern with the target language (Athanasopoulos, Damjanovic, et al., 2015b; Bylund & Athanasopoulos, 2014a; Bylund et al., 2013; Flecken, Athanasopoulos, et al., 2015; Park, 2019). For example, Bylund et al. (2013) examined the cognitive patterns of event construal in bilinguals with Afrikaans as the L1 (a language without a grammatical aspect). The results suggested that the more frequently L2-English was used, the more likely participants were to use manner similarity as the criterion when categorizing voluntary motion. In addition to the cognitive domain of motion, the role of language contact is also well observed in other cognitive domains such as colour (Athanasopoulos, 2009; Athanasopoulos et al., 2010), object naming (Athanasopoulos, 2007; Pavlenko & Malt, 2011) and spatial cognition (Flecken, Athanasopoulos, et al., 2015; Park & Ziegler, 2014).

However, the current study did not report a significant role of language proficiency in modulating the process of cognitive restructuring. One possible explanation is that, as mentioned in the methodology section, Cantonese-English bilinguals in the current study had already achieved an advanced proficiency level in their L2 English due to their early exposure to and frequent use of English at school and in daily communication . This seemed to be well-supported by their self-reported scores of L2 proficiency (cf. Table 2), as most of the participants regarded themselves as proficient bilinguals with equal levels of proficiency in each of their languages.

#### 10.4.2. Factors that modulate cognitive restructuring in motion event lexicalization and categorization in the multilingual mind

With regard to multilingual speakers, results suggested that similar to bilingual speakers, short-term experimental manipulation (i.e. language context) did not play a significant role in modulating multilinguals' conceptual restructuring in either event lexicalization or conceptualization. And the degree of conceptual restructuring in both verbal and non-verbal tasks was associated with the amount of language use with L3-Japanese and L2-English language use on a daily basis. In other words, the more frequently participants used an L3 in daily communication and interaction, the more likely were L3-based linguistic and non-linguistic patterns to be displayed in the L1.

With regard to short-term language manipulation, the lack of an effect of language context suggests that cognitive restructuring in the bi- and multilingual mind is a long-term process and not be subject to short-term language mediation. One possible explanation is that from the perspective of processing costs and benefits, bi- and multilingual speakers tend to build up a convergent mode of thinking rather than retain each language in a separate mode for ease of cognitive processing (Filipović, 2011; Lai et al., 2014; Wang & Li, 2019). In the current case, Cantonese is an E-language and situated on the continuum of S- and V-languages. Thus, the partial overlap across each language may accelerate the process of convergence of different linguistic systems into a unitary one, as bi- and multilingual speakers are able to recognise similarities and differences across different languages due to an enhanced level of multilingual awareness (Cook & Li, 2016; Pavlenko, 2016).

However, for long-term effects of language learning, results suggested that the degree of conceptual restructuring and the source of reverse transfer in both verbal and non-verbal task were significantly modulated by the amount of L2 and L3 use.

This can be interpreted via the account of associative and attentional learning that language-specific representations are built up, or emerge, in an up-regulation fashion, due to continuous exposure to numerous form-meaning pairings (Lupyan, 2012; Smith & Samuelson, 2006). Thus, continuous exposure to novel events will facilitate learners to form new form-meaning associations based on the statistical regularity of the co-occurring associations in different contexts. From the perspective of L2 or additional language learning, the frequent exposure to and cumulative learning experience of the

L2- or L3-based associations may bring about constant restructuring of conceptual categories associated with the L1 (Athanasopoulos, Damjanovic, et al., 2015b).

In the current case, both the L2 and L3 were actively used in participants' daily interaction and for the academic purposes, and participants had achieved quite high levels of language proficiency in both their L2 and L3 (i.e. above the B2 level). Thus, L1-based patterns of lexicalization and conceptualization tended to exhibit more L2-based instances when the L2 was used more actively, but L3-based patterns when the L3 was used as the predominant language. Given the contrastive typological differences between L2 English (satellite-framed) and L3-Japanese (verb-framed), the frequent use of Japanese may activate L3-based associations in L1-Cantonese, whereas the frequent use of an L2 may hinder this process. Thus, the amount of language contact with or use of the L2 and L3 are on opposite sides in modulating the restructuring process in L1-based patterns in lexicalization and conceptualization.

The findings lend support to the associative learning account (Langacker, 2000, 2008), that the frequent use of the target linguistic forms may lead to the entrenchment corresponding conceptual categories. And the associations between language and conceptual representations may be strengthened by a high level of exposure to and frequent use of the target language.

The results are in line with previous studies that the more frequently a target language (or language-specific features) is used, the more likely participants are to exhibit associated conceptualization patterns (Athanasopoulos, Damjanovic, et al., 2015b; Bylund & Athanasopoulos, 2014a; Bylund et al., 2013; Bylund & Jarvis, 2011). For example, Bylund and Athanasopoulos (2014a) examined the effect of grammatical aspect on event categorization with multilingual speakers learning two aspect languages based on the context of multilingualism in Africa. The results showed a close link between the frequent use of aspect marking in linguistic encoding and attention allocation in categorization, indicating that aspect marking is an important grammatical cue to construe event and frequent language use could strengthen the associations between language-specific features and their corresponding conceptual categories.

However, the current findings did not report any effects of language proficiency on cognitive restructuring of motion cognition in the bi- and multilingual mind. There are two possible reasons that count for this. From the measurement perspective, one possible explanation could be related to the specific type of measurement and operationalization of language proficiency in different studies. Following Athanasopoulos et al. (2015) and other well-established research, the current thesis operationalized language proficiency as learner's general or global proficiency and it was measured by a self-assessed language background questionnaire. This type of measurement may not be sensitive enough to capture the subtle effects of this factor, particularly for learners with relatively high-level proficiency. As the effects of language proficiency on cognition is a complex issue (Bylund & Athanasopoulos, 2014a; Brown & Gullberg, 2012; Pavlenko, 2011), it is important to incorporate different types of measurements (i.e. self-reported scores and standardized proficiency tests) to assess both general and domain-specific knowledge of the specific linguistic property when addressing the possible effects of this factor in further research.

In addition, from the language threshold perspective, the other possible explanation is that the effect of language proficiency on cognition is not linear and not always positively correlated with the degree of cognitive restructuring (Brown & Gullberg, 2012). That is, there might be possible intervals or ranges that these effects are the most prominent. However, once the proficiency exceeds a certain point, the effects may level out (Bylund & Athanasopoulos, 2015a). In the current study, bi- and multilingual speakers were functional language users with high language skills and use their L2 and L3 frequently in daily communication and social interaction. It is highly likely that their proficiency level has already passed a certain point and its effects on cognition may not be visible any longer.

### 10.5. Conclusion and implications

The current study extends language-and-thought research to the domain of multilingualism and takes a first step in exploring the effects of additional language learning on cognitive restructuring in Cantonese-English-Japanese multilinguals in the domain of motion events. More specifically, it addresses the questions of whether, and to what extent, the acquisition of an L2-English (satellite-framed) in early childhood and an L3-Japanese (verb-framed) in adulthood recalibrates the lexicalization and conceptualization patterns acquired through the L1-Cantonese (equipollently-framed). In order to further explore how the long-term effects of language learning and the short-term effects of language manipulation interact with each other, and whether cognitive restructuring is a long or short-term process, the current study combines these two lines of enquiry and aims to address how early Cantonese-English bilinguals and Cantonese-English-Japanese multilinguals lexicalize and conceptualize motion events in different language contexts. Specifically, the current study examines how bi- and multi-linguals in a monolingual (L1) and a bilingual (L1 and L2; L1 and L3) context lexicalize and categorize motion compared with monolinguals of each language as a function of recent L2 or L3 activation. In addition, it also addresses how other long-term factors, such as language use with each language and L2/L3 proficiency, tend to affect bilinguals' performances in both event lexicalization and conceptualization.

Two tasks were manipulated with a cartoon-based verbal encoding and a non-verbal triads-matching task. Linguistic data were analysed in terms of frequency of manner selection and event structures. To further explore how language-specific patterns in lexicalization affect different levels of cognition, two types of measurements were used in event conceptualization: a categorical measurement of similarity judgements and a continuous measurement of RT.

The results for monolingual data confirmed the typological status of each language in both linguistic encoding and non-linguistic categorization. Instead of being either a satellite- or a verb-framed language, Cantonese is an equipollent-framed language with both satellite- and verb-framed properties. Results from the bi- and multilingual groups showed that bi- and multilinguals' L1-based patterns of linguistic encoding and categorization were under the influence of both L2- and L3-based patterns, indicating an effect of multiple language learning on the ongoing conceptual restructuring associated with the L1 in the bi- and multi-lingual mind.

In order to resolve the controversy over whether cognitive restructuring is a long- or short-term process, different language contexts (i.e. a monolingual and a bilingual language context) were manipulated to establish the correlations between various factors in learners' language learning trajectories with motion event lexicalization (i.e. the frequency of manner encoding and semantic distribution of manner verbs), as well as motion event conceptualization (i.e. the categorical choices of manner- and path-preferences and processing efficiency as measured by the RT).  Results showed that the degree of cognitive restructuring in bi- and multilingual speakers was not manipulated by short-term experimental manipulation or recent linguistic exposures such as language context; rather, it was modulated by long-term language learning effects, such as the amount of language contact with the L2 and L3 in daily communication or interaction.

The current findings demonstrate that learning an additional language may have different degrees of influence on bi- and multilinguals' native language when the target language is actively involved in the decision-making process. In other words, learners are able to acquire not only relevant linguistic structures of the target language, but also associated thinking patterns when provided with sufficient language-specific instances in daily communication (Athanasopoulos, Damjanovic, et al., 2015b; Bylund & Athanasopoulos, 2014a; Cadierno, 2010; Park, 2019). And the degree of cognitive restructuring is modulated by various factors in learners' language learning trajectories.

The current thesis explores the ongoing theoretical and methodological debates on the mechanisms via which language affects cognition and expands the boundaries of language-and-thought research to additional language teaching and learning. It will contribute to the current literature on the language-thought debate in the following ways.

First of all, the current study demonstrates that cross-linguistic differences in lexicalization can go beyond the linguistic domain and modulate cognitive representations when the access to language is not blocked during the decision-making process. On the whole, the current results lend support to the hypothesis of thinking-for-speaking and associative and attentional learning that the effects of language on thought are flexible and context-bound. The association between linguistic and non-linguistic performance can at least be observed when the language is actively involved in the decision-making process (i.e. online processing).

Second, this study demonstrates that the effects of speaking on thinking are not limited to the languages in immediate operation, but also applicable to a common pattern of 'thinking-for-speaking' developed through a whole lifetime of experience. Thus, cognitive restructuring, by its very nature, is a long-term process and may not be subject to short-term task manipulation and language mediation (Jarvis, 2011, 2016; Jarvis & Pavlenko, 2008). Rather, it is open to continuing changes as a function of the great variations in learners' language-learning history, such as proficiency and language use.

Third, this study has a pedagogical impact on second language teaching and learning by showing that the cognitive restructuring of previously established conceptual representations in SLA is an important outcome of L2 or additional language acquisition. Thus, the acquisition of a certain aspect of language may not only hinge on the acquisition of vocabulary and the grammatical structure itself, but also the interaction between constructions of linguistic meaning and conceptual representations of the external world. This may increase the awareness of L2 language teachers, SLA researchers, curriculum designers and policymakers of the fact that linguistic knowledge is inter-connected with a language-specific way of thinking and conceptual representations in the bi- or multilingual mind are highly interactive and dynamic. This illustrates the importance of adopting a multimodal approach to understand learner languages and highlights the role of multimodal pedagogical tools and techniques to bring linguistic knowledge and language-specific conceptual categories together as a means to facilitate language teaching and learning.

Previous research has documented the benefits of applying audio-visual tools in grammar and vocabulary teaching (Blyund & Athanasopoulos, 2015; Lantolf, 2010;

Montero Perez et al., 2013; Tyler, 2012). For example, Blyund and Athanasopoulos (2015) reported that constant audio-visual exposure to L2-based concepts could facilitate learners' cognitive restructuring towards the target language in the acquisition of new view frames and aspect marking. The findings suggest that the use of audio-visual materials, such as television and the Internet, helps engage language learners in the process of noticing the cross-linguistic differences in their perceptual judgments. In view of the current thesis, this domain of interest should be expanded to lexical representations of motion events.

The application of audio-visual tools in different conceptual domains is in line with the account of associative learning that conceptual representations are multimodal, dynamic and highly inter-correlated across different modalities (Bylund & Athanasopoulos, 2014a; Bylund et al., 2013; Bylund & Jarvis, 2011; Casasanto, 2008). Thus, in order to facilitate learners' internalization of new concepts and linguistic frames related to the external world, teachers should try to engage learners in a wide range of activities that allow them to access language-related novel concepts in the modalities that are not directly related to language (Athanasopoulos, Damjanovic, et al., 2015b; Jarvis & Pavlenko, 2008). For example, the multimodal activities may include exercises of grouping or classifying newly-acquired concepts or thoughts based on different conceptual categories (i.e. motion events) such that learners can pay enough attention to different aspects and dimensions of this novel way of thinking (i.e. the manner- or path-salience in the motion domain). In addition, Tyler (2012) suggest that 'visuals meant to provide memorable, meaningful representation for L2 learners' (Tyler, 2012, p.136). Thus, language teachers may choose to use dynamic video clips with visuals or sounds instead of static pictures or grammatical instructions when teaching novel concepts and a new way of thinking.

Although the current study adopted a triangulation of methods, combining different tasks across different modalities, namely linguistic data of motion event elicitation, non-linguistic data of motion event categorization, and co-verbal data of RT future research may combine the measurement of RT with the use of an eye-tracking technique to further explore participants' attention-allocation patterns during event perception. In addition, the sequence of experimental administration was that participants orally described caused-motion videos immediately prior to their nonlinguistic evaluation of the same scenes. This particular context may thus facilitate the influence that language exerts (implicitly or explicitly) on cognitive processes. Therefore, it would be interesting for future research to add a condition that serves as a check on whether differences are found when it is hard for the participants to use and get access to the target language. An identical task utilizing verbal interference is thus recommended as an additional condition to include in future research on the interplay between language and cognition. Also, in future research, other extra-linguistic factors, such as language proficiency and lengthen of immersion, need to be taken into consideration when examining the dynamic relationship between the progress of language learning and changes to the cognitive state in the bi- or multilingual mind. This will contribute to a more systematic and comprehensive picture on how individual differences may influence the process of cognitive restructuring with bilingual speakers and additional language learners.

## References


Allen, S., Özyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T., & Fujii, M. (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.

Ameel , E., Storms , G., Malt , B., & Sloman, S. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language, 52*, 309-329.

Aronin, L., & Singleton, D. (2012). *Multilingualism*. Amsterdam: John Benjamins.

Aske, J. (1989). *Path predicates in English and Spanish: A closer look.* Paper presented at the Annual Meeting of the Berkeley Linguistics Society.

Athanasopoulos, P. (2006). Effects of the grammatical representation of number on cognition in bilinguals. *Bilingualism: Language and Cognition, 9*, 89-96.

Athanasopoulos, P. (2007). Do bilinguals think differently from monolinguals? Evidence from non-linguistic cognitive categorisation of objects in Japanese-English bilinguals. *Selected papers on theoretical and applied linguistics, 17*(2), 338-345.

Athanasopoulos , P. (2007). Interaction between grammatical categories and cognition in bilinguals: The role of proi ciency, cultural immersion, and language of instruction. *Language and Cognitive Processes, 22*, 689-699.

Athanasopoulos, P. (2009). Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition, 12*(1), 83-95.

Athanasopoulos, P., & Albright, D. (2016). A Perceptual Learning Approach to the Whorfian Hypothesis: Supervised Classification of Motion. *Language Learning, 66*(3), 666-689.

Athanasopoulos, P., & Bylund, E. (2013a). 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. (2013b). The 'thinking' in thinking-for-speaking: Where is it? *Language, Interaction and Acquisition, 4*(1), 91-100.

Athanasopoulos, P., Bylund, E., Montero-Melis, G., Damjanovic, L., Schartner, A., Kibbe, A., & Thierry, G. (2015). Two Languages, Two Minds: Flexible Cognitive Processing Driven by Language of Operation. *Psychological Science, 26*(4), 518-526.

Athanasopoulos, P., Damjanovic, L., Burnand, J., & Bylund, E. (2015a). Learning to Think in a Second Language: Effects of Proficiency and Length of Exposure in English Learners of German. *Modern Language Journal, 99*(sS), 138-153.

Athanasopoulos, P., Damjanovic, L., Burnand, J., & Bylund, E. (2015b). Learning to Think in a Second Language: Effects of Proficiency and Length of Exposure in English Learners of German. *Modern Language Journal, 99*, 138-153.

Athanasopoulos, P., Damjanovic, L., Krajciova, A., & Sasaki, M. (2011). Representation of colour concepts in bilingual cognition: The case of Japanese blues. *Bilingualism: Language and Cognition, 14*(1), 9-17.

Athanasopoulos, P., Dering, B., Wiggett, A., Kuipers, J.-R., & Thierry, G. (2010). Perceptual shift in bilingualism: Brain potentials reveal plasticity in pre-attentive colour perception. *Cognition, 116*(3), 437-443.

Athanasopoulos, P., & Kasai, C. (2008). Language and thought in bilinguals: The case of grammatical number and nonverbal classification preferences. *Applied Psycholinguistics, 29*(1), 105-123.

Aveledo, F., & Athanasopoulos, P. (2016). Second language influence on first language motion event encoding and categorization in Spanish-speaking children learning L2 English. *International Journal of Bilingualism, 20*(4), 403-420.
Baayen, H., Davidson, D., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.
Baddeley, A. (1992). Working Memory: The Interface between Memory and Cognition. *Journal of Cognitive Neuroscience, 4*(3), 281-288.
Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders, 36*(3), 189-208.
Barner , D., Inagaki, S., & Li., P. (2009). Language, thought, and real nouns. *Cognition, 111*, 329-344.
Bassetti, B., & Cook, V. (2011). Language and cognition: The second language user. In V. J. Cook & B. Bassetti (Eds.), *Language and Bilingual Cognition* (pp. 143-190). Oxford, UK: Psychology Press.
Baten, K., Hofman, F., & Loeys, T. (2011). Cross-linguistic activation in bilingual sentence processing: The role of word class meaning. *Bilingualism: Language and Cognition 14*, 351–359.
Beavers, J., Levin, B., & Shiao. (2010). The typology of motion expressions revisited *Journal of Linguistics, 46*(2), 331-377.
Berlin, B., & Kay, P. (1969). *Basic Color Terms. Their Universality and Evolution*. Berkeley: UCP.
Berman, R. A., & Slobin, D. I. (1994). Narrative structure. In R. A. Berman & D. I. Slobin (Eds.), *Relating events in narrative: A crosslinguistic developmental study* (Vol. 39, pp. 84). Hillsdale, NJ: Lawrence Erlbaum.
Berthele, R., & Stocker, L. (2017). The effect of language mode on motion event descriptions in German–French bilinguals. *Language And Cognition, 9*(4), 648-676.
Boroditsky, L. (2001). Does Language Shape Thought? Mandarin and English Speakers' Conceptions of Time. *Cognitive Psychology, 43*(1), 1-22.
Boroditsky, L., Fuhrman, O., & McCormick, K. (2011). Do English and Mandarin speakers think about time differently? *Cognition, 118*(1), 123-129.
Boroditsky, L., Ham, W., & Ramscar, M. (2002). What is universal in event perception? Comparing English & Indonesian speakers',. In W. D. Gray & C. D. Schunn (Eds.), *Proceedings of the 24th Annual Conference of the Cognitive Science Society,* (pp. 136-141). Lawrence: Erlbaum.
Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Research, 1479*, 72-79.
Brown, A. (2015). Universal Development and L1-L2 Convergence in Bilingual Construal of Manner in Speech and Gesture in Mandarin, Japanese, and English. *Modern Language Journal, 99*(S1), 66-82.
Brown, A., & Chen, J. (2013). Construal of Manner in speech and gesture in Mandarin, English, and Japanese. *Cognitive Linguistics, 24*(4), 605-631.
Brown, A., & Gullberg, M. (2008). Bidirectional cross-linguistic influence in L1-L2 encoding of manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition, 30*(2), 225-251.
Brown, A., & Gullberg, M. (2010). Changes in encoding of path of motion in a first language during acquisition of a second language. *Cognitive Linguistics, 21*(2), 263-286.

Brown, A., & Gullberg, M. (2011). Bidirectional Cross-Linguistic Influence in Event Conceptualization? Expressions of Path among Japanese Learners of English. *Bilingualism: Language and Cognition, 14*(1), 79-94.

Brown, A., & Gullberg, M. (2013). L1-L2 convergence in clausal packaging in Japanese and English. *Bilingualism: Language and Cognition, 16*(3), 477-494.

Brown, R. W., & Lenneberg, E. H. (1954). A study in language and cognition. *The Journal of Abnormal and Social Psychology, 49*(3), 454-462.

Bylund, E. (2009a). Effects of Age of L2 Acquisition on L1 Event Conceptualization Patterns. *Bilingualism: Language and Cognition, 12*(3), 305-322.

Bylund, E. (2009b). Maturational Constraints and First Language Attrition. *Language Learning, 59*(3), 687-715.

Bylund , E. (2011). Segmentation and temporal structuring of events in early Spanish–Swedish bilinguals *International Journal of Bilingualism, 15*(1), 56-84.

Bylund, E., & Athanasopoulos, P. (2014a). Language and thought in a multilingual context: The case of isiXhosa. *Bilingualism: Language and Cognition, 17*(2), 431-441.

Bylund, E., & Athanasopoulos, P. (2014b). Linguistic Relativity in SLA: Toward a New Research Program. *Language Learning, 64*(4), 952-985.

Bylund, E., & Athanasopoulos, P. (2015a). Introduction: Cognition, Motion Events, and SLA. *Modern Language Journal, 99*, 1-13.

Bylund, E., & Athanasopoulos, P. (2015b). Introduction: Cognition, Motion Events, and SLA. In *Mod. Lang. J.* (Vol. 99, pp. 1-13).

Bylund, E., & Athanasopoulos, P. (2015c). Motion event categorisation in a nativised variety of South African English. *International Journal of Bilingual Education and Bilingualism*, 1-14.

Bylund, E., & Athanasopoulos, P. (2015d). Televised Whorf: Cognitive Restructuring in Advanced Foreign Language Learners as a Function of Audiovisual Media Exposure. *Modern Language Journal, 99*, 123-137.

Bylund, E., Athanasopoulos, P., & Oostendorp, M. (2013). Motion Event Cognition and Grammatical Aspect: Evidence from Afrikaans. *Linguistics: An Interdisciplinary Journal of the Language Sciences, 51*(5), 929-955.

Bylund, E., & Jarvis, S. (2011). L2 effects on L1 event conceptualization. *Bilingualism: Language and Cognition, 14*(1), 47-59.

Cadierno, T. (2008). Learning to talk about motion in a foreign language. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 239-275). London: Routledge.

Cadierno, T. (2010). Motion in Danish as a second language: Does the learner's L1 make a difference? In Z. Han & T. Cadierno (Eds.), *Linguistic relativity in SLA : Thinking for speaking*. Bristol, UK: Multilingual Matters.

Cadierno, T., & Ruiz, L. (2006). Motion events in Spanish L2 acquisition. *Annual Review of Cognitive Linguistics, 4*(1), 183-216.

Carroll, M., Von Stutterheim, C., & Nuese, R. (2011). The language and thought debate: A psycholinguistic approach. In T. Pechmann & C. Habel (Eds.), *Multidisciplinary Approaches to Language Production* (pp. 183-218).

Carroll, M., Weimar, K., Flecken, M., Lambert, M., & Von Stutterheim, C. (2012a). Tracing trajectories: Motion event construal by advanced L2 French-English and L2 French-German speakers. *LIA Language, Interaction and Acquisition, 3*(2), 202-230.

Carroll, M., Weimar, K., Flecken, M., Lambert, M., & Von Stutterheim, C. (2012b). Tracing Trajectories: Motion Event Construal by Advanced L2 French-English and L2 French-German Speakers. *LIA: Language, Interaction and Acquisition/Langage, Interaction et Acquisition, 3*(2), 202-230.

Casasanto, D. (2008). Who's Afraid of the Big Bad Whorf? Crosslinguistic Differences in Temporal Language and Thought. *Language Learning, 58*, 63-79.

Casasanto, D. (2016). A Shared Mechanism of Linguistic, Cultural, and Bodily Relativity. *Language Learning, 66*(3), 714–730.

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition, 106*(2), 579-593.

Chen , J.-Y., Su , J.-J., Lee , C.-Y., & O'Seaghdha, P. (2012). Linguistically directed attention to the temporal aspect of action events in monolingual English speakers and Chinese–English bilingual speakers of varying English proi ciency. *Bilingualism: Language and Cognition, 15*(2), 413-421.

Chen, L., & Guo, J. (2009). Motion events in Chinese novels: Evidence for an equipollently-framed language. *Journal of Pragmatics, 41*(9), 1749-1766.

Choi, S. (2011a). Language-specifcity of motion event expressions in young Korean children. *LIA Language, Interaction and Acquisition, 2*(1), 157-184.

Choi, S. (2011b). Language-specificity of Motion event expressions in young Korean children. *Language, Interaction & Acquisition, 2*(1), 157-184.

Choi, S. (2017). Language-specificity in motion expression: Early acquisition in Korean compared to French and English. In E. Veneziano & J. H (Eds.), *Variation in first language acquisition: languages, contexts, learners* (pp. 103-122). Amsterdam & Philadelphia: John Benjamins.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition, 41*(1), 83-121.

Chomsky, N. (1988). *Language and problems of knowledge: The Managua lectures* (Vol. 16): MIT press.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning & Verbal Behavior, 8*(2), 240-247.

Cook, V. (1992). Evidence for Multicompetence. *Language Learning, 42*(4), 557-591.

Cook, V. (1999). Going beyond the native speaker in language teaching. *TESOL Quarterly, 33*, 185–209.

Cook, V. (2002). Background to the L2 user. In V. Cook (Ed.), *Portraits of the L2 User* (pp. 1-28). Clevedon: Multilingual Matters.

Cook, V. (2003). 'The changing L1 in the L2 user's mind. In V. J. Cook (Ed.), *Effects of the Second Language on the First* (pp. 1-18). Clevedon: Multilingual Matters.

Cook, V. (2016). Premises of multi-competence. In V. Cook & L. Wei (Eds.), *The Cambridge Handbook of Linguistic Multi-competence* (pp. 1-25). New York Cambridge University Press.

Cook, V., Bassetti, B., Kasai, C., Sasaki, M., & Takahashi, J. A. (2006). Do bilinguals have different concepts? The case of shape and material in Japanese L2 users of English. *International Journal of Bilingualism, 10*(2), 137-152.

Cook, V., & Wei, L. (2016). *Cognitive consequences of multi-competence*. Cambridge: Cambridge: Cambridge University Press.

Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research, 28*(3), 369–382.

Daller, M. H., Treffers-Daller, J., & Furman, R. (2011). Transfer of conceptualization patterns in bilinguals: The construal of motion events in Turkish and German *. *Bilingualism: Language and Cognition, 14*(1), 95-119.

Engemann, H., Harr, A., K. , & Hickmann, M. (2012). Caused motion events across languages and learner types. In L. Filipovic & K. Jaszczolt, M. (Eds.), *Space and Time in Languages and Cultures: Linguistic Diversity* (pp. 263-287). John Benjamins, Amsterdam/Philadelphia.

Engemann, H., Hendriks, H., Hickmann, M., Soroli, E., & Vincent, C. (2015). How language impacts memory of motion events in English and French. *Cognitive Processing, 16*, S68-S68.

Evans, V., & Green, M. (2006). *Cognitive linguistics: An introduction*. Edinburgh: Edinburgh University Press.

Evans,V., Bergen, B.K., & Zinken,J. (2007). *Cognitve linguistic reader*. London: Equinox.

Everett, C. (2013). *Linguistic relativity: Evidence across languages and cognitive domains*. Berlin: De Gruyter Mouton.

Filipovic, L. (2018). Speaking in a second language but thinking in the first language: Language-specific effects on memory for causation events in English and Spanish. *International Journal of Bilingualism, 22*(2), 180-198.

Filipović, L. (2011). Speaking and remembering in one or two languages: bilingual vs. monolingual lexicalization and memory for motion events. *International Journal of Bilingualism, 15*(4), 466-485.

Filipović, L. (2018). Speaking in a second language but thinking in the first language: Language-specific effects on memory for causation events in English and Spanish. *International Journal of Bilingualism, 22*(2), 180-198.

Filipović, L., & Geva, S. (2012). Language-specific effects on lexicalization and memory of motion events. In L. Filipović & K. Jaszczolt (Eds.), *Space and time across languages and cultures Vol2: Language, culture and cognition* (pp. 269–282). Amsterdam, The Netherlands: John Benjamins.

Finkbeiner, M., Nicol, J., Greth, D., & Nakamura, K. (2002). The role of language in memory for actions. *Journal of Psycholinguistic Research, 31*(5), 447-457.

Flecken, M. (2011). Event conceptualization by early dutch-german bilinguals: Insights from linguistic and eye-tracking data. *Bilingualism: Language and Cognition, 14*(1), 61-77.

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.

Flecken, M., Carroll, M., Weimar, K., & Von Stutterheim, C. (2015). Driving along the Road or Heading for the Village? Conceptual Differences Underlying Motion Event Encoding in French, German, and French-German L2 Users. *Modern Language Journal, 99*(S1), 100-122.

Flecken, M., Gerwien, J., Carroll, M., & Von Stutterheim, C. (2015). Analyzing gaze allocation during language planning: a cross-linguistic study on dynamic events. *Language And Cognition, 7*(1), 138-166.

Flecken, M., Von Stutterheim, C., & Carroll, M. (2014). Grammatical aspect influences motion event perception: findings from a cross- linguistic non-verbal recognition task. *6*(1), 45-78.

Frank, M., Everett, D., Fedorenko, E., & Gibson, E. (2008). Number as a cognitivetechnology: Evidence from Pirah˜a language and cognition. *Cognition, 108*, 819-824.

Frederickson, N., & Cline, T. (2008). *Special educational needs, inclusion and diversity: A textbook.* Buckingham, UK: Open University Press.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition, 1*(2), 131-149.

Guillaume, T., Panos, A., Alison, W., Benjamin, D., & Jan-Rouke, K. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences, 106*(11), 4567.

Gullberg, M. (2011). Thinking, speaking and gesturing about motion in more than one language. In A. Pavlenko (Ed.), *Thinking and speaking in two languages* (pp. 143-170). Bristol, UK: Multilingual Matters.

Habel, C., & Tappe, H. (1999). Processes of segmentation and linearization in describing events. In K. Ralf & C. v. Stutterheim (Eds.), *Processes in Language Production* (pp. 117-152). Opladen: Westdeutscher Verlag.

Hamano, S. (1998). *The sound symbolic system of Japanese* Stanford, Calif : Center for the Study of Language and Information ; Tokyo Kurosio.

Han, Z., & Cadierno, T. (2010). *Linguistic relativity in SLA: thinking for speaking.* Bristol: Multilingual Matters.

Harnad, S. R. (1987). Psychophysical and cognitive aspects of categorical perception: A critical overview. In S. R. Harnad (Ed.), *Categorial perception: The groundwork of cognition* (pp. 1-28). Cambridge, UK: Cambridge University Press.

Heider, E. R., & Olivier, D. (1972). The Structure of the Color Space in Naming and Memory for Two Languages. *Cognitive Psychology, 3*(2), 337-354.

Hendriks, H., & Hickmann, M. (2011). Space in second language acquisition. In V. Cook & B. Bassetti (Eds.), *Language and Bilingual Cognition* (pp. 315-339). Hove, UK: Psychology Press.

Hendriks, H., & Hickmann, M. (2015). Finding One's Path Into Another Language: On the Expression of Boundary Crossing by English Learners of French. *Modern Language Journal, 99*(S1), 14-31.

Hendriks, H., Hickmann, M., & Demagny, A. C. (2008). How English native speakers learn to express caused motion in English and French. *Acquisition et Interaction en Langue Étrangère, 27*, 15-41.

Hickmann, M. (2003). *Children's discourse : Person, space and time across languages.* Cambridge: Cambridge University Press.

Hickmann, M., Engemann, H., Soroli, E., Hendriks, H., & Vincent, C. (2017). Expressing and categorizing motion in French and English: verbal and non-verbal cognition across languages. In I. Ibarretxe-Antunano (Ed.), *Motion and space across Languages and applications* (pp. 61-94). Amsterdam: John Benjamins.

Hickmann, M., & Hendriks, H. (2006). Static and dynamic location in French and in English. *First Language, 26*(1), 103-135.

Hickmann, M., & Hendriks, H. (2010). Typological constraints on the acquisition of spatial language in French and English. *Cognitive Linguistics, 21*(2), 189-215.

Hickmann, M., Hendriks, H., & Champaud, C. (2009). Typological constraints on motion in French child language. In J. Guo, E. Lieven, S. Ervin-Tripp, N. Budwig, Ş. Özçalışkan, & K. Nakamura (Eds.), *Crosslinguistic approaches to the psychology of language:*

*Research in the tradition of Dan Isaac Slobin* (pp. 209-224). New York: Psychology Press.

Hickmann, M., Hendriks, H., Harr, A., K., & Bonnet, P. (2018). Caused Motion across Child Languages: A Comparison of English, German, and French. *Journal of Child Language, 45*(6), 1247-1274.

Hickmann, M., Taranne, P., & Bonnet, P. (2009). Motion in First Language Acquisition: Manner and Path in French and English Child Language. *Journal of Child Language, 36*(4), 705-741.

Hohenstein, J. (2005). Language-related motion event similarities in English- and Spanish-speaking children. *Journal of Cognitive Development, 6*, 403-425.

Hohenstein, J., Eisenberg, A., & Naigles, L. (2006). Is He Floating across or Crossing Afloat? Cross-Influence of L1 and L2 in Spanish-English Bilingual Adults. *Bilingualism: Language and Cognition, 9*(3), 249-261.

Ibarretxe-Antuñano, I. (2009). Path salience in motion events. In E. J. Guo, N. Lieven, S. Budwig, Ervin-Tripp, K. Nakamura., & S. Özçalışkan. (Eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 403-414). New York: Psychology Press.

Inagaki, S. (2001). Motion Verbs with Goal PPs in the L2 Acquisition of English and Japanese *Studies in Second Language Acquisition, 23*(2), 153-170.

Inagaki, S. (2002). Japanese learners' acquisition of English manner-of-motion verbs with locational/directional PPs. *Second Language Research, 18*(1), 3-27.

Jackendoff, R. (1990). Semantic structures. Cambridge, M.A.: MIT Press.

Jarvis, S. (2007). Theoretical and methdological issues in the investigation of conceptual transfer. *Vigo International Journal of Applied Linguistics, 4,* 43-71

Jarvis, S. (2011). Conceptual transfer: Crosslinguistic effects in categorization and construal *Bilingualism: Language and Cognition, 14*(1), 1-8.

Jarvis, S. (2016). Clarifying the Scope of Conceptual Transfer. *Language Learning Journal, 66*, 608-635.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York/London: Routledge.

Ji, Y. (2017). Motion event similarity judgments in one or two languages: An exploration of monolingual speakers of English and Chinese vs. L2 learners of English. *Frontiers in Psychology, 8*, 909.

Ji, Y. (2019). Cognitive Representation of Spontaneous Motion in a Second Language: An Exploration of Chinese Learners of English. *Frontiers in Psychology*. doi:https://doi.org/10.3389/fpsyg.2019.02706

Ji, Y., Hendriks, H., & Hickmann, M. (2011a). Children's expression of voluntary motion events in English and Chinese. *Journal of Foreign Languages, 34*(4), 2-20.

Ji, Y., Hendriks, H., & Hickmann, M. (2011b). The expression of caused motion events in Chinese and in English: Some typological issues. *Linguistics, 49*(5), 1041-1077.

Ji, Y., Hendriks, H., & Hickmann, M. (2011c). How children express caused motion events in Chinese and English: Universal and language-specific influences. *Lingua, 121*(12), 1796-1819.

Ji, Y., & Hohenstein, J. (2014). The syntactic packaging of caused motion components in a second language: English learners of Chinese. *Lingua, 140*, 100-116.

Ji, Y., & Hohenstein, J. (2017). Conceptualising voluntary motion events beyond language use: A comparison of English and Chinese speakers' similarity judgments. *Lingua, 195*, 57-71.

Ji, Y., & Hohenstein, J. (2018). English and Chinese children's motion event similarity judgments. *Cognitive Linguistics, 29*(1), 45.

Kempe, V. (1999). Processing of morphological and semantic cues in Russian and German. *Language and Cognitive Processes, 14*(2), 129-171.

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.

Kurinski, E., & Sera, M. D. (2011). Does learning Spanish grammatical gender change English-speaking adults' categorization of inanimate objects? *Bilingualism: Language and Cognition, 14*, 203-220.

Lakoff, G. (1990). The Invariance Hypothesis: is abstract reason based on image-schemas? *Cognitive Linguistics, 1*(1), 39-74.

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.

Lamarre, C. (2007). The linguistic encoding of motion events in Chinese: With reference to cross-dialectal variation. *Typological studies of the linguistic expression of motion events, 1*, 3-33.

Lamarre, C. (2008). The linguistic categorization of deictic direction in Chinese - With reference to Japanese. In X. Dan (Ed.), *Space in Languages of China: Cross-linguistic, Synchronic and Diachronic Perspectives* (pp. 69-97).

Langacker, R. W. (1987). *Foundations of cognitve grammar,Vol1.* Stanford, CA: Stanford Uiveristy.

Langacker, R. W. (1991). *Foundations of cognitve grammar,Vol2.* Stanford, CA: Stanford Uiveristy.

Langacker, R. W. (2000). *Grammar and conceptualization. B*erlin: Mouton de Gruyter.

Langacker, R. W. (2008). *Cognitive grammar a basic introduction*. Oxford: Oxford University Press.

Larrañaga, P., Treffers-Daller, J., Tidball, F., & Ortega, M.-C. G. (2012). L1 transfer in the acquisition of manner and path in Spanish by native speakers of English. *International Journal of Bilingualism, 16*(1), 117-138.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. (1996). Perspective taking and ellipsis in spatial descriptions. In M. A. P. Bloom, L. N. Peterson, & M. F. Garrett (Eds.), *Language and space* (pp. 77-107). Cambridge, MA: MIT Press.

Levin, B. (1993). *English verb classes and alternations*. Chicago: University of Chicago Press.

Levinson, S. C. (2001). Covariation between spatial language and cognition. In *Language acquisition and conceptual development* (pp. 566-588): Cambridge University Press.

Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity*. Cambridge Cambridge University Press.

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.

Li, T. (2013). AN EMPIRICAL STUDY ON STATIONARY MOTION EVENTS IN MANDARIN CHINESE*. *International Journal of Cognitive Linguistics, 4*(2), 133-156.

Loucks, J., & Pederson, E. (2011). Linguistic and non-linguistic categorization of complex motion events. In J. Bohnemeyer & E. Pederson (Eds.), *Event representation in language and cognition* (pp. 108-133). Cambridge: Cambridge University Press.

Lucy, J., A. (1992a). *Grammatical Categories and Cognition: A Case Study of the Linguistic Relativity Hypothesis*: Cambridge University Press.

Lucy, J., A. (1992b). *Language Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis*: Cambridge University Press.

Lucy, J., A. (1996). The scope of linguistic relativity: An analysis and review ofempirical research. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 37-69). Cambridge, UK: Cambridge University Press.

Lucy, J., A. (2014). Methodological approaches in the study of linguistic relativity. In F. c. L. & P. u. M. (Eds.), *Multilingual cognition and language use: Processing and typological perspectives* (pp. 17-44). Amsterdam: John Benjamins.

Lucy, J., A. (2016). Recent Advances in the Study of Linguistic Relativity in Historical Context: A Critical Assessment. *Language Learning, 66*(3), 487-515.

Lucy, J., A., & Gaskins, S. (2001). Grammatical categories and the development of classification preferences: a comparative approach. In M. Bowerman & S. Levinson (Eds.), *Language Acquisition and Conceptual Development* (pp. 257-283). Cambridge: Cambridge University Press.

Lucy, J., A., & Gaskins, S. (2003). Interaction of language type and referent type in the development of nonverbal classification preferences. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 465-492). Cambridge, MA: MIT Press.

Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology, 3*, 54.

Lupyan, G., & Clark, A. (2015). Words and the World: Predictive Coding and the Language-Perception-Cognition Interface. *Current Directions in Psychological Science, 24*(4), 279-284.

Lupyan, G., Rakison, D., & McClelland, J. (2007). Language is not just for talking - Redundant labels facilitate learning of novel categories. *Psychological Science, 18*(12), 1077-1083.

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., Imai, M., Haryu, E., Vanegas, S., . . . Sanchez-Davis, B. (2010). A developmental shift from similar to language-specific strategies in verb acquisition: A comparison of English, Spanish, and Japanese. *Cognition, 114*(3), 299-319.

Malt, B., & Sloman, S. (2003). Linguistic diversity and object naming by nonnative speakers of English. *Bilingualism: Language and Cognition, 6*, 47-67.

Malt, B., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language, 40*(2), 230-262.

Matsumoto, Y. (1996). Subjective motion and English and Japanese verbs. *Cognitive Linguistics, 7*(2), 183-226.

Matsumoto, Y. (2017). Types of motion-event expressions and expressions of path in Japanese In Y. Matsumoto (Ed.), *The typology of motion expressions* (pp. 247-276). Tokyo: Kurosio Publishers.

Matsumoto, Y. (2018). Motion event descriptions in Japanese from typological perspectives. In P. Pardeshi & T. Kageyama (Eds.), *Handbook of Japanese Contrastive Linguistics* (pp. 273–290). Berlin, Boston: De Gruyter.

Matsumoto, Y., Akita, K., & Takahashi, K. (2017). The functional nature of deictic verbs and the coding patterns of Deixis: An experimental study in English, Japanese, and Thai In I. Ibarretxe-Antuñano (Ed.), *Motion and Space across Languages: Theory and applications* (pp. 59-122). John Benjamins.

Matthews, S. (2006). On serial verb constructions in Cantonese. In Y. Alexandra, Aikhenvald, & R. M. W. Dixon (Eds.), *Serial Verb Constructions. A Cross-Linguistic Typology* (pp. 69-87). Oxford & New York: Oxford University Press.

Matthews, S., & Yip, V. (2011). *Cantonese: A comprehensive grammar*. London: Routledge.

Montero-Melis, G., & Bylund, E. (2017). Getting the ball rolling: the cross-linguistic conceptualization of caused motion. *Language And Cognition, 9*(3), 446-472.

Montero-Melis, G., Eisenbeiss, S., Narasimhan, B., Ibarretxe-Antuñano, I., Kita, S., Kopecka, A., Bohnemeyer, J. (2017). Satellite- vs. Verb-Framing Underpredicts Nonverbal Motion Categorization: Insights from a Large Language Sample and Simulations. *Cognitive Semantics, 3*(1), 36-61.

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.

Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System, 41*, 720-739.

Munnich, E., Landau, B., & Dosher, B. A. (2001). Spatial language and spatial representation: A cross-linguistic comparison. *Cognition, 81*(3), 171-208.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115*, 39-57.

Ochsenbauer, A. K., & Engemann, H. (2011). The impact of typological factors in monolingual and bilingual first language acquisition: Caused motion expressions in English and French. *LIA Language, Interaction and Acquisition, 2*(1), 101-128.

Özçalışkan, Ş., & Slobin, D. I. (2003). Codability effects on the expression of manner of motion in Turkish and English. In.

Panayiotou, A. (2004a). Bilingual emotions: The untranslatable self. *Estudios de Sociolinguistica, 5*, 1-19.

Panayiotou, A. (2004b). Switching codes, switching code: Bilinguals' emotional responses to English and Greek. *Journal of Multilingual and Multicultural Developmental Psychology, 25*, 124-139.

Papafragou, A. (2002). Mindreading and Verbal Communication. *17*(1-2), 55-67.

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does Language Guide Event Perception? Evidence from Eye Movements. *Cognition, 108*(1), 155-184.

Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle, 'n' roll: The representation of motion in language and cognition. *Cognition, 84*(2), 189-219.

Papafragou, A., & Selimis, S. (2010). Event categorisation and language: A cross-linguistic study of motion. *Language and Cognitive Processes, 25*(2), 224-260.

Park, H. I. (2019). How do Korean–English bilinguals speak and think about motion events? Evidence from verbal and non-verbal tasks. *Bilingualism: Langauge and Cognition*. doi:https://doi.org/10.1017/S1366728918001074

Park, H. I., & Ziegler, N. (2014). Cognitive shift in the bilingual mind: Spatial concepts in Korean–English bilinguals. *Bilingualism: Language and Cognition, 17*(2), 410-430.

Pavlenko, A. (2003). Eyewitness memory in late bilinguals: Evidence for discursive relativity. *International Journal of Bilingualism, 7*, 257-281.

Pavlenko, A. (2011). Thinking and Speaking in Two Languages: Overview of the field. In A. Pavlenko (Ed.), *Thinking and speaking in two languages* (pp. 237-257). Bristol: Multilingual Matters.

Pavlenko, A. (2014). *The bilingual mind: And what it tells us about language and thought*. Cambridge: Cambridge University Press.

Pavlenko, A. (2016). Whorf's Lost Argument: Multilingual Awareness. *66*(3), 581-607.

Pavlenko, A., & Malt, B. C. (2011). Kitchen Russian: Cross-Linguistic Differences and First-Language Object Naming by Russian-English Bilinguals. *Bilingualism: Language and Cognition, 14*(1), 19-45.

Pavlenko, A., & Volynsky, M. (2015a). Motion Encoding in Russian and English: Moving Beyond Talmy's Typology. *99*(S1), 32-48.

Pavlenko, A., & Volynsky, M. (2015b). Motion Encoding in Russian and English: Moving Beyond Talmy's Typology. *Modern Language Journal, 99*(1 [Supplement]), 32-48.

Peyraube, A. (2006). Motion events in Chinese: Diachronic study of directional complements. In M. Hickmann & S. Robert (Eds.), *Space in languages: LIinguistic systems and cognitive categories* (pp. 121-135). Philadelphia: Benjamins.

Pederson, E. (2007). Cognitive linguistics and linguistic relativity. In D. Geeraerts & H. Cuyckens (Eds.), *Oxford handbook of cognitive linguistics* (pp. 1012-1044). New York: Oxford University Press.

Pourcel, S. (2009). Motion scenarios in cognitive processes. In V. Evans & S. Pourcel (Eds.), *New directions in cognitive linguistics* (pp. 371-394). Amsterdam & Philadelphia: John Benjamins.

Pulvermüller, F., Shtyrov, Y., & Ilmoniemi, R. (2005). Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience, 17*(6), 884.

Regier, T., & Kay, P. (2009). Color categories are culturally diverse in cognition as well as in language. *Cultural Research, 39*, 56-71.

Roberson, D. (2005). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory & Cognition, 28*(6), 977-986.

Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory & Cognition, 28*(6), 977-986.

Roberson, D., Davies, I., & Davidoff, J. (2000). Color categories are not universal: replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General, 129*, 369–398.

Rousselet, G., Fabre-Thorpe, M., & Thorpe, S. (2002). Parallel processing in high-level categorization of natural images. *Nature Neuroscience, 5*(7), 629.

Ruiz, L. (2006). Motion events in Spanish L2 acquisition. *Annual Review of Cognitive Linguistics, 4*(1), 183-216.

Sachs, O., & Coley, J. (2006). Envy and Jealousy in Russian and English: Labeling and Conceptualization of Emotions by Monolinguals and Bilinguals: Emotional Experience, Expression, and Representation. In A. Pavlenko (Ed.), *Bilingual minds : emotional experience, expression, and representation* (pp. 209-231). Clevedon: Multilingual Matters.

Samuelson, L. K. (2002). Statistical Regularities in Vocabulary Guide Language Acquisition in Connectionist Models and 15–20-Month-Olds. *Developmental Psychology, 38*(6), 1016-1037.

Shibatani, M. (2003). Directional verbs in Japanese. In E. Shay & U. Seibert (Eds.), *Motion, Direction and Location in Languages* (pp. 259-286). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Siok, W. T., Paul, K., William, S. Y. W., Alice, H. D. C., Lin, C., Kang-Kwong, L., & Li Hai, T. (2009). Language regions of brain are operative in color perception. *Proceedings of the National Academy of Sciences, 106*(20), 8140.

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

Slobin, D. I. (1996a). From"thought and language" to "thinking for speaking". In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70-96). Cambridge: Cambridge University Press.

Slobin, D. I. (1996b). Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani & S. A. Thompson (Eds.), *Grammatical constructions: Their form and meaning* (pp. 195-220). Oxford: Clarendon Press.

Slobin, D. I. (2000). Verbalised events: A dynamic approach to linguistic relativity and determinism. In S. Niemeier. & R. Dirven. (Eds.), *Evidence for linguistic relativity* (pp. 107-138). Amsterdam, The Netherlands: John Benjamins.

Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D.Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 157-192). Cambridge, MA: MIT Press.

Slobin, D. I. (2004). The Many Ways to Search for a Frog: Linguistic Typology and the Expression of Motion Events. In S. Strömqvist & L. Verhoeven (Eds.), *Relating events in narrative, Vol. 2. Typological and contextual perspectives* (pp. 219-257). Mahwah, NJ: Erlbaum.

Slobin, D. I. (2005). How people move: Discourse effects of linguistic typology. In A. Martinović-Zić & C. L. Moder (Eds.), *Discourse across languages and cultures* (pp. 191-206). Amsterdam/Philadelphia: John Benjamins.

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). Amsterdam & Philadelphia, PA: John Benjamins.

Slobin, D. I., & Hoiting, N. (1994). *Reference to movement in spoken and signed languages: Typological considerations.* Paper presented at the Annual Meeting of the Berkeley Linguistics Society.

Smith, L. B. (2010). *Learning How to Learn Words: An associative crane*. New York: Oxford University Press.

Smith, L. B., & Samuelson, L. (2006). An Attentional Learning Account of the Shape Bias: Reply to Cimpian and Markman (2005) and Booth, Waxman, and Huang (2005). *Developmental Psychology, 42*(6), 1339-1343.

Soroli, E. (2012). Variation in spatial language and cognition: exploring visuo-spatial thinking and speaking cross-linguistically. *Cognitive Processing, 13*(1), 333-337.

Soroli, E., & Hickmann, M. (2010). Language and spatial representations in French and in English: evidence from eye-movements. In G. Marotta, A. Lenci, L. Meini, & F. Rovai (Eds.), *Space in language* (pp. 581-597). Pisa, Italy: Editrice Testi Scientifici.

Soroli, E., Hickmann, M., & Hendriks, H. (2019). Casting an eye on motion events: Eye tracking and its implications for linguistic typology. In M. Aurnague & Stosić (Eds.), *The semantics of dynamic space in French: Descriptive, experimental and formal studies on motion expression* (pp. 297-353). Amsterdam: John Benjamins.

Soroli, E., Hickmann, M., Hendriks, H., Engemann, H., & Vincent, C. (2015). Language effects on spatial cognition? Cross-linguistic evidence and eye tracking. *NINJAL International Symposium: Typology and Cognition in Motion Event Descriptions, January 2015, Tokyo*.

Soroli, E., Sahraoui, H., & Sacchett, C. (2012). Linguistic encoding of motion events in English and French: Typological constraints on second language acquisition and agrammatic aphasia. *LIA Language, Interaction and Acquisition, 3*(2), 261-287.

Stocker, L., & Berthele, R. (2019). The roles of Language mode and dominance in French–German bilinguals' motion event descriptions. *Bilingualism: Language and Cognition*, 1-13. doi:doi:10.1017/S1366728919000294

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Grammatical categories and the lexicon. Language typology syntactic description* (Vol. 3, pp. 57-149). Cambridge: Cambridge University Press.

Talmy, L. (2000). *Toward a cognitive semantics*. Cambridge: 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.

Talmy, L. (2012). Main verb properties. *International Journal of Cognitive Linguistics, 3*(1), 1-23.

Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J. (2009). Unconscious effects of language-specific terminology on pre-attentive color perception. *Proceedings of the National Academy of Sciences, 106*, 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.

Toratani, K. (2016). 10. The L2 acquisition of Japanese Motion event descriptions by L1 English speakers: An exploratory study. In *Cognitive-Functional Approaches to the Study of Japanese as a Second Language* (Vol. 46).

Treffers-Daller, J., & Calude, A. (2015). The Role of Statistical Learning in the Acquisition of Motion Event Construal in a Second Language. *International Journal of Bilingual Education and Bilingualism, 18*(5), 602-623.

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.

Vanek, N. (2019). Changing Event Categorization in Second Language Users Through Perceptual Learning. *Language Learning*. doi:10.1111/lang.12377

Vanek, N., & Hendriks, H. (2015). Convergence of temporal reference frames in sequential bilinguals: Event structuring unique to second language users *Bilingualism: Langauge and Cognition, 18*(4), 753-768.

Verkerk, A. (2013). Scramble, Scurry and Dash: The Correlation between Motion Event Encoding and Manner Verb Lexicon Size in Indo-European. In *Language Dynamics and Change* (Vol. 3, pp. 169-217).

von Humboldt, W. (1963). *Humanist without portfolio: An anthology of the writings of Wilhelm*

*von Humboldt.* Detroit, MI: Wayne State University Press.

Von Stutterheim, C., Andermann, M., Carroll, M., Flecken, M., & Schmiedtova, B. (2012). How grammaticized concepts shape event conceptualization in language production: Insights from linguistic analysis, eye tracking data, and memory performance. *Linguistics, 50*(4), 833-867.

Von Stutterheim, C., Bouhaous, A., & Carroll, M. (2017). From time to space: The impact of aspectual categories on the construal of motion events: The case of Tunisian Arabic and Modern Standard Arabic. *Linguistics, 55*(1), 207-249.

Von Stutterheim, C., & Nuse, R. (2003). Processes of conceptualization in language production: language-specific perspectives and event construal. *Linguistics, 41*(5), 851-881.

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.

Whorf, B. L. (1956). *Language, thought, and reality.* Cambridge, MA: 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.

Wolff, P., & Malt, B. C. (2010). The language-thought interface: an Introduction. In P. Wolff & B. C. Malt (Eds.), *Words and the Mind: How Words Capture Human Experience* (pp. 3-15). Oxford: Oxford University Press.

Xu, D. (2006). *Typological change in Chinese syntax*. Oxford: Oxford University Press.

Yiu, C. (2013). Directional Verbs in Cantonese: A Typological and Historical Study. *Language and Linguistics, 14*(3), 511-569.

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). Berlin/Boston: De Gruyter Mouton.

Yoshinari, Y. (2016). Influence of L1 English on the descriptions of motion events in L2 Japanese with focus on deictic expressions. In *Cognitive-Functional Approaches to the Study of Japanese as a Second Language* (Vol. 46).

Zlatev, J., & Yangklang, P. (2004). A third way to travel: The place of Thai in motion event typology. In S. Strömqvist & L. Verhoeven (Eds.), Relating events in narrative (pp. 159-190). Mahwah, NJ: Erlbaum.

**Appendix A. Dynamic stimuli for voluntary motion in linguistic encoding-Control Item (N=18)**

| Item | Control Stimuli |
|---|---|
| 1. | A boy is walking in a white background (from left to right) |
| 2. | A boy is walking in a white background (from right to left) |
| 3. | A boy is running in a white background (from left to right) |
| 4. | A boy is running in a white background (from right to left) |
| 5. | A boy is jumping in a white background (from left to right) |
| 6. | A boy is jumping in a white background (from right to left) |
| 7. | A boy is hopping in a white background (from left to right) |
| 8. | A boy is hopping in a white background (from right to left) |
| 9. | A boy is marching in a white background (from left to right) |
| 10. | A boy is marching in a white background (from right to left) |
| 11. | A boy is crawling in a white background (from left to right) |
| 12. | A boy is crawling in a white background (from right to left) |
| 13. | A boy is skateboarding in a white background (from left to right) |
| 14. | A boy is skateboarding in a white background (from right to left) |
| 15. | A boy is roller-skating in a white background (from left to right) |
| 16. | A boy is roller-skating in a white background (from right to left) |
| 17. | A boy is cycling in a white background (from left to right) |
| 18. | A boy is cycling in a white background (from right to left) |

**Appendix B. Dynamic stimuli for caused motion in the linguistic encoding task-Control Items (N=18)**

| Item | Control Stimuli |
|---|---|
| 1. | A boy is pushing a box in a white background (from left to right) |
| 2. | A boy is pushing a wheel in a white background (from left to right) |
| 3. | A boy is pushing a chair in a white background (from right to left) |
| 4. | A boy is pushing a trolley in a white background (from right to left) |
| 5. | A boy is pulling a suitcase in a white background (from left to right) |
| 6. | A boy is pulling a toy car in a white background (from left to right) |
| 7. | A boy is pulling a boat in a white background (from right to left) |
| 8. | A boy is pulling a trolley in a white background (from right to left) |
| 9. | A boy is dragging a boat in a white background (from left to right) |
| 10. | A boy is dragging a toy car in a white background (from left to right) |
| 11. | A boy is dragging a suitcase in a white background (from right to left) |
| 12. | A boy is pulling a bundle of woods in a white background (from right to left) |
| 13. | A boy is kicking a ball in a white background (from left to right) |
| 14. | A boy is kicking a toy car in a white background (from left to right) |
| 15. | A boy is kicking a ball in a white background (from right to left) |
| 16. | A boy is pulling a toy car in a white background (from right to left) |
| 17. | A boy is rolling a wheel in a white background (from right to left) |
| 18. | A boy is rolling a bundle of woods in a white background (from left to right) |

**Appendix C. Dynamic stimuli for voluntary motion in the non-linguistic categorization task-Control Items (N=6)**

1. Contrast Manner of motion with Ground

| Item | Target | Manner-match alternate | Ground-match alternate |
|---|---|---|---|
| 1. | Swim across a river in Background A | Swim across a river in Background B | Surf across a river in Background A |
| 2. | Walk out of a room in Background A | Walk out of a room in Background B | Jump out of a room in Background A |
| 3. | Cycle towards a river in Background A | Cycle towards a river in Background B | Skateboard towards a river in Background A |

2. Contrast Path of motion with Ground

| Item | Target | Path-match alternate | Ground-match alternate |
|---|---|---|---|
| 1. | Swim across a river from right to left in Background A | Swim across a river from right to left in Background B | Swim across a river from near-screen side to afar in Background A |
| 2. | Jump away from a tree in Background A | Jump away from a tree in Background B | Jump towards a tree in Background A |
| 3. | Cycle out of a castle in Background A | Cycle out of a castle in Background B | Cycle into a castle in Background A |

**Appendix D. Dynamic stimuli for caused motion in the non-linguistic categorization task-Control Items (N=6)**

1. Contrast Manner of cause with Ground

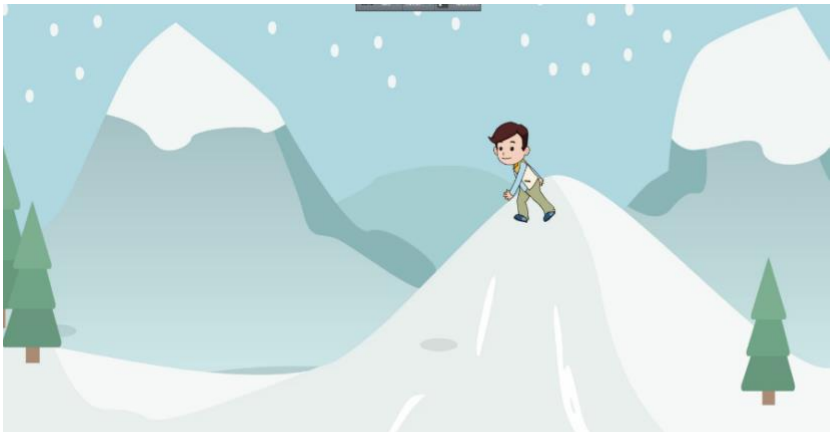
| Item | Target | C-Manner-match alternate | Ground-match alternate |
|---|---|---|---|
| 1. | Push a box up a hill in Background A | Push a box up a hill in Background B | Pull a box up a hill in Background A |
| 2. | Pull a boat across an icy river in Background A | Pull a boat across an icy river in Background B | Drag a boat across an icy river in Background A |
| 3. | Roll a wheel into a cave Background A | Roll a wheel into a cave Background B | Kick a wheel into a cave Background A |

2. Contrast Path of motion with Ground

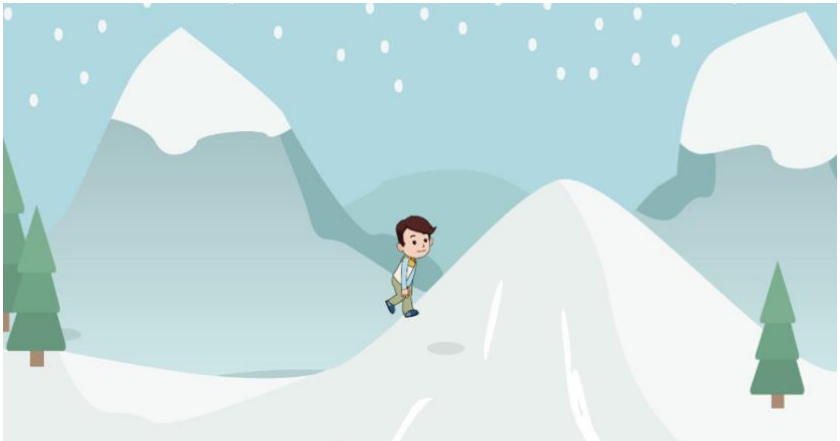
| Item | Target | Path-match alternate | Ground-match alternate |
|---|---|---|---|
| 4. | Push five woods away from fire in Background A | Push five woods away from fire in Background B | Push five woods towards fire in Background A |
| 5. | Push a trolley down a slope Background A | Push a trolley down a slope Background B | Push a trolley up a slope Background A |
| 6. | Drag a boat across the river from right to left in Background A | Drag a boat across the river from right to left in Background B | Drag a boat across the river from left to right in Background A |

**Appendix E. An illustration of the video stimuli in the non-linguistic task-test items-voluntary motion**

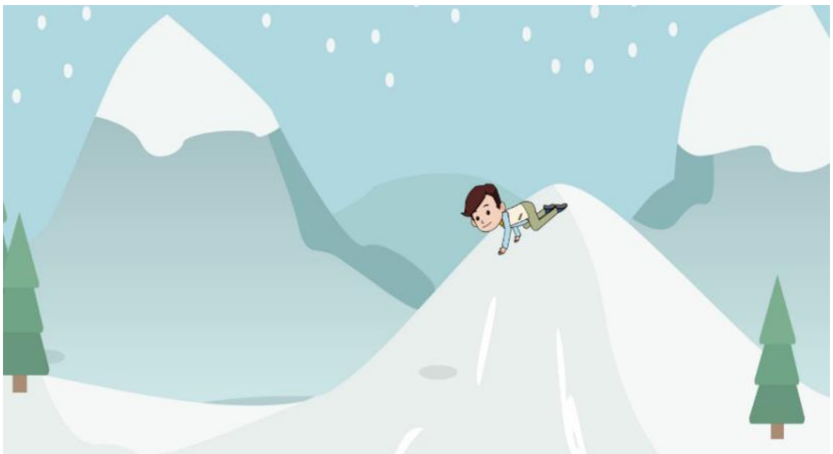
**Target Item: Tom walks down the snow hill.**



**Manner-match: Tom walks UP the hill.**



**Path-match: Tom CRAWLS down the hill.**



**Appendix F. An illustration of the video stimuli in the non-linguistic task-test items-caused motion**

**Target Item: Tom pulls a chair out of the room.**


**Manner-match: Tom pulls a chair INTO the room**


**Path-match: Tom PUSHES a chair out of the room**

**Appendix G. An illustration of the video stimuli in the non-linguistic task-control items-voluntary motion**

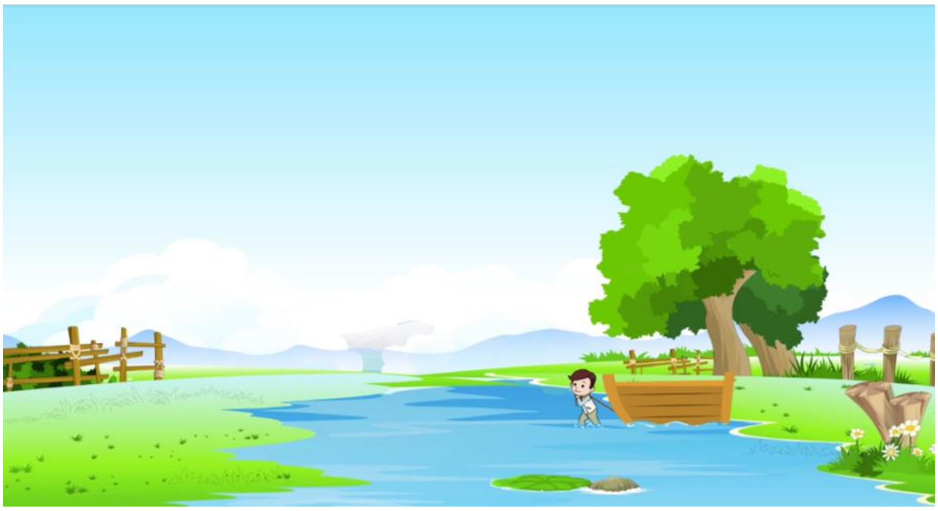
**Target Item: Tom swims across the river from right to left.**

**Path-match: Tom swims across the river from right to left in a NEW GOURND.**

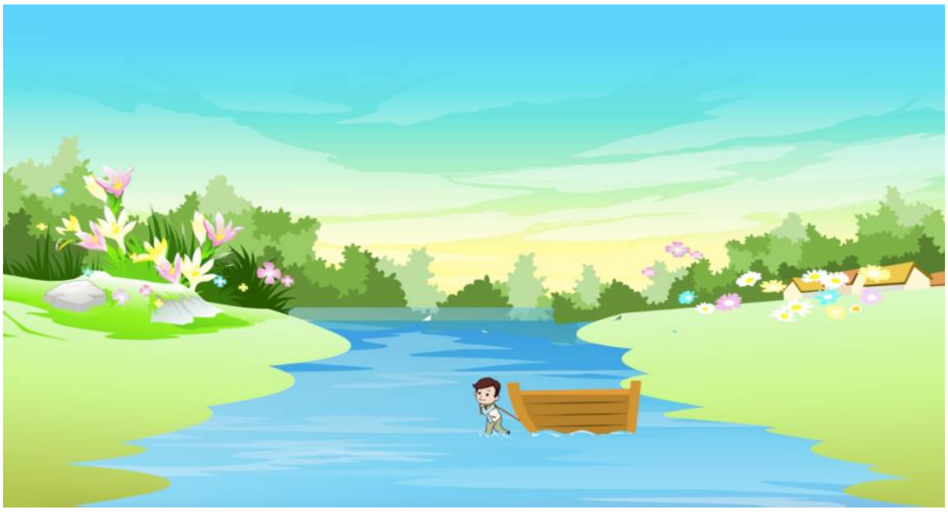
**Ground-match: Tom swims across the river AWAY FROM ME.**

**Appendix H. An illustration of the video stimuli in the non-linguistic task-control items-caused motion**

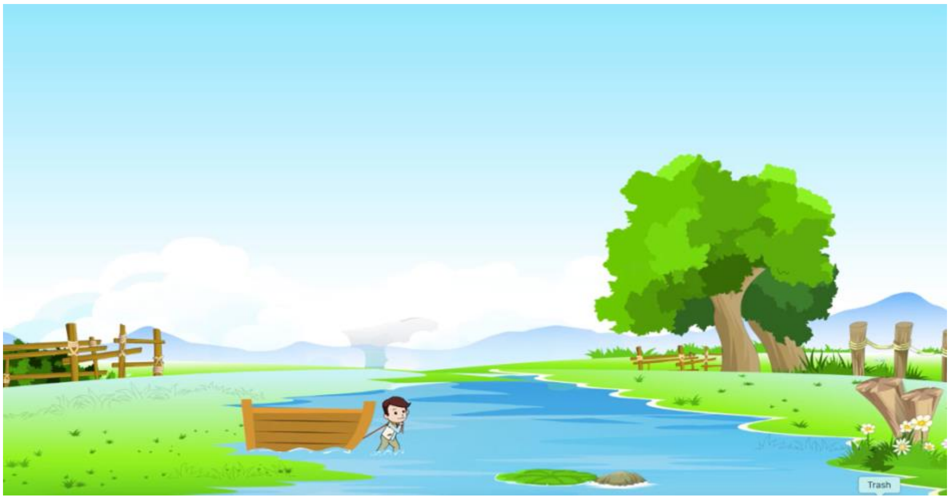
**Target Item: Tom drags a boat across the river from right to left.**



**Path-match: Tom drags a boat across the river from right to left in a NEW GROUND.**



**Ground-match: Tom drags a boat across the river FROM LEFT TO RIGHT.**



## Appendix I. Language background questionnaire-sample version for Cantonese-English bilingual speakers

### I. Personal Information

Gender: • Male • Female Age: ________________________________

Degree: • BA • MA • PhD Major: ______________________________

### II. Background of Language Learning

1. How many languages can you speak? _______________________________
2. Please define the languages you have learned and put them in a chronological order.

   First language (s) (mother tongue) ____________________________________

   Second language ________________________________________________

   Third language __________________________________________________

   Other languages (If any) ___________________________________________
3. Please rate your own proficiency in languages you know according to the following scale.

   Very poor Poor Limited Functional Good Very good Highly-advanced

   ______1________2______3_______ 4__________5_______6_________ 7______

| Language | speaking | listening | reading | writing |
|---|---|---|---|---|
| | | | | |
| | | | | |
| | | | | |

4. Have you attended any standardized language proficiency tests about the languages you have learnt? Please specify the tests and scores you have obtained.

| Language | Exam | Listening | Speaking | Reading | Writing | Overall |
|---|---|---|---|---|---|---|
| | | | | | | |
| | | | | | | |
| | | | | | | |

5. Have you travelled or studied in countries other than Hong Kong for more than three months? If yes, please indicate the name of the country, your length of residence and the language you use in daily communication.

| Name of the country | Length of Residence | Language |
|---|---|---|
| | | |
| | | |
| | | |

6. Please indicate the language used by teachers for formal classroom instructions at different educational levels.

Primary school__________________ Middle school ___________________________

Secondary school _______________ University______________________________

7. Please specify the total number of years you have spent on using the languages you have learnt.

| Languages you have learnt | Years of use |
|---|---|
| | |
| | |
| | |

8. Please indicate the age you use the languages you have learnt in the following environments:

| Language | At home | At school | At work | With friends | For entertainment |
|---|---|---|---|---|---|
| | | | | | |
| | | | | | |
| | | | | | |

9. Please estimate how many hours per day you spend in the following activities using the languages you have learnt.

| Activities | Language 1 (hrs) | Language 2 (hrs) | Language 3 (hrs) |
|---|---|---|---|
| Watching television | | | |
| Surfing the internet | | | |
| Reading for fun | | | |
| Reading for school | | | |
| Writing email to friends | | | |
| Writing for school | | | |

10. Please estimate how many hours per day you spend talking with the following group of people using the languages you have learnt:

| Group of people | Language 1 (hrs) | Language 2 (hrs) | Language 3 (hrs) |
|---|---|---|---|
| Family members | | | |
| Friends | | | |
| Classmates | | | |

**Thank you very much for your efforts and cooperation !**

**Appendix J. Equations for all mixed-effects models used in the current thesis**

A full list of statistical models used in the current study for the analysis of the linguistic encoding and non-linguistic categorization of motion events in monolingual, bilingual and multilingual speakers.

[1] model1<-glmer(FrequencyPath~Group+(1|Subject) +(1|Item), family=binomial, data=MonolingualVTask1, REML=FALSE)

[2] model2<-glmer(FrequencyManner~Group+(1|Subject) +(1|Item), family=binomial, data=MonolingualVTask1, REML=FALSE)

[3] model3<- glmer(FramingS ~Group+(1|Subject) +(1|Item), family=binomial, data=MonolingualVTask1, REML=FALSE)

[4] model4<- glmer(Choice~Group+(1|Subject) +(1|Item), family=binomial, data=MonolingualVTask2, REML=FALSE )

[5] model5<- lmer(log(RT) ~factor(Choice)*Group+(1|Subject)+(1|Item), data=MonolingualVTASK2, REML=FALSE)

[6] model6<-glmer(FrequencyPath~Group+(1|Subject) +(1|Item), family=binomial, data=MonolingualCTask1, REML=FALSE)

[7] model7<-glmer(FrequencyManner~Group+(1|Subject) +(1|Item), family=binomial, data=MonolingualCTask1, REML=FALSE)

[8] model8<- glmer(FramingS ~Group+(1|Subject), family=binomial, data=MonolingualCTask1, REML=FALSE)

[9] model9<- glmer(Choice~Group+(1|Subject) +(1|Item), family=binomial, data=MonolingualCTask2, REML=FALSE )

[10] model10<- lmer(log(RT) ~factor(Choice)*Group+(1|Subject)+(1|Item), data=MonolingualCTASK2, REML=FALSE)

[11] model11<- lmer(log(RT) ~factor(Choice)+(1|Subject)+(1|Item), data=ENGLISH, REML=FALSE)

[12] model12<- lmer(log(RT) ~factor(Choice)+(1|Subject)+(1|Item), data=JAPANESE, REML=FALSE)

[13] model13<- lmer(log(RT) ~factor(Choice)+(1|Subject)+(1|Item), data=CANTONESE, REML=FALSE)

[14] model14<-glmer(FrequencyManner~Group+(1|Subject) +(1|Item), family=binomial, data=BilingualVTask1, REML=FALSE)

[15] model15<-glmer(FrequencyPath~Group+(1|Subject) +(1|Item), family=binomial, data=BilingualVTask1, REML=FALSE)

[16] model16<- glmer(FramingS ~Group+(1|Subject) +(1|Item), family=binomial, data=BilingualVTask1, REML=FALSE)

[17] model17<- glmer(Choice~Group+(1|Subject) +(1|Item), family=binomial, data=BilinguallVTask2, REML=FALSE )

[18] model18<- lmer(log(RT) ~factor(Choice)*Group+(1|Subject)+(1|Item), data=BilingualVTASK2, REML=FALSE)

[19] model19<- lmer(log(RT) ~factor(Choice)+(1|Subject)+(1|Item), data=BilingualB, REML=FALSE)

[20] model20<- lmer(log(RT) ~factor(Choice)+(1|Subject)+(1|Item), data=Biln gualM, REML=FALSE)

[21] model21<-glmer(FrequencyManner~LanguageContext+Amount of English use+ English proficiency +(1|Subject) +(1|Item), family=binomial, data=BilingualTask1)

[22] model22<-glmer(MannerPreference~LanguageContext+Amount of English use+ English proficiency +(1|Subject) +(1|Item), family=binomial, data=BilingualTask2)

[23] model23<-lm(RTdifferences~LanguageContext+Amount of English use+ English proficiency, data=BilingualTask2)

[24] model24<-glmer(FrequencyManner~Group+(1|Subject) +(1|Item), family=binomial, data=BilingualCTask1, REML=FALSE)

[25] model25<-glmer(FrequencyPath~Group+(1|Subject) +(1|Item), family=binomial, data=BilingualCTask1, REML=FALSE)

[26] model26<- glmer(FramingS ~Group+(1|Subject) +(1|Item), family=binomial, data=BilingualCTask1, REML=FALSE)

[27] model27<- glmer(Choice~Group+(1|Subject) +(1|Item), family=binomial, data=BilinguallCTask2, REML=FALSE )

[28] model28<- lmer(log(RT) ~factor(Choice)*Group+(1|Subject)+(1|Item), data=BilingualCTASK2, REML=FALSE)

[29] model29<- lmer(log(RT) ~factor(Choice)+(1|Subject)+(1|Item), data=Bilingual-M, REML=FALSE)

[30] model30<- lmer(log(RT) ~factor(Choice)+(1|Subject)+(1|Item), data=Bilingual-B, REML=FALSE)

[31]model31<-glmer(FrequencyManner~LanguageContext+Amount of English use+ English proficiency +(1|Subject) +(1|Item), family=binomial, data=BilingualCTask1)

[32] model32<-lm(RTdifferences~LanguageContext+Amount of English use+ English proficiency, data=BilingualCTask2)

[33] model34<-glmer(FrequencyManner~Group+(1|Subject) +(1|Item), family=binomial, data=MultilingualVTask1, REML=FALSE)

[34] model34<-glmer(FrequencyPath~Group+(1|Subject) +(1|Item), family=binomial, data=MultilingualVTask1, REML=FALSE)

[35] model34<-glmer(FramingS~Group+(1|Subject) +(1|Item), family=binomial, data=MultilingualVTask1, REML=FALSE)

[36] model36<- glmer(Choice~Group+(1|Subject) +(1|Item), family=binomial, data=MultilingualVTask2, REML=FALSE )

[37] model37<- lmer(log(RT) ~factor(Choice)*Group+(1|Subject)+(1|Item), data=MultilingualVTASK2, REML=FALSE)

[38] model38<- lmer(log(RT) ~factor(Choice)+(1|Subject)+(1|Item), data=Multilingual B, REML=FALSE)

[39] model39<- lmer(log(RT) ~factor(Choice)+(1|Subject)+(1|Item), data=Multilingual M, REML=FALSE)

[40] model40< glmer(FrequencyManner~LanguageContext+Amount of English use+ English proficiency+ Amount of Japanese use + Japanese proficiency +(1|Subject) +(1|Item), family=binomial, data=MultilinguallTask1)

[41] model41<-glmer(MannerPreference~LanguageContext+Amount of English use+ English proficiency+ Amount of Japanese use + Japanese proficiency +(1|Subject) +(1|Item), family=binomial, data=MultilingualTask2)

[42] model42<-lm(RTdifferences~LanguageContext+Amount of English use+ English proficiency+ Amount of Japanese use + Japanese proficiency, data=MultilingualTask2)

Institute of Education

# Cognitive restructuring in multilingual mind: motion event construal in Cantonese-English-Japanese multilingual speakers

Jun 2017-Sep 2019

## Information sheet for the Chinese-English bilingual group

**Who is conducting the research?**

My name is Wang Yi and I am inviting you to take in part in the research project: **Cognitive restructuring in multilingual mind: motion event construal in Cantonese-English-Japanese multilingual speakers.**

I am a PhD student from the department of Applied Linguistics, Institute of Education, University College London. The UCL Institute of Education is well-known for its world-leading scholars and research excellence. The research project is my PhD project and it is a self-funded project. The main research findings will be reported in my PhD thesis. I am the one who conduct the research and have plenty of experience on conducting research in the field of language acquisition through doing a Master degree and working as a research assistant.

I am hoping to find out the interplay between language and thought in the context of bilingualism and multilingualism. And further cultivate the possible factors that may influence the interaction between language and thought.

I very much hope that you would like to take part. This information sheet will try and answer any questions you might have about the project, but please don't hesitate to contact me if there is anything else you would like to know.

**Why are we doing this research?**

The research is of great importance as it can deepen our understanding of the interplay between language and cognition in a bi- and multilingual setting and shed some light on language teaching in a multilingual classroom.

**Why am I being invited to take part?**

We are now recruiting 18 participants who can speak two languages from early childhood: Chinese and English. If you are proficient speaker of both languages, you are the target participants we want in the study.

**What will happen if I choose to take part?**

The study will be conducted on a laptop in a quiet room and will be finished within 20 minutes. The participants will be asked to complete two tasks and a language background questionnaire. In each task, participants watch the video clip first and then describe 'what happened' to the researcher. The participants will be informed that they are required to take part in all activities.

**Will anyone know I have been involved?**

Participants will be kept anonymous throughout the study and they will only be identified by their study codes. A unique study ID automatically generated by a software will be allocated to each participant prior to the study and the study code/ID will keep the participants from being recognized.

**Could there be problems for me if I take part?**

As the study doesn't deal with any sensitive issues, there won't be potential risks for the participants. But participants still reserve the right to stop at any time they want without reporting the reasons.

**What will happen to the results of the research?**

Primary results will be reported in the upgrading documents. Final results will be reported in the PhD thesis and other academic sources, such as academic conferences and research publications. A brief summary of the results will be disseminated to the participants after the project is completed. Participants will be kept anonymous throughout the study. All sources of data will be kept secure: the digital data will be stored in the personal laptop and the UCL desktop space with security codes; the manual personal data will be stored in the locked units.

**Do I have to take part?**

It is entirely up to you whether or not you choose to take part. We hope that if you do choose to be involved then you will find it a valuable experience. If you choose not to take part, we will fully respect your choices and there won't be any negative effects on you.

**Thank you very much for taking the time to read this information sheet.**

**If you would like to be involved, please complete the following consent form and return to dtnvywa@ucl.ac.uk.**

**If you have any further questions before you decide whether to take part, you can reach me at dtnvywa@ucl.ac.uk.**

**This project has been reviewed and approved by the UCL IOE Research Ethics Committee.**

Institute of Education

UCL

# Cognitive restructuring in the multilingual mind: motion event construal in Cantonese-English-Japanese multilingual speakers

## Consent Form

To participate in this study, please complete this consent form and return to Yi Wang in person or at the address below dtnvywa@ucl.ac.uk.

| | Yes | No |
|---|---|---|
| I have read and understood the information leaflet about the research. | ☐ | ☐ |
| I understand that my participation will involve a language background questionnaire, a description task, and a similarity judgement task, and I agree that the researcher may use the results as described here and in the Information Sheet. | ☐ | ☐ |
| I understand that if any of my words are used in reports or presentations they will not be attributed to me. | ☐ | ☐ |
| I understand that I can withdraw from the project at any time, and that if I choose to do this, any data I have contributed will not be used. | ☐ | ☐ |
| I understand that I can contact Yi Wang at any time and request for my data to be removed from the project database. | ☐ | ☐ |
| I understand that the results will be reported as part of a thesis or dissertation, may be presented at academic conferences, and may be subject to publication in journal, book, or other scholarly format. This information will not be traceable to me as an individual participant. | ☐ | ☐ |
| I agree for the data I provide to be archived at the UCL Data Safe Haven after the ensuing PhD thesis has been submitted. | ☐ | ☐ |

--------------------------------------------------------------------------------------------------------------

Name ________________________Signed ________________________

Date _____________________

**UCL Institute of Education**

20 Bedford Way, London WC1H 0AL

+44 (0)20 7612 6000 | enquiries@ioe.ac.uk | www.ucl.ac.uk/ioe