

Running-head: Uyghur-Chinese bilingual children's motion event encoding

Title: Uyghur-Chinese early successive bilingual children's acquisition of voluntary motion expressions

Author: Alimujiang Tusun

Affiliation: Pembroke College, University of Cambridge.

*Acknowledgments: The study reported here grew out of my PhD research conducted at the University of Cambridge. I gratefully acknowledge the Gates Cambridge Trust, the Cambridge Trust and Hughes Hall for supporting me with various scholarships. I also thank Ji Yinglin for her helpful feedback on an earlier draft of the paper, and three anonymous *BLC* reviewers for their invaluable comments on the manuscript. Usual disclaimers apply.

Address for correspondence: at648@cam.ac.uk

Key words: motion events, Uyghur, Chinese, early successive bilingual

1. Introduction

The crosslinguistic encoding of motion events (Talmy, 2000) has provided a fruitful venue for exploring bilingual language acquisition and use (Daller, Treffers-Daller & Furman, 2011; Hohenstein, Eisenberg & Naigles, 2006; Wang & Wei, 2021). This study extends this line of research to the context of Uyghur-Chinese¹ bilingualism. Uyghur and Chinese belong to different language families and differ markedly in their general linguistic profiles. They also represent different language types in their dominant lexicalisation patterns for encoding motion events (verb-framed vs. equipollently-framed) although they display structural overlap (verb-framing). Moreover, Uyghur-Chinese bilingualism is embedded in a socio-political milieu where the dichotomy assumed in much of previous research between societally dominant versus non-dominant language is less sharp (van Dijk, van Wonderen, Koutamanis, Kootstra, Dijkstra & Unsworth, 2022). This study thus aims to shed light on the implications of such linguistic and sociological factors for bilinguals' acquisition of motion expressions in L1 and L2. Although we are interested in bilingualism-specific issues such as crosslinguistic influence, we also seek to relate bilingual children's acquisition of motion expression to what is generally known about spatial language development in childhood. Specifically, drawing on insights from child language research that the acquisition of motion expressions is shaped by both language-specific and language-universal factors, and that certain aspects of motion expression develop throughout childhood (Hendriks, Hickmann & Pastorino-Campos, 2021), we adopt a developmental approach by including four age groups (4-, 6-, 8-, 10- year-olds). Our overall objective is to offer a more comprehensive characterisation of bilingual children's developmental trajectories in L1 and L2 while highlighting potentially universal patterns in child language development.

2. Motion expressions across languages

Talmy (2000) defined a motion event as involving a Figure moving along a Path with reference to a Ground in a particular Manner, also known as voluntary motion (Hendriks 2021). Of these components, Path is considered the framing event and Manner the co-event, and depending on whether Path is encoded in the verb or a satellite (e.g., particle, prefix), Talmy categorised the world's languages into satellite-framed (S-languages: e.g., English) and verb-framed languages (V-languages: e.g., Spanish). English is an S-language because speakers typically express Path in a particle and Manner in the main verb, as in (1); Spanish is a V-language as Path is typically expressed in the main verb and Manner (if at all) in an adjunct, as in (2). Subsequent research noted that V-languages license satellite-framed constructions if the denoted motion does not entail crossing a spatial boundary (Aske, 1989; Slobin & Hoiting, 1994), as in (3).

(1) Mary ran [Manner] into [Path] the classroom.

(2) Mary entró [Path] a la clase corriendo [Manner].

‘Mary entered the classroom while running.’

(3) Maria corrió [Manner] a [Path] la clase.

‘Maria ran to the classroom.’

The implication of Talmy's typology for language use has been explored in numerous studies, mostly in relation to Slobin's *thinking-for-speaking* hypothesis (Bunger, Skordos, Trueswell & Papafragou, 2021; Filipović, 2011; Ji & Hohenstein, 2017; 2018; Slobin, 1996; 2003; 2004; 2006; von Stutterheim, Bouhaous & Carroll, 2017; Wang & Wei, 2021). One recurrent observation in this research is that S-language speakers typically provide semantically denser motion descriptions (i.e., two components - Manner and Path)

than their V-language counterparts (i.e., one component – typically Path only). According to the *thinking-for-speaking* hypothesis, such crosslinguistic differences is rooted in language-specific ways of conceptualising experience for the purpose of verbalisation. Specifically, to express Manner and Path simultaneously, S-language speakers have compact constructions at their disposal whereas V-language speakers have to use syntactically complex constructions (e.g., subordination) that incur greater processing load (Özçalışkan & Slobin, 2003; Slobin, 2004, 2006). During online production, speakers fit their conceptualisation of the event in constructions that are most readily accessible in their language. Thus, facilitated by compact structures, S-language speakers habitually profile both Manner and Path, thereby producing semantically dense motion descriptions. V-language speakers, due to typological constraints, typically omit Manner and profile only Path - the component carried in the obligatory element of a sentence, thereby producing semantically less rich motion descriptions (Allen, Özyürek, Kita, Brown, Furman, Ishizuka & Fujii, 2007; Özçalışkan, 2015; Tusun, 2022a; Tusun & Hendriks, 2019; 2022; Wang & Wei, 2021). How early successive bilingual children develop such language-specific tendencies, and by implication, *thinking-for-speaking* patterns, will be the focus of this study.

3. Motion expression in Uyghur and in Chinese

Uyghur is a Turkic language of the southeastern branch. It is spoken in northwestern China's Xinjiang Uyghur Autonomous Region (Xinjiang) by at least 10 million native speakers (nearly half of Xinjiang's population). It is co-official with Chinese and is the lingua franca among other ethnic minorities. The language is also used in the local printing press, radio, and television broadcasting (cf. Ragagnin, 2016; Yakup, 2021). As is typical of Turkic, Uyghur is head-final with rich agglutinative morphology. Examples (4) is the

equivalent of (2) where Path is expressed in the verb, additional Path information (i.e., Goal and Source of motion) via case marking, and Manner in a converb, the functional equivalent of gerundives in European languages (Johanson, 1995); in (5), Manner is expressed in the verb while Path (i.e., Goal) is expressed in a case marker. Thus, (4) is a verb-framed construction and (5) a satellite-framed construction, and usage-based studies on Uyghur (Tusun, 2022a; Tusun & Hendriks, 2019) have shown that while Uyghur licenses satellite-framed constructions, it is a typical V-language both in terms of lexicalisation and semantic density.

- (4) Meryem sinip-qa yügür-üp [Manner] kir-d-i [Path]².
 Mary classroom-DAT run-CONV enter-PST-3SG
 Mary entered the classroom while running.

- (5) Meryem sinip-qa [path] yüger-di. [Manner]
 Mary classroom-DAT run-PST-3SG
 Mary ran to the classroom.

Talmy (2000) originally categorised Chinese as an S-language. In (6), Manner and Path are expressed in a resultative verb compound (RVC) and Talmy took the Path-encoding morpheme (V2) *jin4* to be a satellite to the main verb (V1) *pao3* expressing Manner, a pattern characteristic of Germanic languages. He maintained that, similar to Path satellites in English, the V2 morphemes form a closed class and the V2 slot is where semantic categories such as aspect and resulting state are expressed. Due to the absence of morphological marking in Chinese, however, establishing the grammatical status of the V2 morphemes is not straightforward. Moreover, unlike Germanic Path satellites, Chinese

3, they largely follow language-specific lexicalisation patterns (Allen et al., 2007; Bowerman & Choi, 2003; Guo & Chen, 2009; Hickmann, Hendriks & Harr, 2018). However, children's ability to produce semantically dense motion descriptions develop over time, and as mentioned in Section 2, typological constraints play a major role. For example, children acquiring S-languages have been found to reach the adult level of semantic density earlier than those speaking V-languages (Harr, 2012; Hickmann et al., 2018), but children learning an E-language has been found to outperform their S-language peers. Thus, Ji, Hendriks and Hickmann (2011a, 2011b) compared age-matched Chinese and English children (aged 3, 4, 5, 6, 8, 10) and showed that, across age groups, the former consistently produced more high-density descriptions than the latter; in fact, Chinese children reached the adult level of density already from age 3. Ji and colleagues attributed this to the facilitative effect of readily accessible linguistic devices in Chinese (i.e., RVC). Finally, beyond the impact of typological factors, children also display certain universal tendencies: younger children experience greater difficulty encoding motion events that involve a categorical change of location, i.e., crossing a spatial boundary, as compared to a gradual change of location (Ji et al., 2011b; Hendriks et al., 2021).

Relevant research on bilingual speakers concerned the extent to which they think for speaking in language-specific ways, and whether, to what extent and why there is crosslinguistic influence (CLI), defined as the overuse of morphosyntactic structures in bilinguals' one language under the influence of the other language (Serratrice, 2013). Regardless of whether it is simultaneous child bilinguals (Engemann, 2021; Miller, Furman & Nicoladis, 2018) or successive child (Aktan-Erciyas, 2020; Aktan-Erciyas et al., 2020; Avelledo & Athanasopoulos, 2015; Engemann, 2016) or adult bilinguals (Daller et al., 2011; Hohenstein et al., 2006), the general understanding is that bilinguals largely

follow language-specific lexicalisation patterns, but they also exhibit CLI. To illustrate, studies almost always involved one V-language and one S-language and compared to monolinguals, bilinguals displayed in-between encoding tendencies where they used more Manner verbs in their V-language and more Path verbs in their S-language. The dimension of bilingual speakers' semantic density, especially with respect to how it is affected by CLI, is not well-understood, but preliminary evidence suggests that child bilinguals' semantic density in the V-language can be augmented under the influence of the S-language, although this implicates the use of target-deviant lexicalisation patterns (Engemann, 2016).

Factors proposed as underlying CLI include, inter alia, structural overlap, the amount of relative exposure and language dominance. Structural overlap is important because, despite their preferred motion encoding strategies, languages tend to share lexical (and syntactic) resources (e.g., Manner verbs, Path verbs, cf. Beavers et al., 2010). Whilst bilinguals do capitalize on crosslinguistically shared options (Filipović, 2022), this seems to be modulated by language dominance and the relative amount of exposure. For example, Hohenstein et al. (2006) found Spanish-English bilingual adults living in the U.S. to display L2 to L1 influence, which they attributed to the sociolinguistic setting wherein English was the dominant language. Similarly, Daller et al. (2011) showed that Turkish-German adult bilinguals resident in Germany tend to use lexicalisation patterns characteristic of German (S-language) when verbalising motion events in L1 Turkish whereas those living in Turkey tend to use patterns typical of Turkish in their L2 German, thereby reflecting the typical pattern of the (societally) dominant language. Two other studies (Aktan-Erciyes, 2020; Aktan-Erciyes et al., 2020) on Turkish-English child bilinguals (aged 5 vs. 7, AoO=3) in Turkey also argued for the impact of language

dominance on CLI: 5-year-old bilinguals exhibited an L2 to L1 influence (more Manner verbs, fewer Path verbs) while the 7-year-olds displayed L1 to L2 influence. Aktan-Erciyes and colleagues explained that this was because the 5-year-olds had total immersion in L2 English (8 hours daily) whereas the 7-year-olds' quantity of L2 exposure dropped (2 hours daily) when they attended Turkish-dominant schools. Somewhat related are findings from Aveledo (2015) and Aveledo and Athanasopolous (2016) on Spanish-English child bilinguals (aged 5-7 vs. 8-9, AoO=3-4) in Venezuela that only the older bilinguals showed L2 to L1 influence due to their increased L2 exposure (16 hours weekly) compared to the younger bilinguals (8 hours weekly). Thus, shifts in language dominance, typically associated with the amount of relative exposure to a given language, shape CLI in bilinguals' motion expression.

5. Uyghur-Chinese early successive bilingual children's acquisition of motion expressions

The above-mentioned studies have undoubtedly improved our understanding of bilingual expression of motion, but most of them included a V-language and an S-language that were genetically related (Aveledo & Athanasopoulos, 2015; Engemann, 2016; Hohenstein et al., 2006) while a better appreciation of the role of language-specific factors in bilingual language acquisition calls for more diverse language combinations (Serratrice, 2013; Yip & Matthews, 2022). Additionally, most of the studies involved adult bilinguals while those on children tended not to include many age groups. We know little about how language-universal factors found to operate in monolingual child language acquisition (Hendriks et al., 2021) inform bilingual language acquisition, how the aspect of 'semantic density', known to develop later than the acquisition of lexicalisation patterns *per se* in monolinguals, develops in bilingual children, how this is affected by CLI, and indeed, how

CLI plays out developmentally, an issue much debated in the context of the acquisition of morphosyntax (Chondrogianni, 2023; van Dijk et al., 2022) but relatively unexplored in the motion domain. We therefore need a developmental perspective. Finally, reflective of the field of bilingualism research, most previous studies concerned Western immigration contexts where one of the languages is societally dominant and the other the minority/heritage language, but it is doubtful that this dichotomy is readily applicable to bilingualism situations in non-Western communities. And we need more information on such communities and how affordances specific to their own sociolinguistic realities shape the acquisition and use motion expressions (Foroodi-Nejad & Paradis, 2009; Paradis & Nicoladis, 2007).

This study contributes to closing these gaps. First, it focuses on a hitherto unexplored language combination a verb-framed language and an equipollently-framed language that are distant genetically (Turkic vs. Sino-Tibetan), and distinct in their general linguistic profile (agglutinating vs. isolating). Second, it adopts a developmental perspective with a view to shedding light both on bilingualism-related issues such as the role of CLI during acquisition (Hulk, 2017; van Dijk et al., 2022), and on what is potentially common/universal in children's spatial language development, be it monolingual or bilingual. Third, Uyghur-Chinese bilingualism presents a non-Western bilingual situation where the distinction between societally dominant versus non-dominant language is blurred, not least because, as mentioned in Section 3, Uyghur is co-official with Chinese in Xinjiang and is a regional lingua franca, and Uyghurs constitute nearly half of Xinjiang's population, and attach great importance to promoting and maintaining their language (see Elterish, 2015 and Zang, 2015 for insights into the sociology and politics of Uyghur-Chinese bilingualism). Within this sociolinguistic milieu, Uyghur children typically grow

up speaking their L1 Uyghur, and the regional educational policy, at least in urban areas, is such that at around age 3, they attend full immersion Chinese kindergartens and subsequently full immersion Chinese schools at around age 6 (cf. Ma, 2012; Zheng, 2011). They are therefore early successive bilinguals (Chondrogianni & Vasić, 2016; Meisel, 2018) who acquire their L2 naturalistically. A relevant affordance of this unique bilingual setting, which contrasts with much of previous research, concerns bilingual children's relative exposure to and use of their two languages. Sociolinguistic research on Uyghur-Chinese bilinguals' language use (Elterish, 2015, 2016) reports that they tend to exclusively use Uyghur outside the school context, and by virtue of their schooling from kindergarten onwards, bilinguals' exposure to L2 Chinese remains constant (about 8 hours daily). That is, bilinguals' waking hours are somewhat naturally divided into 8 hours of Chinese immersion and 8 hours of Uyghur outside school. And in light of recent insights that bilinguals' relative amount of language exposure and use can serve as a proxy for language proficiency and language dominance (Unsworth et al., 2016, 2018), the sociolinguistic setting in question is arguably more conducive to more balanced bilingualism (Filipović, 2019), and as such, Uyghur-Chinese bilingual children's acquisition and use of motion expressions may be different from their peers in other bilingual contexts. Against this backdrop, this study asks the following research questions:

RQ1: Whether and at what age do Uyghur-Chinese early successive bilingual children's motion expressions in L1 Uyghur become adult-like both in terms of lexicalisation pattern and semantic density?

RQ2: Whether and at what age do children's motion expressions become adult-like in L2 Chinese and what is the role of crosslinguistic influence in the acquisition process?

In relation to RQ1, we predicted that, like children learning V-languages (Hendriks et al., 2021 for French; Aktan-Erciyes et al., 2021 for Turkish), bilinguals from the earliest age tested would follow the adult pattern of expressing Path in the verb and Manner in the converb. Additionally, they would express additional Path information (e.g., Source, Goal) via case markers (cf. Furman, 2011 for Turkish). However, the adult-like ability to simultaneously express Manner and Path would develop much later (cf. Harr, 2012). In terms of RQ2, several predictions could be entertained. Given the early and systematic exposure to the L2, it was predicted that bilinguals from the earliest age tested would be fully adult-like with no L1 influence: they would predominantly use equipollently-framed constructions (i.e., RVCs) and much less frequently, verb-framed constructions (i.e., Path in verb and Manner in subordinate clause), as did Chinese monolingual children (cf. Ji et al., 2011a). They would therefore predominantly produce semantically dense descriptions. Alternatively, given the structural overlap of verb-framing between Uyghur and Chinese, bilinguals could use such constructions more frequently than Chinese adults, thereby displaying CLI. This means that they would express only Path in the verb and additional Path information via satellites (e.g., prepositions) but will omit Manner; they would therefore produce low-density descriptions (cf. Hendriks et al., 2021; Slobin, 2004). In terms of how this CLI would manifest developmentally, two possibilities were considered. As per the claim that CLI is part and parcel of the bilingual experience (cf. Chondrogianni, 2023; van Dijk et al., 2022), CLI would persist throughout childhood such that Uyghur-Chinese bilingual children across age groups would consistently use verb-framed constructions more than Chinese adults. Alternatively, in light of L2 studies showing a decrease of CLI as a function of increased proficiency (Montero-Melis & Jaeger, 2020; Park, 2020), bilingual children would use verb-framed constructions more frequently than adults at the early stages, but this would decrease over time while their use of

equipollently-framed constructions would increase, eventually converging on the target equipollent system. This would be compatible with the hypothesis that CLI is a developmental phenomenon (Hulk, 2017). In this case and considering the strong influence of language-specific factors on motion expression (Hendriks et al., 2021; Ji et al., 2011a), bilinguals' semantic density would be higher in Chinese than Uyghur.

6. The study

6.1 Participants

The participants consisted of three groups: Uyghur-Chinese bilingual children, Uyghur adults and Chinese adults. The Uyghur adult group contained 24 speakers, of which 20 were postgraduate students who had recently come to the UK for postgraduate studies, and 4 were based in Xinjiang. In addition to Uyghur, those tested in the UK spoke Chinese and English while those in Xinjiang spoke Chinese. The adult speakers were therefore not monolinguals, but monolinguals are hard to come by in Xinjiang due to its widespread bilingual education (Ma, 2012) that also entails learning English as a foreign language (Feng & Adamson, 2017; Sunuodula & Cao, 2015). Our adult participants' multilingual profiles are thus reflective of that of the younger generation of Uyghurs in Xinjiang. The Chinese adult group included 12 speakers who were university students in Beijing³.

The bilinguals consisted of four age groups with each containing 24 participants: 4-year-olds (B04; age range 3;11-4;7; mean age 4;6), 6-year-olds (B06; age range 5;9-6;6; mean age 6;5), 8-year-olds (B08; age range 7;9-8;4; mean age 8;4) and 10-year-olds (B10; age range 9;8-10;7; mean age 10;6). They were recruited from Chinese immersion kindergartens and primary schools in Ürümchi, Xinjiang and were early successive bilinguals as the exposure to their L2 started at a mean age of 3;2 (range = 3;1-3;4)

(Chondrogianni & Vasić, 2016; Meisel, 2018). The recruitment process started with an initial teachers' screening that involved identifying those who grew up in Uyghur families, and for the 4-year-olds, those who were perceived as highly proficient in Chinese. Upon identifying the appropriate students, their parents were invited to complete a questionnaire on family language practice, literacy activities and parent's ratings of children's proficiency in Uyghur and Chinese (on a scale from 1–10). Based on their responses, we selected only those who were exclusively exposed to and used Uyghur outside school and at home, thereby balancing out their 8 hours of daily Chinese immersion at school, and those whose proficiency ratings in both languages were 8 or above. They were therefore relatively balanced bilinguals (cf. Unsworth et al., 2018).

6.2 *Materials and procedure*

Data were elicited using a set of 18 short video clips in which a protagonist moved along vertical (UP/DOWN) or boundary-crossing paths (ACROSS) in a particular manner. Each path type was represented 6 times in the whole set, resulting in a total of 18 experimental items (see Appendix S1 for the full list). They were randomised into six test orders and were assigned to the participants randomly, and each bilingual performed the same task twice: once in Uyghur and once in Chinese. To minimise task repetition effects, half of the bilinguals performed the task first in Uyghur and the other half first in Chinese. The interval between the two experimental sessions for each bilingual participant was about 1-2 weeks.

The participants were met individually in a quiet room and the cartoons were presented on a computer screen. To ensure that they maximally relied on linguistic means than on gestures, adults and older children had to narrate to an imaginary addressee who had to

reconstruct the whole event based on their speech alone. The youngest children described the clips to an adult who sat opposite them and therefore had no visual access to the cartoons. Each session started with a training item and whenever necessary, the participants were probed so that they minimally noticed the manipulated components (Manner and Path). To sustain the flow of children’ speech, some general questions were asked (e.g., “What happened?”, “And then?”). Great care was taken to induce maximally monolingual mode throughout.

6.3 *Coding and analysis*

All the responses were transcribed into CHAT format (CHILDES; MacWhinney, 2000) and were first segmented into clauses, with a clause defined as a unit containing one verb and its arguments (Hickmann et al., 2018). Thus, responses exemplified in (8) and (9) were segmented into two clauses, a subordinate clause and a matrix clause. When occasionally participants gave more than one response for an item, two criteria were applied hierarchically, i.e., richness and relevance of Path. For example, based on ‘richness’, R2 in (10) would be selected as the ‘target response’ because it simultaneously expressed Manner and Path. However, when responses contained either Manner or Path (1.4%), as in (11), we chose, as per ‘relevance’, R2 as the ‘target response’ (cf. Talmy, 2000). In all cases, R1 was marked as ‘potential target response’, but was not included in our analysis.

| | | | | |
|-----|------|----------|---------------|----------------|
| (8) | Éyiq | dereχ-ke | yamiš-ip [c1] | čiq-t-i [c2] |
| | bear | tree-DAT | climb-CONV | ascend-PST-3SG |

The bear ascended the tree by climbing.

- (9) Ta1 hua2 zhe1 bing1 [c1] guo4 he2 [c2]
 he slide ASP_{dur} ice cross river
 He crossed the river by sliding.

- (10) Adem yol-din öt- t-i. [R1]
 man road-ABL cross-PST-3SG
 That man crossed the road.
 Yügür-üp öt-t-i [R2].
 run-CONV cross-PST-3SG
 He crossed the road running.

- (11) Ta1 pao3bu4 [R1] ran2hou4 guo4qu1 le [R2]
 He run then cross-go ASP_{perf}
 He ran and then crossed the road.

Each target response was coded in terms of the semantic information expressed in various linguistic devices (information locus) and the total number of motion components expressed (semantic density). In relation to information locus, following previous studies (e.g., Hendriks et al., 2021; Wang & Wei, 2021), two loci were identified: the main verb (the verb locus) versus the satellite, defined as all other devices outside the main verb (the OTH locus). Both V1 and V2 elements of an RVC in Chinese were coded as two verbs, and the OTH locus included dative/ablative case markers, converbs, adverbials (e.g., *feisu4de* ‘quickly’) and prepositional phrases (e.g., *cong2 you4bian1* ‘from the right side’). In terms of motion information in the verb locus, responses in Uyghur fell into two categories, those encoding Path and those encoding Manner (see 12-13), while those in Chinese fell into three categories, i.e., Path, Manner, or Path+Manner (see 14-16). With respect to the

OTH locus, responses across the two languages were categorised into those expressing Path (see 17 and 21), Manner (see 18 and 22), Path+Manner (see 19 and 23), and Zero – a residual category for responses with no satellite devices (‘bare verb constructions’ à la Hohenstein et al., 2006) and thus no spatial information in this locus (see 20 and 24). For semantic density, only semantic information from distinct categories was considered (irrespective of the linguistic devices used) such that multiple mentions of Path within one response counted as density 1 (SD1, see 12 and 21) while one mention of Path and one of Manner counted as density 2 (SD2, see 13 and 23).

- (12) Čaşqan üstel-ge çiq-t-i [Path].
 mouse table-DAT ascend-PST-3SG

The mouse ascended to the table.

- (13) Qurut dereχ-ge yamaš-t-i [Manner].
 caterpillar tree-DAT climb-PST-3SG

The caterpillar climbed the tree.

- (14) Zhe4ge4ren2 cong2 ma3lu4 shang4 guo4qu4 [Path] le1.
 This man from road on cross-go ASP_{perf}

This man crossed from the road.

- (15) Song1shu3 pa2 [Manner] shu4
 squirrel climb tree

The squirrel climbed the tree.

- (16) Hou2zi1 cong2 shu4 shang4 pa2xia4lai2 [Path+Manner].
 monkey from tree on climb-descend-come
 The monkey climbed down from the tree.
- (17) Müšük türük-niñ üsti-din [Path] čüş-t-i.
 cat pole-GEN top-ABL descend-PST-3SG
 The cat descended from the top of the tree.
- (18) Bala tilin-ip [Manner] öt-t-i.
 boy slide-CONV cross-PST-3SG
 The boy crossed by sliding.
- (19) Bir bala yol-din [Path] yügürüp [Manner] a_yaq-qa [Path] ket-t-i.
 one boy road-ABL run-CONV that side-DAT go-away-PST-3SG
 A boy went to that side of the road by running.
- (20) Éyiq čüş-üp ket-t-i. (Zero, BU1016)
 bear descend-CONV ASPV-PST-3SG
 The bear descended.
- (21) Chong2zi1 cong2 cao3 shang4 [Path] pa2xia4lai2.
 caterpillar from grass on crawl-descend-come
 The caterpillar crawled down from the grass.
- (22) Bao3bao1 pa2 zhe1 [Manner] guo4 ma3lu4.

baby crawl ASP_{dur} cross road

The baby crossed the road while crawling.

(23) Yun4dong4yuan2 fei1su4de [Manner]

sportsman quickly

cong2 ma3lu4 [Path] zuo3bian1 pao3dao4 you4bian1.

from road left side run-arrive right side

A sportsman ran quickly from the left side of the road to the right side.

(24) Lao3shu3 xia4qu4 le1.

mouse descend ASP_{perf}

The mouse descended.

For the statistical analysis, our independent variable was age whereas dependent variables were the mean occurrence of Path verbs, Manner verbs, Path satellites, Manner satellites as well as SD1 and SD2 responses. The count data were analysed by fitting generalised linear mixed-effects models with a Poisson distribution, using R (R Core Team, 2017), the `glmer()` function in the `lme4` library. We first fitted a model with the fixed effects in question against a reduced model without the fixed effects in question to the same dataset. We then compared the relative goodness of fit of the two models using a likelihood ratio test via the `anova()` command, which revealed the relative fits (expressed as log likelihood) of the two models to test the statistical significance of the fixed effect removed in the reduced model. For all models fitted, random intercept for participant and item were included. Planned contrasts with Bonferroni adjustment were specified where more than

two factors were compared (cf. Appendix S2). We report the chi-square statistics, degrees of freedom and p value for the tests. All model outputs are provided in the ‘Appendix S3’.

7. Results

7.1 Information in the verb locus in Uyghur and in Chinese

Figure 1a shows information expressed in the verb locus in Uyghur by age⁴. A two-way packaging (Path, Manner) x age (4yrs, 6yrs, 8yrs, 10yrs, adults) analysis revealed a significant interaction ($\chi^2(4)=68.998, p < .001$), suggesting that the two lexicalisation patterns varied by age. Further analyses found an age effect only for Manner ($\chi^2(4)=29.14, p < .001$) as 4- and 6-year-olds encoded this component more frequently than adults ($\beta_{4\text{yrs-AD}} = -0.81, SE = 0.32, \text{Wald } z = -2.53, p = .011; \beta_{6\text{yrs-AD}} = -1.26, SE = 0.27, \text{Wald } z = -4.69, p = <.001$). That is, children fully established their L1 verb-framed pattern from age 4 while their early tendency to encode Manner dropped to the adult level at age 8.

Figure 1b represents information expressed in the verb locus in Chinese by age. A two-way packaging (Path, Manner, Path+Manner) x age (4yrs, 6yrs, 8yrs, 10yrs, adults) interaction analysis was significant ($\chi^2(8)=306.11, p < .001$), indicating that children’s lexicalisation patterns varied by age. Further analyses found age effects for Path ($\chi^2(4)=35.932, p < .001$), Manner ($\chi^2(4)=54.398, p < .001$) and Path + Manner ($\chi^2(4)=68.828, p < .001$). Further analyses revealed that 4-, 6- and 8-year-olds encoded Path more frequently than adults ($\beta_{4\text{yrs-AD}} = -0.78, SE = 0.17, \text{Wald } z = -4.40, p < .001; \beta_{6\text{yrs-AD}} = -0.59, SE = 0.17, \text{Wald } z = -3.35, p < .001; \beta_{8\text{yrs-AD}} = -0.63, SE = 0.21, \text{Wald } z = -2.96, p = .003$), and 4- and 6-year-olds encoded Manner more frequently than adults ($\beta_{4\text{yrs-AD}} = -1.58, SE = 0.33, \text{Wald } z = -4.79, p < .001, \beta_{6\text{yrs-AD}} = -1.45, SE = 0.30, \text{Wald } z = -4.76, p < .001$). Finally, only 4- and 6-year-olds used the Path+Manner pattern less

frequently than adults ($\beta_{4\text{yrs-AD}} = -0.98$, $SE = 0.15$, $Wald\ z = 6.24$, $p < .001$; $\beta_{6\text{yrs-AD}} = 0.58$, $SE = 0.11$, $Wald\ z = 4.87$, $p < .001$), the steady increase of this pattern within the four child groups was significant at each age level ($\beta_{4\text{yrs-6yrs}} = -0.40$, $SE = 0.16$, $Wald\ z = 2.47$, $p = .013$; $\beta_{6\text{yrs-8yrs}} = 0.34$, $SE = 0.13$, $Wald\ z = 2.53$, $p = .011$; $\beta_{8\text{yrs-10yrs}} = -0.35$, $SE = 0.10$, $Wald\ z = -3.34$, $p < .001$). That is, children's verb-framed pattern dropped to the adult level at age 10 while their equipollently-framed pattern (i.e., RVC) increased to the adult level from age 8. The Chinese equipollent framing system (both verb- and equipollently-framed lexicalisation patterns) was fully established at age 10.

<Insert Figure 1 about here>

7.2 Information in the OTH locus in Uyghur and in Chinese

Figure 2a illustrates information expressed in OTH locus in Uyghur by age. A two-way packaging (Path, Manner, Path+Manner, Zero) x age (4yrs, 6yrs, 8yrs, 10yrs, adults) analysis revealed a significant interaction ($\chi^2(12)=198.91$, $p < .001$), reflecting that children's packaging strategies varied by age. Further analyses found age effects only for Path ($\chi^2(4)=77.53$, $p < .001$) and Path + Manner ($\chi^2(4)=46.80$, $p = .001$) such that all the child groups expressed Path more frequently than adults ($\beta_{4\text{yrs-AD}} = -0.92$, $SE = 0.12$, $Wald\ z = -7.46$, $p < .001$; $\beta_{6\text{yrs-AD}} = -0.91$, $SE = 0.33$, $Wald\ z = -3.00$, $p < .001$; $\beta_{8\text{yrs-AD}} = -1.01$, $SE = 0.12$, $Wald\ z = -8.30$, $p < .001$; $\beta_{10\text{yrs-AD}} = -0.74$, $SE = 0.14$, $Wald\ z = -5.23$, $p < .001$), and Path + Manner less frequently than adults ($\beta_{4\text{yrs-AD}} = 0.98$, $SE = 0.13$, $Wald\ z = 7.40$, $p < .001$; $\beta_{6\text{yrs-AD}} = 0.75$, $SE = 0.12$, $Wald\ z = 6.15$, $p < .001$; $\beta_{8\text{yrs-AD}} = 0.60$, $SE = 0.11$, $Wald\ z = 5.17$, $p < .001$; $\beta_{10\text{yrs-AD}} = 0.34$, $SE = 0.11$, $Wald\ z = 3.08$, $p = .002$). That is, children did not fully establish the lexicalisation pattern for the OTH locus in Uyghur even at age 10.

Figure 2b displays information expressed in OTH locus in Chinese by age. A two-way components (Path, Manner, Path+Manner, Zero) x age (4yrs, 6yrs, 8yrs, 10yrs, adults) interaction analysis was significant, ($\chi^2(12)=99.566, p < .001$), showing that the relative frequency of the four patterns varied by age. Further analyses specified the age effects to Manner ($\chi^2(4)=23.06, p < .001$), Path+Manner ($\chi^2(4)=19.79, p < .001$) and Zero ($\chi^2(4)=23.29, p < .001$) such that 4- and 6-year-olds expressed Manner ($\beta_{4\text{yrs-AD}} = 1.21, SE = 0.33, Wald z = 3.60, p < .001$; $\beta_{6\text{yrs-AD}} = 1.69, SE = 0.39, Wald z = 4.32, p < .001$) and Path+Manner less frequently ($\beta_{4\text{yrs-AD}} = 2.91, SE = 1.11, Wald z = 2.61, p = .008$; $\beta_{6\text{yrs-AD}} = 1.80, SE = 0.73, Wald z = 2.44, p = .014$) while all child groups produced Zero more frequently than adults ($\beta_{4\text{yrs-AD}} = -0.41, SE = 0.10, Wald z = -4.06, p < .001$; $\beta_{6\text{yrs-AD}} = -0.37, SE = 0.10, Wald z = -3.57, p < .001$; $\beta_{8\text{yrs-AD}} = -0.21, SE = 0.10, Wald z = -2.06, p = .039$; $\beta_{10\text{yrs-AD}} = -0.21, SE = 0.10, Wald z = -2.00, p = .045$). That is, children converged on the adult pattern in the OTH locus by age 8, although they continued to produce more motion constructions without satellite devices, i.e., Zero, than adults.

<Insert Figure 2 about here>

7.3 Semantic density in Uyghur and in Chinese

Figure 3a depicts semantic density in Uyghur by age group. A two-way density (SD1, SD2) x age (4yrs, 6yrs, 8yrs, 10yrs, adults) analysis showed a significant interaction ($\chi^2(4)=132.93, p < .001$), indicating that semantic density varied by age. Further analyses identified age effects for both SD1 ($\chi^2(4)=54.53, p < .001$) and SD2 ($\chi^2(4)=31.38, p < .001$) such that all child groups produced SD1 descriptions more frequently than adults ($\beta_{4\text{yrs-AD}} = -0.82, SE = 0.10, Wald z = -7.63, p < .001$; $\beta_{6\text{yrs-AD}} = -0.63, SE = 0.13, Wald z = -4.84, p < .001$; $\beta_{8\text{yrs-AD}} = -0.80, SE = 0.12, Wald z = -6.25, p < .001$; $\beta_{10\text{yrs-AD}} = -0.63, SE = 0.12,$

Wald $z = -5.01, p < .001$) but SD2 descriptions less frequently than adults ($\beta_{4\text{yrs-AD}} = -0.69$, SE = 0.10, Wald $z = 6.62, p < .001$; $\beta_{6\text{yrs-AD}} = -0.36$, SE = 0.10, Wald $z = -3.48, p < .001$; $\beta_{8\text{yrs-AD}} = 0.58$, SE = 0.12, Wald $z = 4.77, p < .001$; $\beta_{10\text{yrs-AD}} = -0.34$, SE = 0.09, Wald $z = 3.67, p < .001$). That is, children stopped short of the adult frequency for SD2 descriptions even at age 10.

Figure 3b shows semantic density across age groups in Chinese. A two-way density (SD1, SD2) x age (4yrs, 6yrs, 8yrs, 10yrs, adults) analysis revealed a significant interaction ($\chi^2(4)=362.5, p < .001$), suggesting that semantic density varied by age. Further analyses found age effects for both SD1 ($\chi^2(4)=106.66, p < .001$) and SD2 ($\chi^2(4)=85.22, p < .001$). Specifically, while the stepwise decrease of SD1 descriptions was significant at each age group ($\beta_{4\text{yrs-6yrs}} = -0.17$, SE = 0.08, Wald $z = -2.12, p = .033$; $\beta_{6\text{yrs-8yrs}} = -0.38$, SE = 0.11, Wald $z = -3.20, p = .001$; $\beta_{8\text{yrs-10yrs}} = -0.60$, SE = 0.19, Wald $z = -3.19, p = .001$; $\beta_{10\text{yrs-AD}} = -0.63$, SE = 0.27, Wald $z = -2.33, p = .019$) and all child groups produced SD1 descriptions more frequently than adults ($\beta_{4\text{yrs-AD}} = -1.79$, SE = 0.20, Wald $z = -8.81, p < .001$; $\beta_{6\text{yrs-AD}} = -1.62$, SE = 0.20, Wald $z = -7.90, p < .001$; $\beta_{8\text{yrs-AD}} = -1.23$, SE = 0.24, Wald $z = -5.01, p = .001$; $\beta_{10\text{yrs-AD}} = -0.63$, SE = 0.27, Wald $z = -2.33, p = .019$). Meanwhile, the increase of SD2 descriptions was significant from age 4 to 10 ($\beta_{4\text{yrs-6yrs}} = -0.39$, SE = 0.15, Wald $z = -2.51, p = .011$; $\beta_{6\text{yrs-8yrs}} = -0.38$, SE = 0.12, Wald $z = -3.11, p = .001$; $\beta_{8\text{yrs-10yrs}} = -0.28$, SE = 0.09, Wald $z = -3.21, p = .001$). At age 10, children attained the adult level of frequency for SD2 descriptions.

<Insert Figure 3 about here>

8. Discussion

This study investigated early successive bilingual children's acquisition of a novel language pair (Uyghur vs. Chinese) that is genetically distant (Turkic vs. Sino-Tibetan) and typologically distinct (verb-framed vs. equipollently-framed) with some degree of structural overlap (verb-framing). In light of L1 research that aspects of motion expression develop throughout childhood (cf. Hendriks et al., 2021; Hickmann et al., 2018), we adopted a cross-sectional design to ascertain whether and when Uyghur-Chinese bilingual children become adult-like in their L1 and L2. By taking a developmental perspective, we aimed to illuminate, on the one hand, the nature of CLI in child L2 acquisition (e.g., the role of structural overlap, longevity of CLI), and on the other hand, potentially language-universal processes underlying monolingual and bilingual children's spatial language development. In terms of their L1 Uyghur, we predicted children to be sensitive to the lexicalisation patterns from early on in that they would encode Path in the main verb, additional Path information in case markers, and Manner in a converb. However, since jointly expressing Manner and Path requires complex structures (e.g., subordination) that develop later, children would initially focus on Path until relevant structures were fully developed, leading also to a concomitant increase in semantic density. Given the early and systematic exposure to and use of their L2, and their relatively balanced language profile, we predicted the same early sensitivity such that children would predominantly encode Path and Manner in an RVC (equipollent-framing) displaying no CLI. However, given the structural overlap of verb-framing between L1 and L2, we also predicted that children could employ verb-framed constructions (i.e., expressing Path in verb and additional Path in satellite devices) without Manner, as the relevant devices develop over time. In terms of how this L1 to L2 influence would play out developmentally, we entertained two

possibilities: CLI would either remain stable across childhood, or it would eventually disappear when Chinese equipollent system is fully established.

Our predictions about the acquisition of Uyghur have been confirmed. In terms of their sensitivity to the target lexicalisation pattern, children from age 4 encoded Path in the verb as frequently as adults, and despite an early tendency to encode Manner in the verb locus, a point we will discuss shortly, children encoded Manner primarily in the OTH locus (via converbs). This is consistent with existing studies on children learning V-languages such as Turkish, Japanese, and French (e.g., Allen et al., 2007; Hickmann et al., 2009; Hendriks et al., 2021). Furthermore, children from age 4 provided additional Path information in the OTH locus (via dative/ablative case markers), which reflects their early sensitivity to yet another dimension of encoding motion in morphologically rich verb-final V-languages (cf. Ibarretxe-Antuñano, 2009) and echoes earlier findings on Turkish monolingual (Özçalışkan, 2009) and bilingual children (Woerfel, 2018).

On the measure of semantic density, while there was a slight increase of SD2 descriptions (i.e., combining Path and Manner) over time, children did not reach the adult level even at age 10. Again, this has been previously observed for Turkish (e.g., Özçalışkan & Slobin, 2000) and French (e.g., Harr, 2012; Harr & Hickmann, 2013; Hendriks et al., 2021) monolingual children, as well as Turkish-German and Turkish-French bilingual children (Woerfel, 2018). This could be explained on typological grounds. As mentioned in Section 2, even adult V-language speakers tend to produce semantically less dense motion descriptions because expressing Path and Manner simultaneously would typically require complex structures (e.g., subordination) that incur higher processing load (e.g., Özçalışkan, 2015; Tusun & Hendriks, 2019). Given that children are subject to the same typological constraints

and assuming that their processing capacities (e.g., working memory) essential for complex syntax are still developing (cf. Delage & Frauenfelder, 2019; Gathercole et al., 2004), it is understandable that they didn't display adult-like productivity of SD2 descriptions at age 10. However, Tusun's (2022b) recent finding that even Uyghur-Chinese early successive bilingual adults stopped short of monolingual pattern for semantic density suggests that this may not be a developmental phenomenon but rather a more general tendency in bilingual language use. Put differently, while avoiding the use of syntactically complex motion constructions (unless necessitated by the communicative task) may be a tendency of V-language speakers, it may be more pronounced or persistent in bilingual speakers due to the challenges inherent in dual language processing and use (cf. Filipović, 2022; Filipović & Hawkins, 2019; also see Engemann, 2016; 2022 for simultaneous and successive bilingual children's similar tendencies in expressing caused motion)⁵.

Turning to the acquisition of Chinese, children indeed showed early sensitivity to the target equipollent framing system. From age 4, they expressed Path and Manner in the verb locus via an RVC (i.e., equipollently-framed pattern), and they also expressed Path only in the verb locus (i.e., verb-framed pattern). In the OTH locus, they mirrored adults in that they hardly provided any spatial information and that, when they occasionally did, they expressed additional Path information and sometimes even Manner. However, contrary to our prediction, early sensitivity did not imply early systematicity. While there was a clear increase in children's use of the equipollently-framed pattern over time, they matched adult frequency only at age 8. The predicted L1 to L2 influence did occur as children used the verb-framed pattern significantly more frequently than adults up to age 8. And as predicted, when the verb-framed pattern was used, Manner information was not provided in the OTH locus (via subordinate structures). However, the predicted additional Path encoding (via

prepositions) due to L1 influence did not occur. The L1 to L2 influence was restricted to the main verb locus, sparing the verbal periphery, and consequently, as predicted, children produced low-density descriptions at the early stages. Regarding how CLI would play out developmentally, the results supported the second possibility, i.e., that CLI would eventually disappear and children's use of the verb-framed pattern and equipollently-framed pattern fully converged on the adult level at age 10. Our final prediction, based on the importance of language-specific factors, was also met because children's semantic density in Chinese reached the adult pattern at age 10 (but not in Uyghur). Overall, the combined findings on measures of information locus and semantic density suggest that children established the L2 equipollent system at age 10.

This developmental pattern is distinct from that of Chinese monolingual children. Recall that, using the same elicitation material and analytical framework, Ji et al. (2011a) found that Chinese children were adult-like from age 3 with no development up to age 10. Specifically, their use of the equipollently-framed and verb-framed patterns and the semantic density of motion descriptions matched adult frequencies from age 3 already. In contrast, verb-framing constituted a main strategy for bilingual children up to age 8, which had direct (negative) impact on semantic density as well. Their distinct developmental path therefore seems to stem partly from the influence of verb-framing in L1 to their L2 (Aktan-Erciyes, 2020; Aktan-Erciyes et al., 2020; Aveledo & Athanasopoulos, 2015; Hohenstein et al., 2006; Park, 2020).

Another pattern that contributed to bilinguals' distinct developmental path was their tendency to lexicalise Manner in the verb at ages 4 and 6. A qualitative look at the data revealed that this pattern occurred primarily with ACROSS events (4yrs: 54%, 6yrs: 65%) and to a lesser

extent, UP events (4yrs: 40%; 6yrs: 28%). Further inspection of descriptions of ACROSS events showed that they were typically represented as locative rather than boundary crossing (e.g., ‘He swims in the river’ vs. ‘He swam across the river’). Given that they showed the same tendency in L1 Uyghur (ACROSS: 4yrs: 58%, 6yrs: 40%; UP: 4yrs: 37%, 6yrs: 57%) and that similar tendencies have been documented for young children acquiring Chinese (Ji, 2009), English, French and German (Harr, 2012; Hickmann et al., 2009), it is likely that our bilinguals were constrained by a more universal challenge younger children experience in encoding events that involves categorical change of location (cf. Hickmann et al., 2018; Hendriks et al., 2021). For UP events, children unanimously used the Manner verb *pa2* ‘to climb’, and interestingly, they used its equivalent *yamašmaq* in their Uyghur descriptions. Lewandowski & Mateu (2020) hold that such Manner verbs display a certain degree of Path salience due to encyclopaedic and contextual knowledge, and indeed, previous L1 research has shown that young children capitalise on such structures that enable them to express more information (Manner & Path) with less complex structures (e.g., Hickmann et al., 2009; Hendrik et al., 2021; Özçalışkan & Slobin, 1999; Özyürek, A., & Özçalışkan). The younger bilinguals’ frequent use of Manner verbs may thus reflect this same universal tendency.

Before turning to the more general implications of these developmental patterns, we should highlight one qualitative divergence bilinguals displayed in comparison to monolinguals. When describing ACROSS events, instead of using of Path verbs (e.g., ‘cross’) that unequivocally encode boundary crossing, children would sometimes use deictic verbs such as *ketmek* ‘go away’ (cf. Example 19) in Uyghur and semantically general Path verbs like *dao4* ‘arrive/reach’ in Chinese (cf. Example 23) (Uyghur: 4yrs: 7%; 6yrs: 8%, 8yrs: 6%; 10yrs: 8%; Chinese: 4yrs: 7%, 6yrs: 10%, 8yrs: 21%, 10yrs: 15%). They would then mention the Source and Goal of motion via satellite devices (e.g., case markers in Uyghur and

prepositions in Chinese) so that the notion of boundary crossing could be inferred. This pattern is not found in the respective adult data, nor has it been previously reported for monolingual children acquiring V-languages (cf. Hickmann et al., 2009) or Chinese (cf. Ji et al., 2011a). Given that it occurred in all age groups, and indeed in Uyghur-Chinese adult bilinguals (cf. Tusun, 2022b), the possibility that it is a developmental phenomenon can be ruled out. Rather, it echoes previous findings that bilingual speakers tend to use semantically general verbs that could be applied in various contexts (e.g., Álvarez, 2008; Engemann, 2013; Park, 2020; Woerfel, 2018) with the suggestion that such usage reflects a bilingual strategy to lessen the cognitive burden of processing two languages (cf. Silva-Corvalán, 2014; Filipović & Hawkins, 2019). This seems a plausible explanation because such descriptions occurred exclusively with ACROSS events, which entail a more demanding process of form-meaning mapping due to their inherent conceptual complexity (cf. Ji et al., 2011b; Özçalışkan, 2015).

Taking a more general look at bilingual children's language development, it is clear that language-specific factors, CLI, and child-universal tendencies all informed bilingual children's developmental trajectories. One noteworthy aspect of their L2 development concerns how CLI played out over time. Given that our bilinguals were relatively balanced in their two languages, and that they felicitously used the equipollently-framed pattern from the earliest stages, one wonders why CLI persisted for so long. Several possible reasons could be advanced. It could be that structural overlap is an important factor in CLI (Serratrice, 2013; 2022) and that bilinguals tend to use constructions that work in both languages (Filipović, 2019; 2022). But this does not explain why older children did not do so, and insights from work on crosslinguistic priming may shed light. In trying to explain why bilingual children are more susceptible to CLI than bilingual adults, Hsin et al. (2013) reason that adults usually succeed in suppressing the unwanted structure (due to co-activation) thanks to their more

developed inhibitory control skills. Younger bilinguals cannot necessarily do this because such abilities are slow to develop (cf. Bialystok, Craik & Luk, 2012). If this reasoning is correct, the decline of verb-framing in children's L2 may be linked to their more developed cognitive control with age. Admittedly speculative, this explanation seems plausible considering the absence of CLI in Uyghur-Chinese adult bilinguals' L2 Chinese (Tusun, 2022b).

This brings us to the issue of longevity of CLI. Recent reviews of CLI in early successive bilingualism concluded that CLI is a bilingual phenomenon rather than a developmental one (Chondrogianni, 2023; van Dijk et al., 2022). However, that bilingual children's use of the verb-framed pattern dropped to the adult level at 10, alongside Tusun's (2022b) finding that adult bilinguals showed no CLI suggest that it can indeed be a developmental phenomenon (Hulk, 2017). And I propose that this apparent inconsistency could be explained, at least partly, in terms of the sociological realities of Uyghur-Chinese bilingualism in Xinjiang. The aforementioned reviews sampled studies in Western immigration contexts where bilinguals' L1 is the heritage language and their L2 the societal language. But as discussed in Section 5, this distinction is blurred in Xinjiang because, among other things, Uyghurs represent nearly half of the population; and by virtue of its unique sociolinguistic dynamics (Elterish, 2016) and its educational system, children's exposure to and use of their L1 tend to be in comparable proportions to their L2 from kindergarten onwards. It is possible that these factors conspire to engender more balanced bilingualism (Chondrogianni, 2023; Unsworth et al., 2018) where CLI could become less detectable or less of a bilingual trait.

But beyond L1 influence, children's tuning in to the L2 equipollent system was a gradual process, as evidenced by the stepwise development of the equipollently-framed pattern and the corresponding increase of SD2 descriptions. That is, the process of developing thinking

for speaking in the L2 not only involved overcoming L1 influence but becoming incrementally sensitive to the frequency with which different lexicalisation patterns occur and to their distribution in the L2 (Wulff & Ellis, 2018; von Stutterheim, Lambert & Gerwien, 2021). Now, we know from previous research that Chinese speakers predominantly use the equipollently-framed pattern (i.e., RVCs) and much less frequently, the verb-framed pattern (Ji et al., 2011a; Shi et al., 2018; Wen & Shan, 2021). And the thinking-for-speaking account of language and cognition postulates that experience as a speaker in a given language community leads to the formation of cognitive processing routines or event frames that allow effortless and automatic information retrieval and organisation (Gerwien and von Stutterheim, 2021; Slobin, 1996). An important aspect of this experience, presumed to happen through constant exposure to the language (during language acquisition and daily language use), is the understanding of how frequently speakers of one's own community profile aspects of events by using certain linguistic structures under specific conditions (Gerwien and von Stutterheim, 2021; von Stutterheim, Gerwien, Bouhaous, Carroll & Lambert, 2020). Our bilingual children's task therefore was to understand that Chinese typically profiles both Manner and Path in an equipollent event frame. However, their (8-hour-daily) exposure to the L2 in a community where L1 and L2 share comparable dominance necessarily reduced opportunities for exposure and use of L2-specific event frames, and consequently, attaining automatic retrieval and use of L2-specific event frames may have warranted more accumulated exposure (Serratrice, 2022). This interpretation also accords with findings from L2 research that increased exposure enables learners to adapt to the statistical tendencies underlying L2 lexicalisation patterns (Treffers-Daller & Calude, 2015) and to adjust their initially L1-based predictions of motion encoding towards the target language distribution (Montero-Melis & Jaeger, 2020).

9. Conclusions

The study examined how Uyghur-Chinese early successive bilingual children acquired motion expressions in their L1 and L2, and how CLI shaped their L2 acquisition. Our findings showed that bilinguals' L1 Uyghur acquisition mirrored tendencies of children acquiring other V-languages whereas their L2 Chinese acquisition exhibited a distinct pattern from what is known for monolingual children. In both their L1 and L2, children were influenced by certain universal factors previously identified in child language research while their distinct L2 developmental path was shaped by the additional factor of CLI. Although the L1-to-L2 influence persisted for a period, it ultimately phased out, indicating that CLI can be a developmental phenomenon in naturalistic early successive bilingualism. Children converged on the L2 equipollent system at age 10, meaning that they eventually developed the ability to think for speaking in their L2, but that the same ability was already in place in 3-year-old monolingual Chinese children shows that the process of tuning in to the L2 system was gradual and incremental. Furthermore, that children's semantic density reached the adult level in Chinese but not in Uyghur demonstrated that speaking an E-language didn't boost semantic density in the V-language, suggesting that the development of their L1 versus L2 motion expression, at least in this respect, was relatively independent. This study is one of the few to investigate early successive bilingual children's acquisition of a conceptual/semantic domain from a developmental perspective. As such, it not only complements current research that has overwhelmingly focused on various aspects of morphosyntax, but also highlights potentially universal processes underlying monolingual and bilingual children's spatial language development. Importantly, by studying a non-Western bilingual community in which the boundary between the societally dominant versus the non-dominant language is less clear, its findings underscore the importance of studying bilingualism outside the Global North, and of factoring in the sociolinguistic realities and their unique affordances when

exploring and accounting for patterns of bilingual language acquisition and use. One limitation of the study, however, is the lack of child monolingual controls, especially for Chinese. Although our critical comparisons drew on previous studies on monolingual Chinese children that had utilised identical elicitation materials and analytical framework as the current study, having age-matched monolingual controls would have strengthened our observations, particularly regarding the role of CLI.

Competing interests declaration: The author declares none.

Supplementary Materials: Appendix S1: The list of experimental items; Appendix S2: The list of all planned contrasts; Appendix S3: All model outputs; Appendix S4: Absolute and relative frequencies of different categories for verb, OTH, and semantic density.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the author, without undue reservation.

References

- Aktan-Erciyes, A., Göksun, T., Tekcan, A. I. & Aksu-Koç, A. (2020). Children's thinking-for-speaking: Birectional effects of L1 Turkish and L2 English for motion events. *Linguistic Approaches to Bilingualism*, 11(5), 669-699.
- Aktan-Erciyes, A. (2020). Effects of second language on motion event lexicalisation: Comparison of bilingual and monolingual children's frog story narratives. *Journal of Language and Linguistic Studies*, 16(3), 1127-1145.
- 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*, 1, 16-48.
- Álvarez, E. (2008). The simultaneous development of narratives in English and Spanish. In C. Pérez-Vidal, M. Juan-Garau & A. Bel (eds.), *A Portrait of the Young in the New Multilingual Spain*, pp. 159-182. Bristol: Multilingual Matters.
- Aske, J. (1989). Path predicates in English and Spanish: A closer look. Paper presented at the 15th Annual Meeting of the Berkeley Linguistics Society, University of California, Berkeley.
- 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*, 4, 403-420.
- Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects model. *Frontiers in Psychology*, 4, 1-2.
- Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255-278.
- Bialystok, E., Craik, F. & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Science*, 16(4), 240-250.
- Bowerman, M., & Choi, S. (2003). Space under construction: Language-specific spatial categorization in first language acquisition. In D. Gentner & S. Goldin-Meadow (eds.), *Language in mind: Advances in the study of language and thought*, pp. 387-427. Cambridge Massachusetts: The MIT Press.
- Bullock, B. E., & Toribio, A. J. (2004). Introduction: Convergence as an emergent property in bilingual speech. *Bilingualism: Language and Cognition*, 2, 91-93.

- Bunger, A., Skordos, D., Trueswell, J. C., & Papafragou, A. (2021). How children attend to events before speaking- crosslinguistic evidence from the motion domain. *Glossa: a journal of general linguistics*, 6(1), 1-22.
- Chen, J. (2008). The acquisition of verb compounding in Mandarin Chinese. Ph.D. dissertation, Vrije Universiteit Amsterdam, Amsterdam.
- Chen, S. & Wu, L. (2023). Variable motion encoding within Chinese: a usage-based perspective. *Language and Cognition*, 1-23.
- Choi, S. & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalisation patterns. *Cognition*, 41, 83-121.
- Chondrogianni, V. (2018). Child L2 acquisition. In D. Miller, F. Bayram, J. Rothman & L. Serratrice (eds.), *Bilingual Cognition and Language: The state of the science across its subfields*, pp. 103-125. Amsterdam: John Benjamins.
- Chondrogianni, V. (2023). Cross-linguistic influences in bilingual morphosyntactic acquisition. In I. Elgort, A. Siyanova-Chanturia & M. Brysbaert (eds.), *Cross-linguistic Influences in Bilingual Processing and Second Language Acquisition*, pp. 294-315. Amsterdam: John Benjamins.
- Chondrogianni, V. & Vasić, N. (2016). Cross-linguistic aspects in child L2 acquisition. *International Journal of Bilingualism*, 20(4), 361-368.
- Croft, W., Barðdal, J., Hollmann, W., Sotirova, V., & Taoka, C. (2010). Revising Talmy's typological classification of complex event constructions. In H. C. Boas (eds.), *Contrastive studies in construction grammar*, pp. 201-235. Amsterdam: John Benjamins.

- 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, 1*, 95-119.
- Delage, A., & Frauenfelder, U. (2019). Syntax and Working Memory in typically-developing children. *Language, Interaction and Acquisition, 10*(2), 141–176.
- Elterish, A. (2015). The construction of Uyghur urban youth identity through language use. In J. S. Finley & X. Zang (eds.), *Language, Education and Uyghur Identity in Urban Xinjiang*, pp. 74-94. New York: Routledge.
- Elterish, A. (2016). Language use among Uyghur students in Xinjiang, PR China. In E. S. Ahn & J. Smagulova (eds.), *Language Change in Central Asia*, pp. 197-220. Berlin: De Gruyter Mouton.
- Engemann, H. (2013). Motion event expression in bilingual first language acquisition. Ph.D. dissertation, University of Cambridge.
- Engemann, H. (2016). Learning to think for speaking about space in child bilingualism. *Revue française de linguistique appliquée, 2*, 49-64.
- Engemann, H. (2021). How(not) to cross a boundary: Crosslinguistic influence in simultaneous bilingual children's event construal. *Bilingualism: Language and Cognition 25*, 42-54.
- Engemann H (2022) Entropy convergence in early bilinguals' syntactic packaging. *Frontiers in Psychology, 13:1010002*. doi: 10.3389/fpsyg.2022.1010002
- Feng, A. & Adamson, B. (2017). Language policies and sociolinguistic domains in the context of minority groups in China. *Journal of Multilingual and Multicultural Development, 1*-12.

- Filipović, L. (2011). Speaking and remembering in one or two languages: Bilingual vs. monolingual lexicalization and memory for motion events. *International Journal of Bilingualism*, 4, 466-485.
- Filipović, L. (2019). *Bilingualism in Action: Theory and Practice*. Cambridge: Cambridge University Press.
- Filipović, L. & Hawkins, J. (2019). The complex adaptive system principles model for bilingualism: language interactions within and across bilingual minds. *International Journal of Bilingualism*, 23, 1223–1248.
- Filipović, L. (2022). A multi-factor approach to the study of L2 acquisition of motion verbs and motion constructions: Integration of typological, psycholinguistic and sociolinguistic aspects. *Frontiers in Communication*, 7: 822517.
- Foroodi-Nejad, F., & Paradis, J. (2009). Crosslinguistic transfer in the acquisition of compound words in Persian-English bilinguals. *Bilingualism: Language and Cognition*, 14, 411-427
- Furman, R., Küntay, A. C., & Özyürek, A. (2014). Early language-specificity of children's event encoding in speech and gesture: evidence from caused motion in Turkish. *Language, Cognition and Neuroscience*, 5, 620-634.
- Gathercole, S. E., S. J. Pickering, B. Ambridge & H. Wearing. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, 40(2), 177-190.
- Gerwien, J. & von Stutterheim, C. (2021). Describing motion events. In A. H. Junker & H. Hausendorf (eds.), *Pragmatics of Space*, pp.153-179. Berlin: Walter de Gruyter.
- Guo, J., & Chen, L. (2009). Learning to express motion in narratives by Mandarin-speaking children. In J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura & Ş. Özçalışkan. (eds.), *Crosslinguistic Approaches to the Psychology of*

- Language. Research in the Tradition of Dan Isaac Slobin*, pp. 193-208. New York: Psychology Press.
- Harr, Anne-Katharina. (2012). *Language-specific factors in first language acquisition: The expression of motion events in French and German*. Walter de Gruyter.
- Harr, A., & Hickmann, M. (2013). How German and French children express voluntary motion. In C. Paradis, J. Hudson & U. Magnusson (eds.), *The Construal of Spatial Meaning: Windows into Conceptual Space*, pp. 194-213. Oxford: Oxford University Press.
- 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*, 4, 705-741.
- Hickmann, M., Hendriks, H., & Harr, A. (2018). Caused motion across child languages: a comparison of English, German and French. *Journal of Child Language*, 1-28.
- 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*, 3, 249-261.
- Hulk, A. (2017). Note on cross-linguistic influence: back to ‘MULK’. In E. Blom, L. Cornips, and J. Schaeffer (eds.), *Cross-Linguistic Influence in Bilingualism: In Honor of Aafke*, pp.15–24. Amsterdam: John Benjamins.
- Ibrahim, A. (1995). Meaning and usage of compound verbs in modern Uyghur and Uzbek. Ph.D. dissertation, University of Washington, Ann Arbor.
- Ibarretxe-Antuñano, Iraide. (2009). Path salience in motion events. In J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura & Ş. Özçalışkan (eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin*, 403–414. New York: Psychology Press.

- Ji, Y. (2009). The expression of voluntary and caused motion events in Chinese and in English: Typological and developmental perspectives. Ph.D. dissertation, University of Cambridge.
- Ji, Y. (2014). *The acquisition of motion events in Chinese: Typological and developmental perspectives*. Beijing: The Chinese Social Sciences Press.
- Ji, Y., Hendriks, H., & Hickmann, M. (2011a). Children's expression of voluntary motion events in English and Chinese. *Journal of Foreign Languages*, 4, 1-20.
- Ji, Y., Hendriks, H., & Hickmann, M. (2011b). How children express caused motion events in Chinese and English: Universal and language-specific influences. *Lingua*, 121, 1796-1819.
- Ji, Y., Hendriks, H., & Hickmann, M. (2011c). The expression of caused motion events in Chinese and in English: Some typological issues. *Linguistics*, 5, 1041-1077.
- 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*, 1, 45-76.
- Ji, Y. (2019). Cognitive representation of spontaneous motion in a second language: An exploration of Chinese learners of English. *Frontiers in Psychology*, 10:2706.doi.org/10.3389/fpsyg.2019.02706.
- Johanson, L. (1995). On Turkic converb clauses. In M. Haspelmath & E. König (eds.), *Converbs in cross-linguistic perspective*, pp. 313-348. Berlin & New York: Mouton de Gruyter.

- Johanson, L. (1998). The Structure of Turkic. In L. Johanson & É. Á. Csato (eds.), *The Turkic Languages*, pp. 30-66. London: Routledge.
- Johanson, L. (2009). Uyghur. In K. Brown & S. Ogilvie (eds.), *Concise encyclopaedia of languages of the world*, pp. 1142-1145. Oxford: Elsevier.
- Johanson, L. (2021). *Turkic*. Cambridge: Cambridge University Press.
- Johanson, L. & Csato, É. Á. (eds.) (1998). *The Turkic Languages*. London: Routledge.
- Lamarre, C. (2003). Hanyu kongjian weiy shijian de yuyan biaoda- jianlun shuqushi de jige wenti (Linguistic encoding of motion events in Chinese: Several issues on verb-directional constructions). *Xianzai zhongguo yu yanjiu* (Studies on Modern Chinese), 5, 1-18.
- Lewandowski, W. & J. Mateu. (2020). Motion events again: Delimiting constructional patterns. *Lingua*, 247, 1–25.
- Levin, B. & Rappaport Hovav, M. (2019). Lexicalization Patterns. In R. Truswell (eds.), *The Oxford Handbook of Event Structure*, pp. 395-425. Oxford: Oxford University Press.
- Ma, R. (2012). The development of minority education and the practice of bilingual education in Xinjiang Uyghur Autonomous Region. In G.H.Beckett & G.A.Postiglione (eds.), *China's Assimilationist Policy: The impact on indigenous/minority literacy and social harmony*, pp. 33-74. New York: Routledge.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for Analysing Talk*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Meisel, J. M. (2018). Bilingual Acquisition: A Morphosyntactic Perspective on Simultaneous and Early Successive Language Development. In E.M. Fernández &

- H. S. Cairns (eds.), *The Handbook of Psycholinguistics*, pp. 635-652. London: Blackwell.
- Miller, N., Furman, R., & Nicoladis, E. (2018). French-English bilingual children's motion event communication shows crosslinguistic influence in speech but not gesture. *Language, Interaction and Acquisition/Langage, Interaction et Acquisition, 1*, 69-100.
- Montero-Melis, G. & Jaeger, T. F. (2020). Changing expectations mediate adaption in L2 production. *Bilingualism: Language and Cognition, 23*, 602-617.
- Özçalışkan, Ş. (2009). Learning to talk about spatial motion in language-specific ways. In J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura & Ş. Özçalışkan (eds.), *Cross-linguistic approaches to the psychology of language: Research in the tradition of Dan Issac Slobin*, pp. 263-276. New York & London: Psychology Press.
- Özçalışkan, Ş. (2015). Ways of crossing a spatial boundary in typologically distinct languages. *Applied Psycholinguistics, 36*, 485-508.
- Özçalışkan, Ş. & Slobin, D. (1999). Learning how to search for the frog: Expression of manner of motion in English, Spanish, and Turkish. In A. Greenhill, H. Littlefield & C. Tano (eds.), *Proceedings of the 23rd annual Boston University conference on language development*, pp. 541-552. Somerville, MA: Cascadilla Press.
- Özçalışkan, Ş., & Slobin, D. (2003). Codability effects on the expression of manner of motion in Turkish and English. In A. S. Özsoy, D. Akar, M. Nakipoglu-Demiralp, E. E. Taylan & A. Aksu-Koç (eds.), *Studies in Turkish linguistics*, pp. 259-270. Istanbul: Bogaziçi University Press.
- Özyürek, A., & Özçalışkan, Ş. (2000). How do children learn to conflate manner and path in their speech and gestures? Differences in English and Turkish. In E. Clark

- (eds.), *Proceedings of the 30th Stanford Child Language Research Forum*, pp. 77-85. Palo Alto, California: CSLI Publications.
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, *1*, 155-184.
- Paradis, J. & Nicoladis, E. (2007). The influence of dominance and sociolinguistic context on bilingual preschoolers' language choice. *The International Journal of Bilingualism and Bilingual Education*, *10*, 1–21.
- Park, H. I. (2020). How do Korean-English bilinguals speak and think about motion events? Evidence from verbal and non-verbal tasks. *Cognition*, *23*, 483-499.
- Ragagnin, E. Uyghur language. In R. Sybesma (eds.), *Encyclopaedia of Chinese language and linguistics*, http://dx.doi.org/10.1163/2210-7363_ecll_COM_00000435.
Published online by Brill, 2015.
- R Core Team. (2013). R: A language and environment for statistical computing. Vienna Austria: R Foundation for Statistical Computing. <https://www.r-project.org/>
- Serratrice, L. (2013). Cross-linguistic influence in bilingual development: Determinants and mechanisms. *Linguistic Approaches to Bilingualism*, *1*, 3-25.
- Serratrice, L. (2022). What can syntactic priming tell us about crosslinguistic influence? In K. Messenger (eds.), *Syntactic priming in language acquisition: Representations, mechanisms and applications*, pp. 129-156. Amsterdam: John Benkamins.
- Shi, W., Yang, W., & Su, H. (2018). The typological change of motion expressions: Motion events in Old Chinese and its Modern Chinese translation. *Studies in Language*, *4*, 847-885.
- Silva-Corvalán, C. (2014). *Bilingual Language Acquisition: Spanish and English in the First Six Years*. Cambridge: Cambridge University Press.

- Slobin, D. (1996). 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. (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. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömquist & L. Verhoveen (eds.), *Relating events in narrative: Typological and contextual perspectives*, pp. 219-257. New York: Psychology Press.
- Slobin, D. (2005). Linguistic representations of motion events: What is signifier and what is signified? In C. Maeder, O. Fischer & W. Herlofsky (eds.), *Iconicity inside out: Iconicity in language and literature*, pp. 307-322. Amsterdam: John Benjamins.
- Slobin, D. (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: John Benjamins.
- Slobin, D., & Hoiting, N. (1994). Reference to movement in spoken and signed languages: Typological considerations. *Annual Meeting of the Berkeley Linguistics Society, 1*, 487-505.
- Sunuodula, M. & Cao, Y. (2015). Language learning and empowerment: Languages in Education for Uyghurs in Xinjiang. In A. Feng & B. Adamson (eds.), *Trilingualism in Education in China: Models and Challenges*, pp. 65-102. New York: Springer.

- Talmy, L. (2000). *Toward a cognitive semantics: Typology and process in concept structuring*. Cambridge Massachusetts: The MIT press.
- Talmy, L. (2016). Properties of Main Verbs. *Cognitive Semantics*, 2, 133-163.
- 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.
- Tusun, A. (2022a). Motion events in Modern Uyghur narrative discourse. In C. Shei & S. Li (eds.), *The Routledge Handbook of Asian Linguistics*, pp.93-109. New York: Routledge.
- Tusun, A. (2022b). Uyghur-Chinese adult bilingual's construal of voluntary motion events. *Frontiers in Psychology*, 13:892346. doi: 10.3389/fpsyg.2022.892346.
- Tusun, A., & Hendriks, H. (2019). Voluntary motion events in Uyghur: A typological perspective. *Lingua*, 226, 69-88.
- Unsworth, S. (2016). Amount of exposure as a proxy for dominance in bilingual language acquisition. In C. Silva-Corvalan & J. Treffers-Daller (eds.), *Language Dominance in Bilinguals: Issues of Measurement and Operationalisation*, pp. 156-173. Cambridge: Cambridge University Press.
- Unsworth, S., Chondrogianni, V. & Skarabela, B. (2018). Experiential measures can be used as a proxy for language dominance in bilingual language acquisition research. *Frontiers in Psychology*. 9:1809. doi: 10.3389/fpsyg.2018.01809
- van Dijk, C., van Wonderen, E., Koutamanis, E., Kootstra, G. J., Dijkstra, T. & Unsworth, S. (2022). Cross-linguistic influence in simultaneous and early sequential bilingual children: a meta-analysis. *Journal of Child Language*, 49, 897-929.

- 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*, *1*, 207-249.
- von Stutterheim, C. Gerwien, J., Bouhaous, A., Carroll, M. & Lambert, M. (2020). What makes up a reportable event in a language? Motion events as an important test domain in linguistic typology. *Linguistics*, *58(6)*, 1659-1700.
- von Stutterheim, C., Lambert, M. & Gerwien, J. (2021). Limitations of the role of frequency in L2 acquisition. *Language and Cognition*, 1-31.
- Wang, Y., & Wei, L. (2021). Cognitive restructuring in the multilingual mind: language-specific effects on processing efficiency of caused motion events in Cantonese-English-Japanese speakers. *Bilingualism: Language and Cognition*, 1-16.
- Wen, X. & Shan, X. (2021). The adventure of a third way: Motion events in Mandarin. *Language Sciences*, *85*, 1-24.
- Woerfel, T. (2018). *Encoding Motion Events: The Impact of Language-Specific Patterns and Language Dominance in Bilingual Children*. Berlin: De Gruyter Mouton.
- Wulff, S. & Ellis, N. (2018). Usage-based approaches to second language acquisition. In D. Miller, F. Bayram, J. Rothman & L. Serratrice (eds.), *Bilingual Cognition and Language: The state of the science across its subfields*, pp. 37-56. Amsterdam: John Benjamins.
- Yi, V. & Matthews, S. (2022). Language diversity and bilingual first language acquisition: A commentary on Kiss and Garcia (2022). *First Language*, 1-5.
- Zang, X. (2015). Major determinants of Uyghur ethnic consciousness in Ürümqi. In J. S. Finley & X. Zang (eds.), *Language, Education and Uyghur Identity in Urban Xinjiang*, pp. 34-54. New York: Routledge.

Zheng, L. (2011). *Media and Minkaohan Uyghurs: Representation, Reaction and Resistance*. Ph.D. Dissertation, University of Colorado, Denver.

¹ Throughout the paper, Chinese refers to Mandarin Chinese.

² The symbols for transliteration are based on *The Turkic Languages* (Johanson & Csató, 1998). The capitalized abbreviations used are: ABL-ablative case, ACC-accusative case, ASP_{dur}-durative aspectual marker, ASP_{perf}-perfective aspectual marker, ASPV-aspectual verb, CONV-converb, DAT-dative case, GEN-genitive case, POS-possessee marker, PST-past tense, SG-singular.

³ The Chinese adult data were collected for Ji et al. (2011a), and I thank the authors for sharing their data. I should mention that practical constraints made it impossible to collect age-matched monolingual Chinese data, which may be problematic for addressing the second half of RQ2. However, Ji et al. (2011a) using identical tasks and analytical framework showed that Chinese children's motion expression was fully adult-like from age 3 with no developmental change until age 10 (cf. Section 3). We can therefore safely assume that differences (if any) between bilinguals and Chinese adults in using verb-framed patterns is due to L1 influence, not to developmental factors.

⁴ Error bars represent stand error of means. Information on the absolute and relative frequencies of the different categories for the verb locus, OTH locus and semantic density is given in Appendix S4.

⁵ As an anonymous reviewer rightly points out, our adult speakers also spoke Chinese and English, which may have boosted their semantic density, and the pattern may be different if monolingual Uyghur speakers are incorporated as the benchmark.