Bilingualism: Language and Cognition

Research Article


Received: 18 February 2022
Revised: 17 October 2023
Accepted: 22 October 2023

Keywords: motion events; Uyghur; Chinese; early successive bilingual

Corresponding author:
Alimujiang Tusun
Email: a648@cam.ac.uk

Abstract

This study explores the implications of Talmy’s (2000) motion event typology and its subsequent articulations in relation to Slobin’s (1996, 2006) THINKING-FOR-SPEAKING hypothesis for the early successive bilingual acquisition of Uyghur (verb-framed) and Mandarin Chinese (equipollently-framed). Specifically, it examines how 4-, 6-, 8- and 10-year-old bilingual children acquire motion expressions in their L1 and L2 respectively, and how cross-linguistic influence shapes their L2 acquisition process. Results show that, in their L1 Uyghur, bilinguals follow general developmental trajectories observed for children acquiring verb-framed languages. While sensitive to the equipollent Chinese system from early on, due to L1 and other factors, bilinguals fully converge on the Chinese pattern only at age 10, a feat in place in monolinguals from age 3. Our findings highlight that bilingual children do eventually come to develop language-specific THINKING-FOR-SPEAKING patterns in their L2, but they traverse a distinct developmental path.

1. Introduction

The crosslinguistic encoding of motion events (Talmy, 2000) has provided a fruitful venue for exploring bilingual language acquisition and use (Daller et al., 2011; Hohenstein et al., 2006; Wang & Wei, 2021). This study extends this line of research to the context of Uyghur–Chinese 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 (cf. van Dijk et al., 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 (cf. Hendriks et al., 2022), 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 et al., 2022). 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).
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 (Burger et al., 2021; Filipović, 2011; Ji & Hohenstein, 2017, 2018; Slobin, 1996, 2003, 2004, 2006; von Stutterheim et al., 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 are 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 typically have to use syntactically complex constructions (e.g., subordination) that incur greater processing load (Özçalıskan & Slobin, 2003; Slobin, 2004, 2006). During online production, speakers fit their conceptualisation of an 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 et al., 2007; Özçalıskan, 2015; Tusun, 2022a; Tusun & Hendriks, 2019; 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). 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 convverb, 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 yüger-üp [Manner] kir-di [Path].
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, akin 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 Path-encoding morphemes can function as main verbs, as in (7). It has therefore been argued that the two verbal elements in an RVC share the same grammatical status and formal significance, and that Chinese is an equipollently-framed language (E-language) (Slobin, 2004). In response, Talmy (2009) proposed a set of properties characteristic of main verbs, and recent studies testing his proposal support the claim that Chinese is an E-language (Wen & Shan, 2021; see also Talmy, 2016). Importantly, numerous usage-based studies (e.g., Ji et al., 2011c; Lamarre, 2003; Wen & Shan, 2021) have shown that Chinese speakers’ motion descriptions are semantically as dense as those of S-language speakers, and that the verb-framed option, exemplified in (7), is frequently used in Chinese. For instance, in a study based on a one-million-word corpus, Chen and Wu (2023) report that the verb-framed option accounts for about 24% of their data. Taking these together, I consider Chinese as an E-language with verb-framing tendencies.

(6) Ma3li4ya4 pao3 jin4 le Jiao4shi4.
[V1] [V2]
Mary run-enter classroom ASPperf.
Mary ran into the classroom.

(7) Ma3li4ya4 (pao3zhe1) jin4 le Jiao4shi4.
Mary (run ASPdur) enter classroom
Mary entered the classroom (by running)

4. Motion expressions in L1 and bilingual contexts

L1 research has shown that children’s earliest productions reflect the typological tendencies of their ambient language (Chen, 2008; Choi & Bowerman, 1991), and, by age 3, they largely follow language-specific lexicalisation patterns (Allen et al., 2007; Bowerman & Choi, 2003; Guo & Chen, 2009; Hickmann et al., 2018). However, children’s ability to produce semantically dense motion descriptions develops 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 have been found to outperform their S-language peers. Thus, Ji et al. (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 (Hendriks et al., 2022; Ji et al., 2011b).

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 (Serrattrice, 2013). Regardless of whether it is simultaneous child bilinguals (Engemann, 2021; Miller et al., 2018) or successive child (Aktan-Erciyes, 2020; Aktan-Erciyes et al., 2020; Aveledo & Athanasopoulos, 2016; 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 an 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 an 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 Avedello (2015) and Aveledo and Athanasopoulos (2016) on Spanish–English child bilinguals (aged 5–7 vs. 8–9, AoO=3–4) in Venezuela: that only the older bilinguals showed an 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, 2016; 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 pairings (Serrattrice, 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., 2022) 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 featuring a verb-framed language and an equipollently-framed language that are distantly genetically (Turkic vs. Sino–Tibetan), and distinct in their general linguistic profiles (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 (for insights into the sociology and politics of Uyghur–Chinese bilingualism, see Elterish, 2015; Zang, 2015). 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 naturally. 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 use and exposure can serve as a proxy for language proficiency and language dominance (Unsworth, 2016; Unsworth et al., 2018), the sociolinguistic setting in question is arguably more conducive to more balanced bilingualism (Filipović, 2019), and as such, Uyghur–Chinese bilingual children’ 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., 2022 for French; Aktan-Erciyes et al., 2020 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, 2012 for Turkish). However, their 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 would omit Manner; they would therefore produce low-density descriptions (cf. Hendriks et al., 2022; 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 could 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., 2022; Ji et al., 2011a), bilinguals’ semantic density would be higher in Chinese than in Uyghur.

6. The study

6.1. Participants

The participants fell into 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 Beijing.

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 Ürümqi, 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 had been 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. 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 rather than on gestures,
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 dere₂-ke yami₂-ip [c1] čiq₂-ti [c2]
  bear slide DAT climb CONV ascend PST.3SG
  The bear ascended the tree by climbing.

(9) Ta1 hua2 zhe1 bing1 [c1] guo4 he2 [c2]
  run cross road ABL
  He crossed the river running.

(10) Adem yol-din öt-ti. [R1]
  man road ABL cross PST.3SG
  That man crossed the road.

(11) Ta1 pao₃bu₄ [R1] ran₂hou₄ guo₄qu₁ le [R2]
  boy run cross PST.3SG
  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., 2022; Wang & Wei, 2021), two loci were identified: the main verb (the verb locus); and 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., fei₁su₄de ‘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 loci (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) Çaşqan üstel-ge čiq₂-ti [Path].
  mouse table DAT climb PST.3SG
  The mouse ascended to the table.

(13) Qurut dere₂-ge yamaşı-ti [Manner].
  caterpillar tree DAT climb PST.3SG
  The caterpillar climbed the tree.

(14) Zhe₂₄ge₂ren₂ cong₂ ma₃lu₄ shang₄ guo₄qu₁ [Path] le₁.
  This man from road on cross PST Perf
  This man crossed from the road.

(15) Song₁shu₃ pa₂ [Manner] shu₄
  squirrel climb tree
  The squirrel climbed the tree.

(16) Hou₂zi₁ cong₂ shu₄ [Path] pa₂xia₄laï₂ [Path+Manner].
  monkey from tree on climb descend come
  The monkey climbed down from the tree.

(17) Müsük türük-niñ üst₀₂-ti [Path].
  cat pole GEN top ABL descend PST.3SG
  The cat descended from the top of the tree.

(18) Bala tilin-ip [Manner] öt-ti.
  boy slide CONV cross PST.3SG
  The boy crossed by sliding.

(19) Bir bala yol-din [Path] yuğur₂up [Manner] a yaq qa [Path].
  one boy road ABL run CONV that side DAT
  A boy went to that side of the road by running.

(20) Éyiq ču₀₂-s₂-up ket-ti.
  bear descend CONV PST.3SG
  The bear descended.

  caterpillar from grass on crawl descend come
  The caterpillar crawled down from the grass.
For the statistical analyses, our independent variable was age whereas the 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, 2013), the glmer() function in the lme4 library. We first fitted a model to the same dataset with the fixed effects in question, against a reduced model without the fixed effects in question. 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 age\(^6\). 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 (\(\beta_{6yrs-AD} = -0.81, SE = 0.32, Wald z = -2.53, p = .011; \beta_{8yrs-AD} = -1.26, SE = 0.27, Wald z = -4.69, p < .001\)). 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 (\(\chi^2(4)=68.828, p < .001\)). Follow-up analyses revealed that 4-, 6- and 8-year-olds encoded Path more frequently than adults (\(\beta_{4yrs-AD} = -0.78, SE = 0.17, Wald z = -4.40, p < .001; \beta_{6yrs-AD} = -0.59, SE = 0.17, Wald z = -3.35, p < .001; \beta_{8yrs-AD} = -0.63, SE = 0.21, Wald z = -2.96, p < .003\)), and 4- and 6-year-olds encoded Manner more frequently than adults (\(\beta_{4yrs-AD} = -1.58, SE = 0.33, Wald z = -4.79, p < .001, \beta_{6yrs-AD} = -1.45, SE = 0.30, Wald z = -4.76, p < .001\)). Finally, only 4- and 6-year-olds used the Path+Manner pattern less frequently than adults (\(\beta_{6yrs-AD} = -0.98, SE = 0.15, Wald z = 6.24, p < .001; \beta_{6yrs-AD} = 0.58, SE = 0.11, Wald z = 4.87, p < .001\)), the steady increase of this pattern within the four child groups was significant at each age level (\(\beta_{6yrs-6yrs} = -0.40, SE = 0.16, Wald z = 2.47, p = .013; \beta_{6yrs-8yrs} = 0.34, SE = 0.13, Wald z = 2.53, p = .011; \beta_{8yrs-10yrs} = -0.35, SE = 0.10, Wald z = -3.34, p < .001\)). 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.

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 ($\beta_{4\text{yrs}-\text{AD}} = -0.92$, SE = 0.12, Wald $z = -7.46$, $p < .001$; $\beta_{6\text{yrs}-\text{AD}} = -0.91$, SE = 0.33, Wald $z = -3.00$, $p < .001$; $\beta_{8\text{yrs}-\text{AD}} = -1.01$, SE = 0.12, Wald $z = -8.30$, $p < .001$; $\beta_{10\text{yrs}-\text{AD}} = -0.74$, SE = 0.14, Wald $z = -5.23$, $p < .001$), and Path + Manner less frequently than adults ($\beta_{4\text{yrs}-\text{AD}} = 0.98$, SE = 0.13, Wald $z = 7.40$, $p < .001$; $\beta_{6\text{yrs}-\text{AD}} = 0.75$, SE = 0.12, Wald $z = 6.15$, $p < .001$; $\beta_{8\text{yrs}-\text{AD}} = 0.60$, SE = 0.11, Wald $z = 5.17$, $p < .001$; $\beta_{10\text{yrs}-\text{AD}} = 0.34$, 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 ($\beta_{4\text{yrs}-\text{AD}} = 1.21$, SE = 0.33, Wald $z = 3.60$, $p < .001$; $\beta_{6\text{yrs}-\text{AD}} = 1.69$, SE = 0.39, Wald $z = 4.32$, $p < .001$) and Path+Manner less frequently ($\beta_{4\text{yrs}-\text{AD}} = 2.91$, SE = 1.11, Wald $z = -2.61$, $p = .008$; $\beta_{6\text{yrs}-\text{AD}} = 1.80$, SE = 0.73, Wald $z = 2.44$, $p = .014$) while all child groups produced Zero more frequently that adults ($\beta_{4\text{yrs}-\text{AD}} = 0.41$, SE = 0.10, Wald $z = -4.06$, $p < .001$; $\beta_{6\text{yrs}-\text{AD}} = 0.37$, SE = 0.10, Wald $z = -3.57$, $p < .001$; $\beta_{8\text{yrs}-\text{AD}} = 0.21$, SE = 0.10, Wald $z = -2.06$, $p = .039$; $\beta_{10\text{yrs}-\text{AD}} = -0.21$, 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.

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 ($\beta_{\text{SYT-AD}}=-0.82$, SE $=0.10$, Wald $z=-7.63$, $p<.001$; $\beta_{\text{SYT-AD}}=-0.63$, SE $=0.13$, Wald $z=-4.84$, $p<.001$; $\beta_{\text{SYT-AD}}=-0.80$, SE $=0.12$, Wald $z=-6.25$, $p<.001$; $\beta_{\text{SYT-AD}}=-0.63$, SE $=0.12$, Wald $z=-5.01$, $p<.001$) but SD2 descriptions less frequently than adults ($\beta_{\text{SYT-AD}}=-0.69$, SE $=0.10$, Wald $z=6.62$, $p<.001$; $\beta_{\text{SYT-AD}}=-0.36$, SE $=0.10$, Wald $z=-3.48$, $p<.001$; $\beta_{\text{SYT-AD}}=0.58$, SE $=0.12$, Wald $z=4.77$, $p<.001$; $\beta_{\text{SYT-AD}}=-0.34$, 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, the stepwise decrease of SD1 descriptions was significant at each age group ($\beta_{\text{SYT-6YR}}=-0.17$, SE $=0.08$, Wald $z=-2.12$, $p=0.033$; $\beta_{\text{SYT-8YR}}=-0.38$, SE $=0.11$, Wald $z=-3.20$, $p<.001$; $\beta_{\text{SYT-10YR}}=-0.60$, SE $=0.19$, Wald $z=-3.19$, $p<.001$; $\beta_{\text{SYT-AD}}=-0.63$, SE $=0.27$, Wald $z=-2.23$, $p=0.019$) and all child groups produced SD1 descriptions more frequently than adults ($\beta_{\text{SYT-AD}}=-1.79$, SE $=0.20$, Wald $z=-8.81$, $p<.001$; $\beta_{\text{SYT-AD}}=-1.62$, SE $=0.20$, Wald $z=-7.90$, $p<.001$; $\beta_{\text{SYT-AD}}=-1.23$, SE $=0.24$, Wald $z=-5.01$, $p<.001$; $\beta_{\text{SYT-AD}}=-0.63$, SE $=0.27$, Wald $z=-2.33$, $p=0.019$). Meanwhile, the increase of SD2 descriptions was significant from age 4 to 10 ($\beta_{\text{SYT-6YR}}=-0.39$, SE $=0.15$, Wald $z=-2.51$, $p=0.011$; $\beta_{\text{SYT-8YR}}=-0.38$, SE $=0.12$, Wald $z=-3.11$, $p=0.001$; $\beta_{\text{SYT-10YR}}=-0.28$, SE $=0.09$, Wald $z=-3.21$, $p<.001$). At age 10, children attained the adult level of frequency for SD2 descriptions.

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., 2022; Hickmann et al., 2018), we adopted a productivite 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).

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 the 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 expected, 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: 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 supported 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 as this study, 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 a 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, 2016; Hohenstein et al., 2006; Park, 2020).

Another pattern that contributed to the 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 swim 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 parallel 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 involve a categorical change of location (cf. Hendriks et al., 2022; Hickmann et al., 2018). For UP events children unanimously used the Manner verb pa2 ‘to climb’, and interestingly, they used its equivalent yaməmaŋq in their Uyghur descriptions. Lewandowski and 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; Hendriks et al., 2022; Özçalıskan & Slobin, 1999; Özyürek & Özçalıskan, 2000). 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 the bilinguals displayed in comparison to the 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 doa4 ‘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. Filipović & Hawkins, 2019; Silva-Corvalán, 2014). 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; Özçalıskan, 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 et al., 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.
that Chinese speakers predominantly use the equipollently-framed pattern to express motion (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 pattern 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 their distribution in the L2 (von Stutterheim et al., 2021; Wulff & Ellis, 2018). 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 & 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 & von Stutterheim, 2021; von Stutterheim et al., 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 could 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 the fact 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.

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.

Competing interests declaration. The author declares none.

Supplementary Material. For supplementary material accompanying this paper, visit https://doi.org/10.1017/S1366728923000780

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.

Notes
1 Throughout the paper, Chinese refers to Mandarin Chinese.
2 The symbols for transliteration are based on The Turkic Languages (Johnson & Csató, 1998). The capitalized abbreviations used are: ABL-ablative case, ACC-
accusative case, ASP\textsubscript{durative} aspectual marker, ASP\textsubscript{perf} -perfective aspectual marker, ASPV-aspectual verb, CONV-converb, DAT-dative case, GEN-genitive case, PST-past tense, SG-singular.

5 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 standard 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.

References


