Bilingualism: Language and Cognition

cambridge.org/bil

Grammatical development of the native L1 in Cantonese–English bilingual children: early costs and long-term gains

Yuqing Liang[1] and Ziyin Mai[2]

[1]Department of Linguistics and Translation, City University of Hong Kong, Hong Kong, China and [2]Department of Linguistics and Modern Languages, Chinese University of Hong Kong, Hong Kong, China

Research Article

Cite this article: Liang, Y. and Mai, Z. (2025). Grammatical development of the native L1 in Cantonese–English bilingual children: early costs and long-term gains. *Bilingualism: Language and Cognition* 1–14. https://doi.org/10.1017/S1366728925100412

Received: 27 August 2024
Revised: 22 May 2025
Accepted: 10 July 2025

Keywords:
early bilingual; Cantonese; English; input; grammar

Corresponding author:
Ziyin Mai;
Email: maggiezymai@cuhk.edu.hk

This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

Abstract

Previous studies show that bilingual toddlers who develop their first language (L1) alongside another language can show early stabilization in the L1. This study investigates grammatical development of L1 Cantonese in children with very early onset of English before age 3 (earlier-onset bilinguals/EB, n = 31), with matched later-onset bilinguals (LB, n = 21) as the baseline. Input characteristics and child development measures at 3;0 and 5;8 were derived from parental reports, caretaker–child toy play and narration tasks. Results show that at 3;0, when the LB children were monolingual, the EB children were below the LB group in general grammatical complexity and seven specific grammatical structures ('early costs'). At 5;8, the EB children converged with the LB children across grammatical measures in Cantonese, while demonstrating superior performance in English ('long-term gains'). Our findings reveal a distinctive velocity of L1 development in early additive bilinguals raised in a bilingual society.

Highlights

- At 3;0, early-onset bilinguals displayed reduced grammatical complexity in L1.
- But they were only weaker than matched monolinguals in a subset of structures.
- Quality of L1 input accounts for variation among the bilinguals after input amount.
- At 5;8, early-onset bilinguals converge with later-onset peers in L1 grammar.
- L1 development in early bilinguals benefited from L1 support in bilingual contexts.

1. Introduction

Compared to monolingual peers, infants and toddlers who regularly hear two languages from caretakers necessarily receive reduced input proportions in each language ('input reduction'). This can lead to later-than-monolingual development in both languages (Byers-Heinlein et al., 2023; Cote & Bornstein, 2014; Hoff et al., 2012). Yet bilingual children's linguistic environments and acquisition outcomes vary. Abundant research has examined bilingual development in 'minority–majority language contexts', in which the child's first language (L1) is a minority language not widely used outside the home and with lower social status, while the child's L2 (or another L1) is the societal majority language. Such bilingual children display later-than-monolingual development in the minority language (e.g., heritage languages of immigrant children, Cote & Bornstein, 2014; Hoff et al., 2012; Montrul, 2016). The development of early bilinguals' L1 in contexts where the L1 enjoys high social status is less studied. In such contexts, caretakers proactively promote early additive bilingualism (the learning of an additional language, aiming for high proficiency in both languages) by addressing children in the majority language (L1) and an additional language from a young age.

This linguistic environment is exemplified in Cantonese–English bilingual children raised by Cantonese-dominant parents in Hong Kong SAR, where both English and Chinese are official languages. High proficiency in both languages is promoted by the government, prized by local parents, and aimed for in education. Cantonese is the societal-dominant language, spoken by around 85% of the population as the primary language. Approximately 49.9% of adults regularly speak English to their children aged below 6 (Census & Statistics Department of the Government of Hong Kong Special Administrative Region, 2022). Although Cantonese acquisition in a bilingual setting is the norm for Hong Kong children, different types of bilingualism occur. A small group of children develop simultaneous bilingualism under the one-parent-one-language practice with one native Cantonese-speaking and one native English-speaking parent (e.g., Yip & Matthews, 2007), while the majority are raised in Cantonese-speaking homes and develop bilingualism through English exposure from preschool. Others are raised under a third input model, differing from the above two groups in that these children receive both native Cantonese and non-native English input during naturalistic interactions with caretakers before the age of 3.

Their Cantonese input providers are (grand)parents who speak Cantonese as their first and dominant language, and English input providers are parents or live-in foreign domestic helpers who speak English as an additional and (usually) weaker language. Many of them acquire Cantonese and English almost simultaneously from birth. Whether English can be considered an L1 alongside Cantonese in this scenario remains debatable. Our study focuses on the development of Cantonese in this third group of bilingual children, who have not hitherto been studied systematically, and will refer to Cantonese as their 'native L1' for clarity of reference. We focused on bilingual children's L1 Cantonese grammar development in a bilingual society where Cantonese is one of the widely spoken languages. These children, like those raised by immigrant parents in the US and UK, develop their L1 under reduced-input conditions, yet they differ from immigrant children, as their L1 is the societal-dominant language supported by mainstream education.

2. Lexical and grammatical development of L1 in bilingual toddlers and preschoolers

Compared to monolingual children, bilingual toddlers learning a minority L1 in an English-dominant society consistently show weaker lexical skills, especially in production, due to input reduction in the native L1. Cote and Bornstein (2014) found that bilingual 20-month-olds in the US acquiring Spanish or Korean had smaller vocabularies than monolingual peers. Miękisz et al. (2019) discovered that, despite a strong presence of L1 at home, Polish–English toddlers (23–27 months) in the UK displayed lower vocabulary scores than monolingual Polish children. Similar findings have been reported for bilingual toddlers in societies with two dominant languages. French–English bilingual toddlers in Montréal showed reduced productive vocabulary sizes and slower development rates than monolinguals even in their dominant language, be it French or English (Byers-Heinlein et al., 2023). Overall, robust evidence indicates input reduction effects on bilingual toddlers' L1 lexical development across bilingual contexts.

Whether input reduction affects grammatical development to the same extent as in lexical development in toddlerhood is less clear. Prior research has examined grammatical development in the majority L2 (e.g., Foursha-Stevenson et al., 2023; Hoff et al., 2012) or the native L1 in older children (e.g., preschool/school age in Hao & Chondrogianni, 2023; school age in Jia & Paradis, 2015; Mai et al., 2022), but findings are mixed (e.g., Thordardottir, 2014; Unsworth, 2014). Few studies have directly examined grammatical development in bilingual toddlers' L1, factoring in variations in the input. The most relevant study is Blom (2010), which investigated grammatical development in four bilingual children (2;0–3;6) learning Turkish in Dutch-speaking Netherlands using a general measure of syntactic complexity (i.e., mean length of utterance in words, MLUw). Two of the bilingual children showed lower MLUw and less steady development than Turkish monolinguals, while the other two bilinguals demonstrated baseline-like performance.

From preschool onwards, language learning increasingly takes place outside the home. The societal linguistic environment (e.g., language spoken in the larger community and as the medium of instruction) progressively influences the rate, path and balance of bilingual development. Existing bilingual–monolingual comparisons have shown that immigrant children lacking consistent access to L1 education at preschool usually show weaker L1 linguistic abilities than monolingual children from the country of origin in the L1. For example, 66-month-old Polish-L1 immigrant children in the UK had smaller expressive vocabularies than age-matched monolinguals in Poland (Mieszkowska et al., 2017). They also had lower productive vocabulary and grammar scores (aged 4–7, Haman et al., 2017). Notably, bilingual children with an earlier onset of bilingual exposure (EB) tend to perform worse in their native L1 than those with a later onset of bilingualism (LB). Armon-Lotem et al. (2021) compared English (L1)–Hebrew EB children (onset before 2;0) who attended Hebrew-speaking preschools and their LB peers (onset between 2;1–4;0) from similar preschools. EB children performed lower than LB children in English expressive vocabulary and sentence repetition. Similarly, in a study on US heritage Arabic–English preschoolers, Albirini (2018) found that, compared to LB children, EB children had lower accuracy rates in subject–verb agreement, plural morphology and relative clauses in Arabic. Conversely, simultaneous bilingual preschoolers (aged 3–5) with French as one of their L1s who attended French daycare in English-dominant Edmonton, Canada, performed comparably to monolingual peers in French receptive vocabulary (Smithson et al., 2014). Whether EB children with L1 support at preschool can catch up with LB peers in grammatical development, in addition to vocabulary development, remains an open question; the effects of onset age have been investigated much more intensively in the development of the L2 than in that of the L1 (e.g., Roesch & Chondrogianni, 2016; Unsworth et al., 2014).

Across studies, bilingual children who lack sustained L1 support at preschool consistently display early plateau, at least in vocabulary (e.g., Hmong–English bilinguals in the US in Kan & Kohnert, 2005; Chinese–English bilinguals in the US in Sheng, 2014; Sheng et al., 2011; Song et al., 2022). However, those with L1 support from preschools continue to develop morphosyntactic abilities and new vocabulary. For example, Rodina et al. (2020) found that bilingual children (aged 3;0–10;0) who received more instruction in Russian, their heritage language, performed better in grammatical gender than those with less Russian instruction. Armon-Lotem and Ohana (2017) found that among English-L1 children in Hebrew-speaking Israel, older children (36–45 months) understood significantly more English words than younger children (24–35 months), indexing growth in English receptive vocabulary from toddlerhood to preschool. Although English is not the majority language in Israel, it assumes a high status with strong institutional support in government and education. The continuous growth observed in L1 English and L1 Russian in these studies is likely to be extendable to the L1 Cantonese of Hong Kong Cantonese–English bilingual children. Our study aims to test this hypothesis.

3. Input–outcome relations in early bilingual development and Chinese–English bilinguals

In early bilingual development, the proportion of input in a language out of the child's total input (i.e., input proportion) strongly predicts skills in the same language (e.g., Pearson et al., 1997; Place & Hoff, 2011). Most research examines vocabulary development in bilingual toddlers or preschoolers, leaving aside grammatical outcomes (e.g., Cote & Bornstein, 2014; Dijkstra et al., 2016). An exception is Hoff et al. (2012), who found a strong link between input proportion and grammatical complexity (measured by MLU). Albeit valid and reliable, MLU and similar measures such as mean length of the three longest utterances in words (MLU3) are general, coarse-grained measures. The extent to which individual grammatical structures are affected differentially by input reduction within a group of bilingual toddlers is unclear.

https://doi.org/10.1017/S1366728925100412 Published online by Cambridge University Press

Importantly, Thordardottir (2011, 2015) discovered that English–French bilingual preschoolers receiving at least 70% input in either language scored within the 'monolingual normal range' (defined as monolingual mean ± 1 standard deviation) in expressive vocabulary in the same language and that those with close-to-equal exposure to both languages (40%–60%) were similar to corresponding monolinguals in MLU. However, French and English are closely related languages with many cognates and shared grammatical structures. Since children learning closely related language pairs need less input to attain monolingual levels than those learning distant language pairs (Blom et al., 2020), it remains to be tested whether the input thresholds reported for English–French children can also be extended to bilingual children learning distant languages like Cantonese and English.

Apart from input proportion, input quality also accounts for individual differences in language development outcomes (see reviews in Paradis, 2023; Rowe & Snow, 2020). For monolingual toddlers and preschoolers, lexical diversity (word type) and syntactic complexity in the input are robust predictors (e.g., Anderson et al., 2021; Hsu et al., 2017; Huttenlocher et al., 2010). Rowe (2012) found that the number of word types in input at 30 months significantly predicted vocabulary scores at 42 months, after controlling for input amount (number of words). In a longitudinal study (2–60 months), Vernon-Feagans et al. (2020) found that maternal MLU partially mediates the relation between maternal education and language skills in American children.

Despite the bulk of research on input–outcome relations in bilingual development, studies tend to examine input quality through indirect measures such as the number of input providers (e.g., Place & Hoff, 2016), the frequency of language and literacy activities like storytelling and reading (e.g., Jia & Paradis, 2015; Song et al., 2022), and media exposure (e.g., Sun et al., 2020). Few have directly examined the samples of child-directed input and correlated input quality with bilingual outcomes (but see Paradis & Navarro, 2003, for an examination of specific constructions in the input). An exception is David and Wei (2008), who found that young bilingual children whose parents produced longer sentences or richer vocabulary possessed a larger expressive vocabulary. However, whether fine-grained input features (e.g., lexical diversity and syntactic complexity) make additional contributions to bilingual development beyond input proportion has not been thoroughly investigated.

Research has shown that Chinese–English bilingual children do not necessarily perform weaker than monolinguals across all grammatical structures. Structures that are crosslinguistic translation equivalents and appear in similar syntactic positions in both languages might benefit from positive transfer, such that bilinguals demonstrate monolingual-like or even accelerated acquisition. This is exemplified by earlier and more productive usage of pronominal right-dislocation constructions (Ge et al., 2017) and the progressive aspect marker *-gan* in Cantonese–English bilingual children (Luk & Shirai, 2018). Note that the translation equivalents do not have to be identical across all linguistic levels to induce crosslinguistic mapping and transfer in bilingual children. Although progressive *-gan* in Cantonese and progressive *-ing* in English are similar, in that both express imperfective meanings and appear postverbally, they differ in many morphosyntactic and semantic aspects (e.g., telicity of co-occurring verbs). Crosslinguistic mappings between similar (rather than identical) structures also occur between the Mandarin perfective aspect marker *-le* and the English past tense marker *-ed* in Mandarin–English bilingual preschoolers (Nicoladis et al., 2020). Conversely, properties specific to Chinese appear to be particularly problematic for bilinguals. For instance, nominal classifiers and postverbal resultative/directional particles (also termed RESULTATIVE VERB COMPOUNDS) are widely attested in Chinese but not in English (Matthews & Yip, 2011). The inventory of such structures is significantly reduced in older Chinese–English bilingual children (Kan, 2019; Shang et al., 2024; Wei & Lee, 2001). How similar two grammatical structures need to be to induce crosslinguistic transfer is an unresolved question (Unsworth, 2023).

4. This study

4.1. Research questions and predictions

Robust and consistent evidence suggests that input reduction from bilingual exposure can result in early costs to native L1 development, particularly in vocabulary. How specific grammatical structures are differentially affected by input reduction among toddlers, however, awaits systematic investigation. Recent cross-sectional studies have begun to show that bilingual children with out-of-home L1 support continue gaining vocabulary and morphosyntax (e.g., Armon-Lotem & Ohana, 2017; Rodina et al., 2020) and might converge with the monolinguals (e.g., Smithson et al., 2014), at least in European language pairs acquired in Western contexts. Longitudinal studies tracking the same group of bilingual children developing their native L1 alongside a linguistically distant language in toddlerhood and the preschool age are scarce. Among bilingual children, while input proportion robustly predicts development, it is unclear how fine-grained qualitative aspects of caretaker input influence bilingual outcomes.

This study addresses multiple research gaps by investigating input and outcomes in the development of Cantonese grammar in Cantonese–English bilingual children who received substantial naturalistic input in both languages at home from infancy (earlier-onset bilingual, EB) at two critical time points (Time 1: end of toddlerhood before entering kindergarten at 3;0 and Time 2: approaching end of kindergarten at 5;8). As a comparison baseline, we included children who were raised in close-to-monolingual Cantonese households with little exposure to English by Time 1 and had developed bilingual proficiency to varying degrees through English exposure at kindergarten by Time 2. This group of later-onset bilingual (LB) children, rather than monolingual Cantonese children, was adopted as the baseline for this study because LB children represent the norm in Hong Kong – English is integral to kindergarten curricula (Curriculum Development Council, 2017) and strictly monolingual Cantonese kindergarteners do not exist. For ease of reference, this group is hereafter called the LB baseline (or simply 'the baseline'). We adopted the newly published Grammatical Analysis of Cantonese Samples (GACS; Wong et al., 2022) as a novel method for a comprehensive analysis of Cantonese grammatical structures. The following are our research questions and predictions:

1) **Early development of L1 Cantonese in toddlerhood**: At 3;0, to what extent did EB children perform lower than the baseline (LB children still at the monolingual stage) in Cantonese productive grammar?

Based on previous research, we predicted that the EB group would exhibit, as a group, lower-than-baseline Cantonese in general grammatical complexity. Given the absence of studies examining bilingual toddlers' L1 grammar in a comprehensive and systematic manner, and the intricate relations between input reduction and crosslinguistic transfer, we do not have precise predictions for all major grammatical structures in Cantonese. However, based on previous studies with older Chinese–English bilingual children, we

https://doi.org/10.1017/S1366728925100412 Published online by Cambridge University Press

predict that bilingual toddlers should perform comparably to the baseline in producing aspect markers and perform lower than the baseline in producing nominal classifiers and postverbal particles. The full set of grammatical structures will be introduced later.

2) **Continued development of L1 Cantonese at preschool age**: At 5;8, did the EB children still perform lower than the LB children in grammatical complexity, and in the structures in which they had lower-than-monolingual performance at 3;0 (if any)? Did the EB children show significant growth in grammatical complexity from 3;0 to 5;8?

Since the EB children in our study should have had substantial exposure to Cantonese in education and society at preschool age, we predict that they will show converging performance with the LB children and significant growth across time points.

3) **Input–outcome associations and individual differences:** How much input in Cantonese was needed by 3;0 for individual EB children to perform within the normal range of the baseline children (LB children at the monolingual stage) in grammatical complexity? To what extent can qualitative aspects of Cantonese input at 3;0 account for the variance in outcomes in Cantonese at 3;0 and 5;8, respectively, after controlling for Cantonese input proportion and other background variables?

Since Cantonese is more distant from English than French is from English, we predict that Cantonese–English EB children needed more than 40%–60% of the input in Cantonese to reach the normal range of the baseline at 3;0. We predict that caretaker input quality measures will account for a significant amount of variance in outcomes, in addition to input proportion.

4.2. Methods

4.2.1. Participants

This study was part of a larger longitudinal study investigating input–outcome relations in early multilingual development involving Chinese languages and English (Mai et al., 2025). The EB group selected for this study included 31 bilingual toddlers (16 girls and 15 boys). They received Cantonese input from their (grand)parents, who were native Cantonese speakers, and English input from parents or domestic helpers, who spoke English as a non-dominant language. The EB children's main English input providers had medium-to-high proficiency in English (self-ratings in Table 1). By 3;0, the children had

Table 1. Descriptive statistics of participants' background variables at 3;0 and 5;8

	Earlier-onset bilingual			Later-onset bilingual baseline		
Measures	*M*	*SD*	Range	*M*	*SD*	Range
Time 1 (3;0)		*n = 31*			*n = 21*	
Maternal education[a]	4.77	.88	3–6	4.62	1.08	2–7
Family income[b]	5.74	2.31	2–10	5.95	1.11	3–10
BRIEF-P GEC (raw score)[c]	104.77	12.3	82–130	102.05	16.01	76–143
Cumulative Cantonese input%[d]	.62	.16	.25–.83	.97	.03	.90–1.0
Cumulative English input%[d]	.37	.17	.13–.75	.03	.03	0–.10
Caretaker self-rated English level[e]	3.55	0.62	2–5	NA	NA	NA
Parent-rated Cantonese level[f]	4.42	.56	3–5	4.81	.40	4–5
Parent-rated English level[f]	4.19	.65	3–5	2.71	.72	1–4
Child MLUw in English[g]	2.02	.542	1.22–3.09	–	–	–
Time 2 (5;8)		*n = 14*			*n = 8*	
Maternal education[a]	4.79	1.05	3–6	5.12	.99	3–7
Family income[b]	5.71	2.23	2–10	7.38	2.07	3–10
Bug Search raw score	44.93	6.86	34–57	48.25	9.47	34–60
Cumulative Cantonese input% after time 1	.61	.13	.29–.75	.79	.13	.59–.96
Cumulative English input% after time 1	.34	.13	.21–.70	.19	.12	.03–.40
Parent-rated Cantonese level[f]	4.93	.27	4–5	4.94	.18	4.5–5
Parent-rated English level[f]	4.32	.50	3.5–5	3.88	.35	3–4
PPVT–5 standard score	90.30	10.16	73–108	76.12	11.81	60–93
EVT–3 standard score	94.07	12.11	74–116	78.88	10.20	70–102

[a]On a 7-point scale (1 = primary/elementary school and below; 7 = doctorate).
[b]On a 10-point scale (1 = below 9,000HKD; 10 = 121,001HKD & above).
[c]GEC = global executive composite.
[d]Formula:

$$\%\ \text{of caretaker input in language } X = \frac{\text{Sum of input hours in X by all main caretakers}}{\text{Total caretaker input hours} + \text{other input hours}}.$$

[e]On a 5-point scale (1 = cannot understand; 2 = very limited; 3 = conversational on everyday topics; 4 = proficient; 5 = (near-)native).
[f]On a 5-point scale (1 = cannot comprehend this language yet; 5 = can say complex sentences and respond fluently).
[g]Derived from 10-minute caretaker–child standard toy play.

https://doi.org/10.1017/S1366728925100412 Published online by Cambridge University Press

accumulated substantial input in Cantonese and English (mean English input proportion: 37%, ranging from 13% to 75%).

The LB baseline included 21 toddlers (10 girls and 11 boys). They were raised in close-to-monolingual Cantonese environments with no more than 10% input in English by 3;0 (mean English input proportion: 3%), but had regular English exposure at kindergarten after 3;0. All children were born full term without known or suspected language or neurobiological disorders. To control for the effects of birth order and sibling input, only firstborns were recruited. To ensure that the adult input samples collected in this study were representative of the children's main sources of input, children who had spent more than 20% of their time outside home (e.g., daycare, playgroup sessions, etc.) by age 3 were not included. This did not exclude many children due to the suspension of childcare services and social distancing during the COVID-19 pandemic. Parents and children of both groups completed a battery of tasks online through Zoom, assisted by the research team, during the pandemic (Time 1).

After the pandemic, all children received an invitation to a follow-up study at 5;8 (Time 2). A total of 22 children (14 EB, 9 girls; 8 LB baseline, 4 girls) returned and completed the tasks in our child-friendly laboratories. At 5;8, all children were attending kindergartens following the curriculum recommended by the Education Bureau of HKSAR, with a strong emphasis on oral and literacy skills in both Chinese and English. More details of the bilingual exposure and proficiency of the participants are reported in the next section.

4.2.2. Procedures and tasks

Caretaker and Input Questionnaire (CIQ). At 3;0, a parental questionnaire elicited information about family, child characteristics and language exposure. Parents identified the child's main caretakers from birth to 3 and reported the duration of care each caretaker provided (in months). Parents then reflected on daily routines of caretaker–child interaction for each main caretaker ('input hours', excluding sleeping and napping time, which involved limited verbal interaction) and the relative proportion(s) of language(s) used in caretaker–child interactions ('language proportion'). Based on the reported caretaking duration, input hours and language proportion, we calculated both the child's total input hours from all main caretakers ('total caretaker input hours') and the child's total input hours with each caretaker in each language ('caretaker input hours in X'). Parents also reported the child's language exposure beyond caretaker–child interactions ('other input hours', e.g., media, playgroup sessions) from birth to age 3. We computed the proportion of caretaker input in language X using the formula in (1) as follows:

$$\frac{\%\text{ of caretaker input}}{\text{in Language X}} = \frac{\text{Sum of input hours in X by all main caretakers}}{\text{Total caretaker input hours} + \text{other input hours}} \quad (1)$$

Caretakers who provided the highest proportion of input in Cantonese were identified as the 'main Cantonese input providers' (49 mothers, 3 fathers). At 5;8, another interview with the parents gathered information about the children's language exposure after Time 1 (from 3;0 to 5;8), including language and hours of interaction with individual caretakers and peers, extracurricular activities and media exposure.

Executive function and cognitive ability. At 3;0, executive function was assessed using the Behavior Rating Inventory of Executive Function – Preschool Version (BRIEF-P; Gioia et al., 2003); parents indicated 'problematic behaviours' of the children in the past six months. At 5;8, the children completed Bug Search, a non-verbal processing speed task from the Wechsler Preschool and Primary Scales of Intelligence – Fourth Edition (WPPSI-IV, Wechsler, 2012).

Grammatical development in Cantonese. At 3;0, the main Cantonese input provider filled in the long form of the Communicative Developmental Inventory (CDI) in Cantonese (Tardif & Fletcher, 2008) in a Zoom interview with the research team. Dale (1991) reported that CDI is highly correlated with the results of direct tests and with measures derived from spontaneous child speech samples. Although CDI is intended for monolingual children under 30 months, previous studies support its validity and utility for bilingual children up to 36 months (Mancilla-Martinez et al., 2011). CDI measures were excluded from group comparisons because the baseline group, who were monolinguals at 3;0, were expected to perform at ceiling. CDI measures were, however, used to address the final research question about individual differences within the EB group, as the bilinguals were not expected to perform at ceiling.

Additionally, caretaker–child dyads played cooking games for 10 minutes at 3;0, using prescribed identical kitchen toy sets delivered to their homes. The main input provider was instructed to play with the child 'in the way they usually do'. The interactions were video-recorded through Zoom by a research assistant who remained muted and invisible and audio-recorded by the caretaker using their smartphones at home. Details of the setup were described in Zhou et al. (2022). At 5;8, children told stories based on the Cat and Baby Goats pictures from the Multilingual Assessment Instrument for Narratives (MAIN; Gagarina et al., 2019) and answered 20 comprehension questions about the two stories. Narration was chosen as a quick and comprehensive measure of productive abilities in Cantonese, based on known associations between narrative skills and later language and academic achievements (e.g., Griffin et al., 2004).

The procedures of the study were approved by the research ethics committee of the Chinese University of Hong Kong. Written parental informed consent was obtained prior to data collection. The authors assert that all procedures contributing to this work comply with the ethical standards of relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975 (revised in 2008).

Table 1 presents children's demographic information at two time points. There were no significant differences in maternal education or family income between EB and baseline groups at either time point (*ps* > .05). At Time 1, the EB children had lower cumulative input in Cantonese than the baseline (EB: *M* = 62%, range: 25%–83%; baseline: *M* = 97%, range: 90%–100%). At Time 2, the EB group's cumulative input in Cantonese between Time 1 and Time 2 was similar to that in Time 1 (*M* = 61%, range: 29%–75%), whereas that of the baseline group dropped (*M* = 79%, range: 59%–96%) due to increased exposure to English at kindergarten, although it was still significantly higher than that of the EB children (*p* = .006). The two groups did not differ in the global executive composite of the BRIEF-P at 3;0 or in Bug Search raw score at 5;8 (*ps* > .05).

4.2.3. Transcription, coding and measures

Caretaker–child toy play at Time 1 and the children's narratives at Time 2 were manually transcribed in CHAT format (MacWhinney, 2000) by trained research assistants and revised for accuracy and consistency by the first author. Child utterances in caretaker–child interactions at 3;0 and narratives at 5;8 were coded based on a scheme adapted from the Grammatical Analysis of Cantonese Samples (GACS, Wong et al., 2022). GACS comprises a list of 62 prominent Cantonese grammatical structures, subsumed under five subscales (noun phrase; verb phrase; sentence structure; questions; and sentence adverbs, conjunction and sentence final

https://doi.org/10.1017/S1366728925100412 Published online by Cambridge University Press

particles). A child receives one point for a structure when a unique and felicitous token of the structure is recorded (i.e., type-based scoring), with a maximum of four points per structure. Wong (2022) showed that GACS positively correlates with MLUw and age in typically developing Cantonese children. Our adapted scheme includes specific descriptions of what counts as a unique token for each structure and representative examples from child utterances to ensure coding consistency and efficiency. The transcripts were coded by the first author, and 10% of the transcripts were second-coded by the corresponding author. Discrepancies were discussed, and relevant entries in the coding scheme were revised and improved. After several rounds of fine-tuning, the inter-coder reliability based on the current version of the coding scheme reached 83.8%. The full scheme is included in Appendix S1 and Table S1 in the supplementary materials.

Notice that in the original, larger study, the EB children also completed a parallel set of tasks in English in separate sessions at 3;0. As shown in the child MLUw in the caretaker–child interaction with English as the intended language (M = 2.02), EB children had developed productive English abilities by 3;0. At 5;8, standardized vocabulary tests in English (Peabody Picture Vocabulary Test – Fifth Edition, PPVT-5, Dunn, 2019; Expressive Vocabulary Test – Third Edition, EVT-3, Williams, 2018) were administered to both groups. The EB group outperformed the LB group in the standard scores of both tests (PPVT-5: M = 90.30 versus 76.12, $t(12.92) = 2.86$, $p = .014$; EVT-3: M = 94.07 versus 78.88, $w = 96$, $p = .007$). For this study, these English measures serve as indicators of the children's general English ability and of the degree of bilingualism at the two time points; they will not be further analysed in the results section. The children's development in English will be further investigated in another article.

Measures of Cantonese were derived from four data sources, summarized in (1). Different combinations of the measures were chosen to address each research question (see next section).

(1) Summary of tasks and derived measures
   i. Parental questionnaire: input proportion;
   ii. CDI form at 3;0: sentence complexity score;
   iii. Caretaker–child 10-min standard toy play at 3;0: word type, word token and MLUw in the input; MLUw and MLU3 of child utterances; GACS subscore (points of a specific GACS grammatical structure, max = 4); and GACS total score (sum of the subscores of the 62 GACS grammatical structures in individual children); and
   iv. Child narration task at 5;8: MLU3, total utterance (indicator of productivity), comprehension score (total number of correctly answered questions to index children's understanding of macro-structure, following Lindgren, 2019; Sheng et al., 2020, max. = 10 per story) and GACS subscore (a selective subset of GACS structures that showed EB versus baseline differences at 3;0; details below)

5. Results

5.1. Performance at 3;0 (Time 1)

General measures. Table 2 presents the descriptive statistics of the two general measures (MLUw and GACS total) at 3;0, among others. The EB group displayed lower means than the baseline in both measures (MLUw: M = 2.72 versus 3.13, GACS total: M = 45.65 versus 62.57). To examine group effects, we conducted multiple regressions in R (version 4.3.0; R Core Team, 2023), with MLUw and GACS as dependent variables. To control for cognitive development and socioeconomic status (SES), BRIEF-P GEC scores and maternal education were entered together with group as independent variables, and the baseline as the reference. Maternal education was chosen over family income to index SES due to its wider acceptance in similar studies (e.g., Lauro et al., 2020; Unsworth et al., 2019). Spearman's correlation revealed a moderate correlation between maternal education level and family income ($r = .51$, $p < .001$). Given that a lower BRIEF-P score indicates better performance, the BRIEF-P GEC scores were reversed by multiplying −1 before entering the models. Two children (one per group) did not provide BRIEF-P measures; the sample size for this regression analysis was reduced to 50. To correct for multiple comparisons, we controlled for false discovery rate (FDR) through the Benjamini–Hochberg procedure. Results revealed a significant group effect in MLUw ($B = -0.41$, SE = 0.14, $\beta = -.39$, $t = -2.94$, $p = .033$) and GACS ($B = -15.87$, SE = 5.17, $\beta = -.41$, $t = -3.07$, $p = .033$) at Time 1 after controlling for maternal education and cognitive measures, indicating that the EB group displayed reduced grammatical complexity in their Cantonese utterances compared to the baseline.

Table 2. Descriptive statistics of group performance at 3;0 and 5;8

	Earlier-onset bilingual			Later-onset bilingual baseline		
Measures	*M*	*SD*	Range	*M*	*SD*	Range
Time 1(3;0)		*n = 31*			*n = 21*	
Toyplay_MLUw	2.72	0.53	1.53–3.46	3.17	0.40	2.36–3.90
Toyplay_MLU3	6.53	1.60	2.67–9.67	7.81	1.55	5.67–11.33
Toyplay_GACS score	45.65	19.69	1–81	62.57	13.13	41–82
CDI_sentence complexity	75.16	7.92	48–83	76.52	5.79	60–84
Time 2 (5;8)		*n = 14*			*n = 8*	
Narration_MLU3	13.35	1.38	10.33–15.50	12.96	2.02	10.83–16.83
Narration_Total utterances	14.46	4.46	7.5–23	16.06	6.60	9–29.5
Narration_Comprehension	17.29	2.09	14–20	16.13	2.64	11–19

Note: ToyPlay = caretaker–child 10-minute standard toy play with Cantonese as the intended language; narration = storytelling in Cantonese (Cat and Baby Goats stories in MAIN). Narrative measures at Time 2 averaged across two stories.

https://doi.org/10.1017/S1366728925100412 Published online by Cambridge University Press

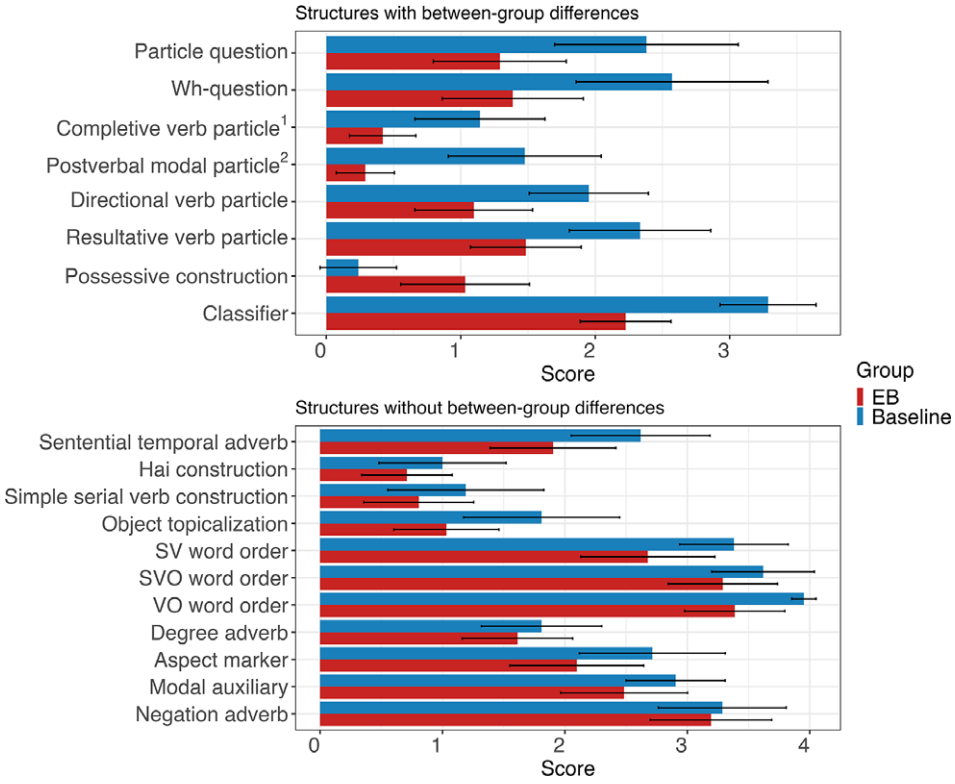

Figure 1. Subscores of 19 grammatical structures of the earlier-onset bilingual (EB) children and the baseline children (later-onset bilingual children at monolingual stage) at 3;0. ([1,2] originally termed 'other verbal particle' and 'potential particle' in the Grammatical Analysis of Cantonese Samples, Wong et al., 2022).

GACS subscores. Mean scores of the 62 GACS structures in both groups of children are presented in Table S2 in the supplementary materials. Forty-three of the structures showed either ceiling or floor effects (mean scores of both groups below 1 or above 3.5 out of 4). These structures were not informative for examining group differences and were excluded in the between-group comparison of grammatical structures.

Figure 1 shows the means of the remaining 19 structures for both groups. A Shapiro–Wilk test showed that the assumption of multivariate normality was violated. Following Choi and Marden (1997), we conducted a MANOVA on the ranked data using the cmanova() function in R (Wilcox & Schönbrodt, 2014) on our full sample (n = 52), with GACS item scores as dependent variables. Results revealed a significant group effect ($H(19) = 80.16$, $p < .001$). Follow-up Wilcoxon signed-rank tests revealed significant group differences in eight structures: classifiers (W = 134, p = .0002, r = −.51), possessive constructions (W = 448.5, p = .008, r = −.37), directional particles (W = 184, p = .007, r = −.38), postverbal modal particles (W = 131.5, p <. 000, r = −.56), resultative particles (W = 193.5, p = .01, r = −.35), completive particles (i.e., adversative *can1*, quantifying *saai3* and *maai3*, W = 193, p = .007, r = −.38), particle questions in the form of [declarative clause-particle] (W = 193, p = .01, r = −.35) and *wh*-questions (W = 191, p = .01, r = −.36). All tests survived FDR control (ps < .05). EB children performed higher than the baseline in possessive constructions and lower than the baseline in seven other constructions. No significant group effects were found in the other 11 structures, namely aspect markers, degree adverbs, modal auxiliaries, negation, simple serial verb constructions, *hai*-constructions (prepositional phrases headed by *hai*), object topicalization, word orders (SVO, SV and VO) and sentential temporal adverbs.

5.2. Performance at 5;8 (Time 2) and growth from 3;0 to 5;8

General measures. Table 2 reports descriptive statistics of performance measures at Time 2, among others. Each child told two stories in Cantonese (Cat and Baby Goats); all measures at Time 2 were averaged across the two stories within each subject. Group means were very close across measures, including MLU3 (EB: 13.35, LB: 12.96), total utterance (EB: 14.46, LB: 16.06) and comprehension score (EB: 17.29, LB: 16.12). Multiple regressions revealed no group effects across these general measures (MLU3: B = 0.48, SE = 0.78, β = .15, t = 0.62, p = .54; total utterance: B = −1.31, SE = 2.52, β= − .12, t = −0.52, p = .61; comprehension score: B = 1.52, SE = 1.03, β= .32, t = 1.48, p = .16), suggesting the EB children had caught up with LB peers by 5;8.

GACS subscores. We coded the children's utterances in narratives for the subset of GACS structures found to be weaker in the EB children at Time 1. Seven structures were analysed for this purpose: classifiers; postverbal resultative, directional, modal and completive (*can1, saai3* and *maai4*) particles; particle questions; and *wh*-questions (see Supplementary Table S2 for mean scores of each structure). We found that both groups showed floor effects (mean scores lower than 1) in four structures (postverbal modal and completive particles, particle questions and *wh*-questions), likely due to a lack of felicitous contexts in the short narration tasks; these four structures were excluded from subsequent analyses. Like the general measures at Time 2, the group means were numerically very close for noun classifiers (EB: 2.71, LB: 3.06), postverbal resultative particles (EB: 1.32, LB: 1.62) and postverbal directional particles (EB: 3.00, LB: 3.06). We conducted a MANOVA on the ranked data through cmanova() in R (Wilcox & Schönbrodt, 2014) with a sample of 22, with GACS item scores as dependent variables. Results revealed no significant group effects ($H(3) = 4.486$, p = .214).

https://doi.org/10.1017/S1366728925100412 Published online by Cambridge University Press

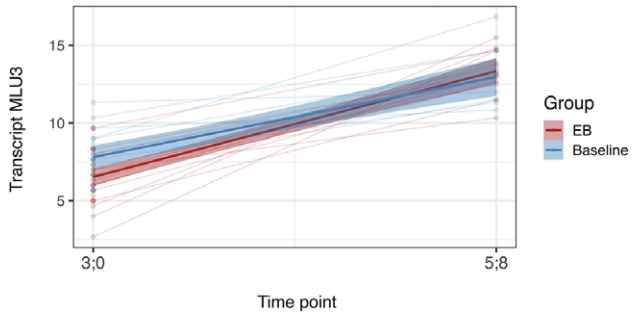

Figure 2. Growth in mean length of the three longest utterances in words (MLU3) from 3;0 to 5;8 in the earlier- and later-onset bilingual (EB and baseline) children.

Growth between 3;0 and 5;8. As shown in Figure 2, both groups showed increased MLU3 at 5;8 (EB: 13.35; LB: 12.96), compared with 3;0 (EB: 6.53; LB: 7.81). Using the *nlme* package in R (Pinheiro et al., 2023), we ran a linear mixed-effect model with time point as the independent variable and by-subject random intercepts and random slope if it improved the model fit as indicated by the Akaike Information Criterion. Results showed significant time point effects in MLU3 ($B = 6.82$, SE = 0.50, $\beta = .9$, $t = 13.77$, $p < .001$), suggesting EB children had significant gains in syntactic complexity across time points.

5.3. Input at 3;0 as predictors of outcomes in the earlier-onset bilinguals (EB)

Input measures. We calculated Cantonese input proportion from parental questionnaires and derived three input measures (word token, word type and MLUw) from transcripts of 10-min caretaker–child toy play. On average, caretakers of the EB children produced 1023.55 words and 199.48 unique words in Cantonese during the sessions, with a mean MLUw of 4.65.

Input proportion at 3;0 ~ outcomes at 3;0 and 5;8. Zero-order correlations were run between Cantonese input proportion at 3;0 and transcript-derived performance measures. Input proportion at 3;0 was significantly correlated with two grammatical complexity measures at 3;0 (CDI sentence complexity and GACS scores, *rs* = .534 and .387, *ps* < .05) but not with any of the outcome measures at 5;8 (see Table 3). Following Thordardottir (2011), we then transformed the EB group's performance measures into z-scores using the means and standard deviations of the baseline group, allowing us to examine how EB children with varying Cantonese input proportions performed relative to the normal range (defined with reference to the baseline group). Note that CDI sentence complexity was excluded from this analysis due to expected ceiling effects in the baseline group at 3;0; only GACS scores were transformed. A linear fit was performed for GACS scores, with Cantonese input proportion as the predictor. The model returned a significant result (GACS: $R^2 = .15$, $p = .03$). Figure 3 plots the linear fit for GACS as a function of Cantonese input proportion. Visual examination of the plot shows that the proportion of Cantonese input needed to reach the lower bound of the baseline's normal range (1 SD below baseline mean) was approximately 70%.

Input quality at 3;0 ~ outcomes at 3;0. CDI sentence complexity score, rather than transcript-based measures, was chosen to index child outcomes at 3;0 in this correlation analysis, because properties in child utterances are expected to correlate with the same properties in child-directed speech (CDS) in the same task due to alignment between interlocutors. Examining associations between transcript-based input measures and independent CDI-based outcome measures bypasses this methodological limitation. Since this part of the analysis involved only EB children, we did not expect ceiling effects in CDI measures as we did for LB children (see below). Zero-order correlations between input quality measures (word type, word token and MLUw) and CDI sentence complexity score at 3;0 are presented in Table 3. Both word type and MLUw in the input correlated with CDI sentence complexity score ($r = .399$ and .375, respectively, *ps* < .05), but word token in the input did not correlate with the outcome. Based on these correlations, we performed hierarchical linear regressions to determine the extent to which input properties at 3;0 accounted for individual variation in grammatical complexity among EB children at 3;0. In step 1, we built a base model with BRIEF-P, maternal education (SES) and input proportion as control variables. In step 2, MLUw and word type in the input were entered into the model one at a time. As shown in Table 4, the base model accounted for 40% of the variance, and MLUw and word type in the input accounted for an additional

Table 3. Zero-order correlations between input properties at 3;0 and acquisition outcomes in Cantonese at 3;0 and 5;8 in the earlier-onset bilingual (EB) children

	Input at Time 1 (3;0)			
Outcomes	Proportion	Word type	Word token	MLUw
Time 1 (n = 30)				
CDI_sentence complexity	.534*	.399*	.195	.376*
ToyPlay_MLUw	.351	–	–	–
ToyPlay_GACS score	.387*	–	–	–
Time 2 (n = 14)				
Narration_MLUw	.300	.085	.112	−.023
Narration_Total Utterance	.128	.296	.144	.011
Narration_Comprehension	−.272	−.150	−.101	.181

Note: ToyPlay = caretaker–child 10-minute standard toy play with Cantonese as the intended language; narration = storytelling in Cantonese (Cat and Baby Goats stories); *$p < .05$.

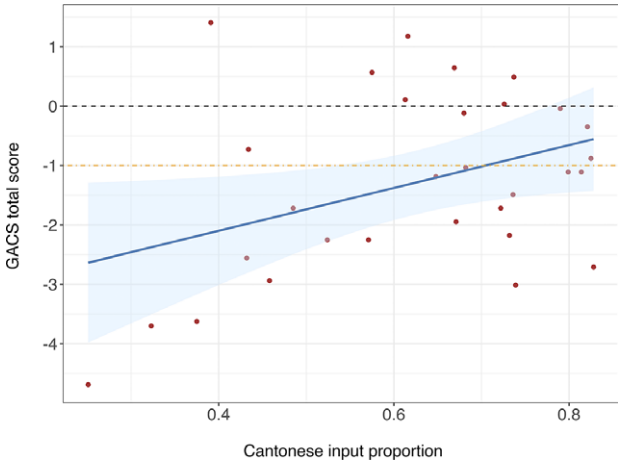

Figure 3. Distribution of GACS total score in Cantonese against input proportion of Cantonese in earlier-onset bilingual (EB) children at 3;0. Scores are standardized using the mean of the baseline children (later-onset bilingual (LB) at the monolingual stage) at 3;0. The grey and orange dashed lines indicate the mean and lower bound of the baseline normal range, respectively.

https://doi.org/10.1017/S1366728925100412 Published online by Cambridge University Press

Table 4. Hierarchical regression with input quality measures as predictors and child grammatical complexity at 3;0 as the response variable within the earlier-onset bilingual (EB) children (n = 30)

	CDI sentence complexity					
Predictors	Estimates	SE	*β*	*p*	$R^2_{adjusted}$	ΔR^2
Step 1					.40*	
Intercept	43.14	13.83		.004		
BRIEF-P GEC	−0.07	0.09	−.10	.49		
Maternal education	0.97	1.25	.11	.44		
Cantonese input%	33.62	7.13	.71	< .000		
Step 2A					.55*	.15*
Intercept	22.69	13.74		.11		
BRIEF-P GEC	−0.01	0.08	−.02	.87		
Maternal education	0.96	1.09	.11	.39		
Cantonese input%	33.00	6.20	.70	<.000		
Input MLUw	5.70	1.86	.39	.005		
Step 2B					.49*	.08*
Intercept	27.97	14.48		.06		
BRIEF-P GEC	−0.96	0.09	−.15	.29		
Maternal education	0.93	1.16	.11	.43		
Cantonese input%	31.63	6.68	.67	< .000		
Input word type	0.07	.03	.31	.03		

Note: *p < .05.

15% and 8% variance of CDI sentence complexity, respectively, after input proportion, SES and cognition.

Input at 3;0 ~ outcomes at 5;8. Zero-order correlations revealed no significant correlations between input measures at 3;0 and narration outcomes at 5;8, suggesting little direct influence of input in toddlerhood on Cantonese outcomes at preschool age for the EB children (see Table 3).[1]

6. Discussion

6.1. Summary of findings

This study investigated the grammatical development of Cantonese as an L1/societal language in children exposed to Cantonese and English before age 3 (EB group). Through dual language input provided by caretakers at home, these children developed productive skills in both languages before age 3. Using parental questionnaire, CDI report, caretaker–child standard toy play, elicited narrations and standardized assessments, we measured bilingual children's performance at 3;0 and 5;8, with reference to a baseline group of LB children who had been raised monolingually in Cantonese until 3;0 and were exposed to English after 3;0. We adopted stringent inclusion criteria, recruiting EB and baseline firstborns matched on demographic factors known to be influential (e.g., SES and executive function at 3;0) and on the type of bilingual kindergarten programme between 3;0 and 5;8.

We first asked whether EB three-year-olds lagged behind the baseline in productive Cantonese grammar at 3;0. Note that the baseline children were monolingual at this point. Our regression analyses found that the EB children, as a group, exhibited reduced grammatical complexity in their utterances (measured by MLUw) and smaller inventories of grammatical structures (operationalized as the total score of 62 grammatical structures based on the adapted GACS scheme) than the baseline. Among the 19 specific grammatical structures that did not display ceiling or floor effects, the EB children showed lower type frequency (lower subscore) in seven structures, including classifiers, four sets of postverbal verb particles, *wh*-questions and particle questions. Meanwhile, EB children performed similarly to baseline children in many other structures (e.g., basic word orders, degree adverbs). Surprisingly, EB children produced possessive constructions more frequently than the baseline. We will discuss these structures in the next section.

Next, we examined continued L1 development in the EB children between 3;0 and 5;8. We hypothesized that EB children would benefit from social and educational L1 support and catch up with LB children by 5;8. A linear mixed-effect model showed that the EB children who returned at Time 2 (n = 14) exhibited considerable growth in grammatical complexity across time points, with MLU3 doubling from 6.53 to 13.35. Multiple regressions showed that the two groups were comparable across all production and comprehension measures (MLU3, total utterance and comprehension score) derived from the MAIN narration task at 5;8. This suggests that EB children's growth in Cantonese grammar was steeper than that of LB children between 3;0 and 5;8.

Finally, we asked how much Cantonese input is needed for EB children to reach the normal range of the baseline group (defined as baseline group mean minus one standard deviation). Figure 3 shows

[1]Zero-order correlations were run between proportion of Cantonese input between 3;0 and 5;8 and acquisition outcomes in Cantonese at 5;8 in the EB children. No significant correlation was found (see Supplementary Table S3).

https://doi.org/10.1017/S1366728925100412 Published online by Cambridge University Press

that the meeting point is around 70% for GACS total score at 3;0, which is higher than the input proportion threshold for monolingual-like performance in morphosyntax identified by Thordardottir (2015, 40%–60%), possibly due to Cantonese and English being linguistically more distant than English and French. We also asked how input quality contributes to outcomes in addition to input proportion. Our regression analyses showed that at 3;0, grammatical complexity (MLUw) and lexical diversity (word type) in the input, rather than caretaker talkativeness (word token), significantly predicted children's grammatical complexity (CDI sentence complexity scores), accounting for an additional 15% and 8% variance, respectively, after controlling for input proportion, SES and cognitive abilities. However, none of the input measures at 3;0 correlated with outcomes at 5;8.

6.2. The costs to the L1 and the 'costly' grammatical structures

Measured by monolingual standards, developing the L1 alongside other language(s) comes with direct 'costs' to the L1, since time spent learning additional language(s) diminishes time spent learning the L1. Those costs were found in the grammatical development of L1 Cantonese in the Cantonese–English bilingual 3-year-olds in our study. Although there have been few prior studies of Cantonese–English toddlers in a bilingual society, our finding comes as no surprise, given previous studies with bilingual toddlers with other language pairs (e.g., Cote & Bornstein, 2014; Miękisz et al., 2019) and the universality of input reduction effects – at least proportionally – in early bilingualism. Our study replicated previous findings with a new language pair.

Nonetheless, input reduction costs to the L1 are not equally extendable to all grammatical structures. By adopting a fine-grained approach, we identified the structures most susceptible to input reduction and thus most 'costly' in early bilingual development. Previous studies have suggested that structural overlaps or shared features between languages enable crosslinguistic transfer, thereby mitigating input reduction effects. The underlying psycholinguistic mechanism may well be that similar structures across languages are paired up (or grouped together) and co-activated in processing through between-language priming, increasing the combined frequency of the similar structures and enhancing learning (e.g., Baroncini & Torregrossa, 2025; Unsworth, 2023). Structures that are unique to Cantonese, however, are less likely to be consistently mapped onto similar English structures and co-activated when processing English. They benefit less from crosslinguistic co-activation, and their acquisition is more reliant on their Cantonese experience. When the amount of Cantonese input is reduced to the extent that it provides insufficient cues for meaningful statistical learning (whether internally guided or not), lower-than-monolingual performance occurs.

Yet crosslinguistic similarity is best conceptualized as a continuum rather than as a binary. How similar must two grammatical structures be to induce crosslinguistic transfer in toddlers? This is a non-trivial question, also raised by Unsworth (2023). Many structures can be straightforwardly positioned at either end of the similarity continuum. For example, both Cantonese and English are SVO languages instantiating SVO, SV and VO sequences, and both have a range of verbal or adverbial elements that frequently appear preverbally to select or modify the lexical verb (e.g., modal auxiliaries, negators, degree and temporal adverbs). The EB children showed monolingual-like scores in producing such 'similar' structures (Figure 1). At the other end of the continuum, nominal classifiers, in-situ *wh*-questions and particle questions ([Declarative-SFP]) are three undoubtedly 'dissimilar' structures, so EB children showed below-monolingual type frequency across these structures unique to Cantonese. We propose that the contrasting degrees of crosslinguistic similarity of these structures and their correspondingly contrasting acquisition outcomes in the bilinguals in our study indicate cause-and-effect relations between the two.

Structural similarity as a causal factor in bilingual outcomes can be disentangled from frequency effects. Theoretically, structures that are similar crosslinguistically tend to be less marked and are more frequent in the language and hence in input, which may impact children's order of acquisition of individual structures. (We thank an anonymous reviewer for raising this point.) However, we expect frequency effects in Cantonese input to impact both the EB and LB children, as both groups received Cantonese input from native Cantonese speakers and were matched in demographic and cognitive factors (Table 1). Note that the 'affected' structures were identified based on EB versus LB differences after screening out ceiling and floor structures, not on the relative ranking of structural scores within the EB children. The differential vulnerability of the structures in the EB children should thus be attributed to linguistic properties uniquely accessible to EB children, but not to monolinguals. The most plausible factor is therefore crosslinguistic similarity, rather than structural frequency in Cantonese.

What about structures that lie further from the poles of the continuum? Possessive constructions, for example, are realized through [possessor-*ge3/classifier*-possessee] in Cantonese and can be loosely mapped onto either or both of possessive's and *of*-structures in English. Our findings show that the EB 3-year-olds were not weaker than the baseline in the possessive constructions, a pattern also found in Babatsouli and Nicoladis (2019), who report a Greek–English child having similar accuracy in English possessives as age-matched monolinguals. The OSV order (object topicalization) is another case of debatable crosslinguistic similarity. OSV is frequent in Cantonese and yet highly restricted in spoken English (e.g., *this pot, you use*). Our EB children were not significantly less productive in OSV than the monolingual children. This differs from previous findings that heritage Mandarin children in English-dominant societies (aged 5–9) performed lower than monolinguals in the OSV order (aged 5–9, Hao & Chondrogianni, 2023), likely because our EB children had qualitatively and quantitatively higher input in their native L1 than the heritage children. Another set of structures with debatable crosslinguistic similarity are the four types of postverbal particles ([V1V2], sometimes termed RESULTATIVE VERB COMPOUNDS or RVC). They encode rich temporal, spatial and modality meanings (e.g., completive, directional, resultative, potential), which are expressed by a variety of linguistic units ranging from words and phrases to clauses in English (see crosslinguistic comparison and bilingual acquisition findings in Yuan & Zhao, 2011; Shang et al., 2024). Our EB bilinguals performed weaker than monolinguals across these postverbal particles, consistent with previous findings with older bilingual children in English-dominant societies (Shang et al., 2024).

Overall, crosslinguistic similarity accurately predicts bilingual–monolingual differences in grammatical structures that are straightforwardly 'similar' or 'dissimilar'. Our findings suggest that bilingual children draw upon crosslinguistic overlaps in a bootstrapping manner to form a shared grammatical representation, mitigating input reduction. Structural similarity is thus an important factor modulating relative vulnerability among L1 structures in early bilingualism. Yet for structures with less straightforward crosslinguistic correspondences, the extent to which translation

https://doi.org/10.1017/S1366728925100412 Published online by Cambridge University Press

counterparts in the other language contribute to explaining outcomes remains unclear. As the first study exploring differential vulnerability in a range of grammatical structures within one group of bilingual toddlers, the degree of crosslinguistic similarity was not included as a pre-determined predictor. However, our findings have identified 19 Cantonese grammatical structures from 62 structures for future investigation. In Supplementary Table S4, we present a post-hoc analysis of crosslinguistic correspondences across the 19 structures, including their typical forms in our speech samples and potential candidates for English counterparts, to provide testable hypotheses for further research on crosslinguistic correspondence in bilingual toddlers.

6.3. Individual differences and the roles of amount and quality of early input

The proportion of Cantonese input among total language input varied across the EB children in our study, ranging from 25% to 83% (Table 1). Although the EB children, as a group, were outperformed by the baseline children at Time 1, they exhibited vast individual differences across measures, which were larger than those among the baseline group (see standard deviations in Table 2). Investigating these individual differences by controlling for SES and cognition, we showed that costs to L1 development are not inevitable at the individual level. Our analysis revealed that bilinguals with 70% input in Cantonese were able to reach the baseline in grammatical complexity and that input quality (utterance length and lexical diversity) explained 15% and 8% additional variance, respectively, in grammatical development after input proportion. However, caution is warranted in generalizing this finding to other bilingual populations. The 70% threshold was identified in the L1 of bilingual toddlers growing up in a L1-majority context; it may not apply to L1-minority contexts, namely immigrant situations. Nevertheless, for communities and societies pursuing early additive bilingualism, our findings highlight the value of investing in home input packages offering 70% of input in the native L1 to reach that goal. Where this ideal amount cannot be achieved, coaching caretakers to provide high-quality input can mitigate the costs of input reduction to L1 development.

The influence of home input, however, decreases upon the onset of formal schooling. None of the performance measures of the EB children at 5;8 was correlated with home input quality or amount at 3;0. It is logical that, as children spend more time at kindergarten, the impact of caretaker input during infancy and early childhood becomes outweighed by school exposure. Yet sustained educational support in native L1 might not be accessible to heritage children, who typically experience a substantial decrease in L1 input upon schooling. In such cases, home input, which remains their main L1 input source, might still exert a strong influence on L1 outcomes (e.g., Mai et al., 2022).

6.4. L1 acquisition from a longer-term perspective and methodological contributions

Developing and maintaining a native L1 is a lifelong endeavour, with numerous opportunities for learners to acquire, automatize and, in some cases, 'lose' L1 knowledge. The growing literature on immigrant children acquiring heritage L1 has demonstrated that the course and attainment of L1 development in these populations are characterized by large variation, moderated by child-internal and child-external factors (e.g., Cote & Bornstein, 2014; Hoff et al., 2012; Paradis, 2023). Whether and to what extent home and institutional L1 support is provided continuously at preschool and school ages (e.g., Armon-Lotem & Ohana, 2017; Gathercole, 2002; Mai et al., 2022) is a critical external factor for L1 development. In our study, although the EB children performed lower than the baseline children in their native L1 at Time 1 ('early costs'), they demonstrated continued (and likely faster) growth afterwards, catching up with the baseline group across macro- and micro-measures of L1 narrative skills in Cantonese while demonstrating superior performance in English at Time 2 ('long-term gains'). As such, whether and to what extent L1 development suffers 'costs' from infant and toddler bilingualism largely depends on the developmental time point selected for comparisons.

A number of caveats should be noted when interpreting the longer-term gains in both languages in our EB children at 5;8. Although results showed continuous growth and convergence with the baseline children at the group level, this pattern might not be applicable to each individual EB child due to wide variations in bilingual development (see Figure 3). Moreover, our EB children had at least 25% Cantonese input during toddlerhood. Convergence with baseline proficiency might not extend to bilinguals with lower proportions of L1 input. Additionally, the EB children's advantage in English at 5;8 may be explained by the early English input they received from relatively proficient L2 English speakers, but whether and for how long such advantages in English will last after 5;8 depends on many input-related and child-internal factors after 5;8.

Our findings reveal a distinctive velocity of L1 development in additive bilinguals raised in a bilingual society at the early stages. Our inclusion of a baseline group that was monolingual at Time 1 and bilingual at Time 2 distinguishes our study from many others. Admittedly, this design prevents us from generalizing bilingual–monolingual differences at Time 2. Nevertheless, our comparison is appropriate and ecologically valid, since acquiring Cantonese in a bilingual setting is the norm, rather than a rarity, among Cantonese-learning children both in Hong Kong and in many overseas Chinese communities. Bilingual acquisition of Mandarin is also the norm for many children in mainland China. Clinicians and practitioners may consider the different developmental patterns of EB and LB children when making recommendations to parents and screening for language disorders. Specifically, the resilient and vulnerable grammatical structures identified in our analysis provide an important empirical base for further identification of clinical markers of language disorders in bilingual children.

Our study is observational and correlational by nature. However, thanks to the longitudinal design, we were able to compare the same group of EB children with the same group of baseline children across two time points. We can thus attribute the rapid growth of L1 Cantonese after 3;0 in EB children to the rich and diverse exposure to both spoken and written Cantonese in kindergarten, which closed the gap between EB and LB Cantonese development by 5;8. Formal education has been shown to promote grammatical development of the native L1 effectively in school-age heritage bilinguals (e.g., Bayram et al., 2017; Gathercole, 2002; Rodina et al., 2020); recently, Montrul and Armstrong (2024) proposed that textual exposure in formal education promotes L1 growth in a minoritized context ('Literacy Enhancement Hypothesis'). Although our study did not test this hypothesis directly, our results are consistent with it, suggesting that it may apply to bilingual children of a much younger age in an L1-majority context.

Due to the sudden outbreak of the pandemic, we could not assess the children directly at Time 1 as originally planned, constituting a methodological limitation. However, this prompted

https://doi.org/10.1017/S1366728925100412 Published online by Cambridge University Press

us to develop reliable and responsible web-based remote data collection protocols. When following a standard protocol, the 10-minute caretaker–child standard play sessions recorded through Zoom in fact offer greater scheduling flexibility and document interactions in a physical environment familiar to the dyads. We recommend this new data collection format for future studies facing physical constraints.

7. Conclusions

Our study has its novelty in examining the grammatical development of the native L1 at two critical time points of early bilingual development through a combination of direct and indirect measures, with a baseline group matched for important demographic and cognitive variables. Methodologically, we innovated a web-based pipeline for recording caretaker–child play sessions, and we analysed a comprehensive set of grammatical structures from child speech samples. Our findings revealed both 'costs' of early bilingualism to the grammatical development of the native L1 around kindergarten entry and 'gains' in both languages 32 months later. By showing the significant yet different roles of input proportion and input quality in the L1 development of bilingual toddlers and the differential vulnerability of specific grammatical structures, our study sheds new light on the intricate and dynamic relations between input and outcomes in language acquisition and provides an empirical basis for future clinical applications and educational interventions. Our study has two important limitations: a relatively small sample size and a lack of direct in-person assessment at Time 1 to examine bilingual versus monolingual differences at the comprehension level. Overall, our study suggests additive bilingualism in Cantonese–English bilingual children in L1-majority contexts, in contrast to the subtractive bilingualism typically observed in L1-minority settings. Our study calls for a longer-term perspective on L1 grammatical development when, for many populations, early additive bilingualism is both the goal and the norm.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S1366728925100412.

Data availability statement. The full coding scheme of the speech samples is described in Appendix S1 and Table S1 in the supplementary materials. Transcripts of the speech samples are included in the Early Additive Child Multilingual Corpus in the Child Language Data Exchange System (CHILDES, doi:10.21415/JEFY-9F39). The data that support the findings of this study are available in OSF at https://osf.io/ujqyt/.

Acknowledgements. This study is derived from the doctoral thesis work of Yuqing Liang, supervised by Ziyin Mai. The larger study was funded by research grants from the Research Grants Council of HKSAR, awarded to Ziyin Mai (GRF/RGC #14615820, ECS/RGC #21604522). We express our deep gratitude to our colleagues for their tremendous support since the beginning of the project: Virginia Yip, Stephen Matthews, Zhuang Wu, Peggy Mok, Cecilia Chan, Jiangling Zhou, Jingyao Liu, Xuening Zhang and Mengyao Shang. The following research assistants participated in data collection and transcription in the larger project: Shiyu He, Lu Zou, Lu Zhao, Lean Luo, Ashley Chan, Katherine Chang, Yingyu Su, Haoxuan Hou, Jiaqi Nie, Yue Cao, Yue Chen, Ranee Cheng, Qiuyun Cai, Mingwei Liang and Elena Vermeer. Some of the findings were presented at the Seventh International Conference on Chinese as a Second Language Research (CASLAR-7) in Beijing, China, and the Child Language Symposium 2024 (CLS2024) in Newcastle, UK. We also thank the editors and reviewers of Bilingualism: Language and Cognition, who provided insightful comments, and Elena Vermeer, who proofread the manuscript. All remaining errors are our own.

Competing interests. The author(s) declare none.

References

Anderson, N. J., **Graham, S. A.**, **Prime, H.**, **Jenkins, J. M.**, & **Madigan, S.** (2021). Linking quality and quantity of parental linguistic input to child language skills: A meta-analysis. *Child Development*, **92**(2), 484–501. https://doi.org/10.1111/cdev.13508.

Albirini, A. (2018). The role of age of exposure to English in the development of Arabic as a heritage language in the United States. *Language Acquisition*, **25**(2), 178–196. https://doi.org/10.1080/10489223.2016.1234620.

Armon-Lotem, S., & **Ohana, O.** (2017). A CDI study of bilingual English-Hebrew children–frequency of exposure as a major source of variation. *International Journal of Bilingual Education and Bilingualism*, **20**(2), 201–217. https://doi.org/10.1080/13670050.2016.1179257.

Armon-Lotem, S., **Rose, K.**, & **Altman, C.** (2021). The development of English as a heritage language: The role of chronological age and age of onset of bilingualism. *First Language*, **41**(1), 67–89. https://doi.org/10.1177/0142723720929810.

Babatsouli, E., & **Nicoladis, E.** (2019). The acquisition of English possessives by a bilingual child: Do input and usage frequency matter? *Journal of Child Language*, **46**(1), 170–183. https://doi.org/10.1017/S0305000918000429.

Baroncini, I., & **Torregrossa, J.** (2025). Language and structure activation explain cross-linguistic influence in bilingual language production: Evidence from within-and across-language priming. *Bilingualism: Language and Cognition*, 1–15. https://doi.org/10.1017/S1366728925000379.

Bayram, F., **Rothman, J.**, **Iverson, M.**, **Kupisch, T.**, **Miller, D.**, **Puig-Mayenco, E.**, & **Westergaard, M.** (2017). Differences in use without deficiencies in competence: Passives in the Turkish and German of Turkish heritage speakers in Germany. *International Journal of Bilingual Education and Bilingualism*, **22**(8), 919–939. https://doi.org/10.1080/13670050.2017.1324403.

Blom, E. (2010). Effects of input on the early grammatical development of bilingual children. *International Journal of Bilingualism*, **14**(4), 422–446. https://doi.org/10.1177/1367006910370917.

Blom, E., **Boerma, T.**, **Bosma, E.**, **Cornips, L.**, **van den Heuij, K.**, & **Timmermeister, M.** (2020). Cross-language distance influences receptive vocabulary outcomes of bilingual children. *First Language*, **40**(2), 151–171. https://doi.org/10.1177/0142723719892794.

Byers-Heinlein, K., **Gonzalez-Barrero, A. M.**, **Schott, E.**, & **Killam, H.** (2023). Sometimes larger, sometimes smaller: Measuring vocabulary in monolingual and bilingual infants and toddlers. *First Language*, **44**(1), 74–95. https://doi.org/10.1177/01427237231204167.

Census & Statistics Department of the Government of Hong Kong Special Administrative Region. (2022). *Hong Kong monthly digest of statistics.* https://www.censtatd.gov.hk/en/data/stat_report/product/FA100270/att/B72402FA2024XXXXB0100.pdf#page=12.10

Choi, K., & **Marden, J.** (1997). An approach to multivariate rank tests in multivariate analysis of variance. *Journal of the American Statistical Association*, **92**(440), 1581–1590. https://doi.org/10.1080/01621459.1997.10473680.

Cote, L. R., & **Bornstein, M. H.** (2014). Productive vocabulary among three groups of bilingual American children: Comparison and prediction. *First Language*, **34**(6), 467–485. https://doi.org/10.1177/0142723714560178.

Curriculum Development Council. (2017). *Kindergarten education curriculum guide: Joyful learning through play balanced development all the way*. Education Bureau HKSARG. https://www.edb.gov.hk/attachment/en/curriculum-development/major-level-of-edu/preprimary/ENG_KGECG_2017.pdf

Dale, P. S. (1991). The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of Speech, Language, and Hearing Research*, **34**(3), 565–571. https://doi.org/10.1044/jshr.3403.565.

David, A., & **Wei, L.** (2008). Individual differences in the lexical development of French–English bilingual children. *International Journal of Bilingual Education and Bilingualism*, **11**(5), 598–618. https://doi.org/10.2167/beb478.0.

Dijkstra, J., **Kuiken, F.**, **Jorna, R.**, & **Klinkenberg, E.** (2016). The role of majority and minority language input in the early development of a bilingual vocabulary. *Bilingualism: Language and Cognition*, **19**(1), 191–205. https://doi.org/10.1017/S1366728915000012.

Dunn, D. M. (2019). *Peabody picture vocabulary test* (5th ed.). Pearson.

Foursha-Stevenson, C., **Nicoladis, E.**, **Trombley, R.**, **Hablado, K.**, **Phung, D.**, & **Dallaire, K.** (2023). Pronoun comprehension and cross-linguistic influence in monolingual and bilingual children. *International Journal of Bilingualism*, **28**(3), 479–493. https://doi.org/10.1177/13670069231170968.

Gagarina, N., **Klop, D.**, **Kunnari, S.**, **Tantele, K.**, **Välimaa, T.**, **Bohnacker, U.**, & **Walters, J.** (2019). *MAIN: Multilingual assessment instrument for narratives – Revised* (Vol. **63**). ZAS Papers in Linguistics.

Gathercole, V. (2002). Grammatical gender in bilingual and monolingual children: A Spanish Morphosyntactic distinction. In K. Oller & R. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 207–219). Multilingual Matters. https://doi.org/10.21832/9781853595721-010.

Ge, H., **Matthews, S.**, **Cheung, L. Y.**, & **Yip, V.** (2017). Bidirectional cross-linguistic influence in Cantonese–English bilingual children: The case of right-dislocation. *First Language*, **37**(3), 231–251. https://doi.org/10.1177/0142723716687955.

Gioia, G. A., **Espy, K. A.**, & **Isquith, P. K.** (2003). *The behavior rating inventory of executive function-preschool version (BRIEF-P)*. Psychological Assessment Resources.

Griffin, T. M., **Hemphill, L.**, **Camp, L.**, & **Wolf, D. P.** (2004). Oral discourse in the preschool years and later literacy skills. *First Language*, **24**(2), 123–147. https://doi.org/10.1177/0142723704042369.

Haman, E., **Wodniecka, Z.**, **Marecka, M.**, **Szewczyk, J.**, **Białecka-Pikul, M.**, **Otwinowska, A.**, **Mieszkowska, K.**, **Łuniewska, M.**, **Kołak, J.**, **Miękisz, A.**, **Kacprzak, A.**, **Banasik, N.**, & **Foryś-Nogala, M.** (2017). How does L1 and L2 exposure impact L1 performance in bilingual children? Evidence from polish-English migrants to the United Kingdom. *Frontiers in Psychology*, **8**, 1444. https://doi.org/10.3389/fpsyg.2017.01444.

Hao, J., & **Chondrogianni, V.** (2023). Comprehension and production of non-canonical word orders in mandarin-speaking child heritage speakers. *Linguistic Approaches to Bilingualism*, **13**(4), 468–499. https://doi.org/10.1075/lab.20096.hao.

Hoff, E., **Core, C.**, **Place, S.**, **Rumiche, R.**, **Señor, M.**, & **Parra, M.** (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, **39**(1), 1–27. https://doi.org/10.1017/S0305000910000759.

Hsu, N., **Hadley, P.**, & **Rispoli, M.** (2017). Diversity matters: Parent input predicts toddler verb production. *Journal of Child Language*, **44**(1), 63–86. https://doi.org/10.1017/S0305000915000690.

Huttenlocher, J., **Waterfall, H.**, **Vasilyeva, M.**, **Vevea, J.**, & **Hedges, L. V.** (2010). Sources of variability in children's language growth. *Cognitive Psychology*, **61**(4), 343–365. https://doi.org/10.1016/j.cogpsych.2010.08.002.

Jia, R., & **Paradis, J.** (2015). The use of referring expressions in narratives by mandarin heritage language children and the role of language environment factors in predicting individual differences. *Bilingualism: Language and Cognition*, **18**(4), 737–752. https://doi.org/10.1017/S1366728914000728.

Kan, P. F., & **Kohnert, K.** (2005). Preschoolers learning Hmong and English: Lexical-semantic skills in L1 and L2. *Journal of Speech, Language, and Hearing Research*, **48**, 372–383. https://doi.org/10.1044/1092-4388(2005/026.

Kan, R. T. (2019). Production of Cantonese classifiers in young heritage speakers and majority language speakers. *International Journal of Bilingualism*, **23**(6), 1531–1548. https://doi.org/10.1177/13670069188080.

Lauro, J., **Core, C.**, & **Hoff, E.** (2020). Explaining individual differences in trajectories of simultaneous bilingual development: Contributions of child and environmental factors. *Child Development*, **91**(6), 2063–2082. https://doi.org/10.1111/cdev.13409.

Lindgren, J. (2019). Comprehension and production of narrative macrostructure in Swedish: A longitudinal study from age 4 to 7. *First Language*, **39**(4), 412–432. https://doi.org/10.1177/0142723719844089.

Luk, Z., & **Shirai, Y.** (2018). The development of aspectual marking in Cantonese-English bilingual children. *International Review of Applied Linguistics in Language Teaching*, **56**(2), 137–179. https://doi.org/10.1515/iral-2014-0018.

MacWhinney, B. (2000). *The CHILDES project: Tools for Analyzing talk* (3rd ed.). Lawrence Erlbaum Associates.

Mai, Z., **Liang, Y.**, **Wu, Z.**, & **Yip, V.** (2025) Trilingual toddlers strategically draw on closely-related languages to catch up. In *Paper presented at the 15th international symposium on bilingualism (ISB15)*, June 9–13. Donostia-San Sebastián, Spain.

Mai, Z., **Zhao, L.**, & **Yip, V.** (2022). The mandarin *Ba*-construction in school-age heritage speakers and their parental input. *Linguistic Approaches to Bilingualism*, **12**(3), 377–405. https://doi.org/10.1075/lab.18025.mai.

Mancilla-Martinez, J., **Pan, B. A.**, & **Vagh, S. B.** (2011). Assessing the productive vocabulary of Spanish–English bilingual toddlers from low-income families. *Applied PsychoLinguistics*, **32**(2), 333–357. https://doi.org/10.1017/S0142716410000433.

Matthews, S., & **Yip, V.** (2011). *Cantonese: A comprehensive grammar* (2nd ed.). Routledge. https://doi.org/10.4324/9780203835012.

Miękisz, A., **Haman, E.**, **Łuniewska, M.**, **Kuś, K.**, **O'Toole, C.**, & **Katsos, N.** (2019). The impact of a first-generation immigrant environment on the heritage language: Productive vocabularies of Polish toddlers living in the UK and Ireland. In *Risk and protective environmental factors for early bilingual language acquisition* (pp. 67–84). Routledge.

Mieszkowska, K., **Łuniewska, M.**, **Kołak, J.**, **Kacprzak, A.**, **Wodniecka, Z.**, & **Haman, E.** (2017). Home language will not take care of itself: Vocabulary knowledge in trilingual children in the United Kingdom. *Frontiers in Psychology*, **8**, 1358. https://doi.org/10.3389/fpsyg.2017.01358.

Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge University Press.

Montrul, S., & **Armstrong, A.** (2024). The literacy enhancement hypothesis in bilingual language development. In E. Babatsouli (Ed.), *Multilingual acquisition and learning: An ecosystemic view to diversity* (pp. 60–89). John Benjamins. https://doi.org/10.1075/sibil.67.02mon.

Nicoladis, E., **Yang, Y.**, & **Jiang, Z.** (2020). Why "jumped" is so difficult: Tense/aspect marking in mandarin-English bilingual children. *Journal of Child Language*, **47**(5), 1073–1083.

Paradis, J. (2023). Sources of individual differences in the dual language development of heritage bilinguals. *Journal of Child Language*, **50**(4), 793–817. http://doi.org/10.1017/S0305000922000708

Paradis, J., & **Navarro, S.** (2003). Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: What is the role of the input? *Journal of Child Language*, **30**(2), 371–393. https://doi.org/10.1017/S0305000903005609.

Pearson, B. Z., **Fernández, S. C.**, **Lewedeg, V.**, & **Oller, D. K.** (1997). The relation of input factors to lexical learning by bilingual infants. *Applied PsychoLinguistics*, **18**(1), 41–58. https://doi.org/10.1017/S0142716400009863

Pinheiro J, **Bates D**, **R Core Team** (2023). NLME: Linear and nonlinear mixed effects models. R package version 3.1-163. Available at: https://CRAN.R-project.org/package=nlme

Place, S., & **Hoff, E.** (2011). Properties of dual language exposure that influence 2-year-olds' bilingual proficiency. *Child Development*, **82**(6), 1834–1849. https://doi.org/10.1111/j.1467-8624.2011.01660.x.

Place, S., & **Hoff, E.** (2016). Effects and noneffects of input in bilingual environments on dual language skills in 2 ½-year-olds. *Bilingualism: Language and Cognition*, **19**(5), 1023–1041. http://doi.org/10.1017/S1366728915000322.

R Core Team (2023). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. Available at: https://www.R-project.org/.

Rodina, Y., **Kupisch, T.**, **Meir, N.**, **Mitrofanova, N.**, **Urek, O.**, & **Westergaard, M.** (2020). Internal and external factors in heritage language acquisition: Evidence from heritage Russian in Israel, Germany, Norway, Latvia and the United Kingdom. *Frontiers in Education*, **5**(20). https://doi.org/10.3389/feduc.2020.00020.

Roesch, A. D., & **Chondrogianni, V.** (2016). "Which mouse kissed the frog?" Effects of age of onset, length of exposure, and knowledge of case marking on the comprehension of wh-questions in German-speaking simultaneous and early sequential bilingual children. *Journal of Child Language*, **43**(3), 635–661. https://doi.org/10.1017/S0305000916000015.

Rowe, M. L., & **Snow, C. E.** (2020). Analyzing input quality along three dimensions: Interactive, linguistic, and conceptual. *Journal of Child Language*, **47**(1), 5–21. https://doi.org/10.1017/S0305000919000655.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, **83**(5), 1762–1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x.

Shang, M., **Zhao, L.**, **Yip, V.**, & **Mai, Z.** (2024). Protracted development in the heritage lexicon: Resultative verb compounds in school-age heritage Chinese children and their input. *Linguistic Approaches to Bilingualism*, **15**(3), 342–369. https://doi.org/10.1075/lab.23001.sha.

Sheng, L. (2014). Lexical-semantic skills in bilingual children who are becoming English-dominant: A longitudinal study. *Bilingualism: Language and Cognition*, **17**(3), 556–571. https://doi.org/10.1017/S1366728913000746.

https://doi.org/10.1017/S1366728925100412 Published online by Cambridge University Press

Sheng, L. I., **Lu, Y.**, & **Kan, P. F.** (2011). Lexical development in Mandarin–English bilingual children. *Bilingualism: Language and Cognition*, **14**(4), 579–587. https://doi.org/10.1017/S1366728910000647.

Sheng, L., **Shi, H.**, **Wang, D.**, **Hao, Y.**, & **Zheng, L.** (2020). Narrative production in Mandarin-speaking children: Effects of language ability and elicitation method. *Journal of Speech, Language, and Hearing Research*, **63**(3), 774–792. https://doi.org/10.1044/2019_JSLHR-19-00087.

Smithson, L., **Paradis, J.**, & **Nicoladis, E.** (2014). Bilingualism and receptive vocabulary achievement: Could sociocultural context make a difference? *Bilingualism: Language and Cognition*, **17**(4), 810–821. https://doi.org/10.1017/S1366728913000813.

Song, L., **Sheng, L.**, & **Luo, R.** (2022). Comprehension skills of Chinese-English dual language learners: Relations across languages and associations with language richness at home. *International Journal of Bilingual Education and Bilingualism*, **27**(1), 19–37. https://doi.org/10.1080/13670050.2022.2137386.

Sun, H., **S, N. G.**, **O'brien, B.**, & **Fritzsche, T.** (2020). Child, family, and school factors in bilingual preschoolers' vocabulary development in heritage languages. *Journal of Child Language*, **47**(4), 817–843. https://doi.org/10.1017/S0305000919000904.

Tardif, T., & **Fletcher, P.** (2008). *Chinese communicative development inventories: User's guide and manual*. Peking University Medical Press.

Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism*, **15**(4), 426–445. https://doi.org/10.1177/1367006911403202.

Thordardottir, E. (2014). The typical development of simultaneous bilinguals. In T. Grüter & J. Paradis (Eds.), *Input and experience in bilingual development* (pp. 141–160). John Benjamins. https://doi.org/10.1075/tilar.13.08tho.

Thordardottir, E. (2015). The relationship between bilingual exposure and morphosyntactic development. *International Journal of Speech-Language Pathology*, **17**(2), 97–114. https://doi.org/10.3109/17549507.2014.923509.

Unsworth, S. (2014). Comparing the role of input in bilingual acquisition across domains. In T. Grüter, & J. Paradis (Eds.), *Input and experience in bilingual development* (pp. 181–201). John Benjamins. https://doi.org/10.1075/tilar.13.10uns

Unsworth, S., **Argyri, F.**, **Cornips, L.**, **Hulk, A.**, **Sorace, A.**, & **Tsimpli, I.** (2014). The role of age of onset and input in early child bilingualism in Greek and Dutch. *Applied PsychoLinguistics*, **35**(4), 765–805. https://doi.org/10.1017/S0142716412000574.

Unsworth, S., **Brouwer, S.**, **De Bree, E.**, & **Verhagen, J.** (2019). Predicting bilingual preschoolers' patterns of language development: Degree of non-native input matters. *Applied PsychoLinguistics*, **40**(5), 1189–1219. https://doi.org/10.1017/S0142716419000225.

Unsworth, S. (2023). Shared syntax and cross-linguistic influence in bilingual children: Evidence from between-and within-language priming. *Linguistic Approaches to Bilingualism*, **15**, 117–151. https://doi.org/10.1075/lab.22093.uns.

Vernon-Feagans, L., **Bratsch-Hines, M.**, **Reynolds, E.**, & **Willoughby, M.** (2020). How early maternal language input varies by race and education and predicts later child language. *Child Development*, **91**(4), 1098–1115. https://doi.org/10.1111/cdev.13281.

Wechsler, D. (2012). *Wechsler preschool and primary scale of intelligence* (4th ed.) The Psychological Corporation.

Wei, L., & **Lee, S.** (2001). L1 development in an L2 environment: The use of Cantonese classifiers and quantifiers by young British-born Chinese in Tyneside. *International Journal of Bilingual Education and Bilingualism*, **4**(6), 359–382. https://doi.org/10.1080/13670050108667738.

Wilcox, R.R., & **Schönbrodt, F.D.** (2014). The WRS package for robust statistics in R (version 0.24). Retrieved from: http://r-forge.r-project.org/projects/wrs/

Williams, K. T. (2018). *Expressive vocabulary test–third edition*. Pearson Assessments.

Wong, A. M.-Y. (2022). The development of grammar in Cantonese-speaking preschool children. In A. M.-Y. Wong (Ed.), *Understanding development and disorder in Cantonese using language sample analysis* (pp. 65–102). Routledge. https://doi.org/10.4324/9780367824013.

Wong, A. M.-Y., **Cheung, C. C.-H.**, **Lo, J. M.-W.**, & **Wan, E. K.-H.** (2022). Grammatical analysis of Cantonese samples. In A. M.-Y. Wong (Ed.), *Understanding development and disorder in Cantonese using language sample analysis* (pp. 19–64). Routledge. https://doi.org/10.4324/9780367824013.

Yip, V., & **Matthews, S.** (2007). Relative clauses in Cantonese-English bilingual children: Typological challenges and processing motivations. *Studies in Second Language Acquisition*, **29**(2), 277–300. https://doi.org/10.1017/S0272263107070143.

Yuan, B., & **Zhao, Y.** (2011). Asymmetric syntactic and thematic reconfigurations in English speakers' L2 Chinese resultative compound constructions. *International Journal of Bilingualism*, **15**(1), 38–55. https://doi.org/10.1177/1367006910371023.

Zhou, J., **Mai, Z.**, **Cai, Q.**, **Liang, Y.**, & **Yip, V.** (2022). Reference production in Mandarin–English bilingual preschoolers: Linguistic, input, and cognitive factors. *Frontiers in Psychology*, **13**, 897031. https://doi.org/10.3389/fpsyg.2022.897031.