




Heritage language development and processing: Non-canonical word orders in Mandarin–English child heritage speakers

Jiuzhou Hao^{1,2} , Vasiliki Chondrogianni¹ and Patrick Sturt¹

¹School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, United Kingdom and ²Department of Language and Culture (ISK), UiT The Arctic University of Norway, Tromsø, Norway

Research Article

Cite this article: Hao, J., Chondrogianni, V., & Sturt, P. (2024). Heritage language development and processing: Non-canonical word orders in Mandarin–English child heritage speakers. *Bilingualism: Language and Cognition*, **27**, 334–349. <https://doi.org/10.1017/S1366728923000639>

Received: 24 December 2022
Revised: 8 August 2023
Accepted: 18 August 2023
First published online: 18 September 2023

Keywords:

Mandarin; child heritage speakers; non-canonical word order; syntax; processing

Corresponding author:

Jiuzhou Hao,
Email: jiuzhou.hao@uit.no

Abstract

Previous research suggests that child HSSs' performance in offline linguistic tasks is typically worse than their age-matched monolingual peers and is modulated by linguistic and child-level factors. This study examined the comprehension and production of three Mandarin non-canonical structures in 5- to 9-year-old Mandarin–English heritage children and Mandarin-speaking monolingual children, including an online processing task. Results showed that heritage children had different performance in production and offline comprehension across structures compared to monolinguals. In online processing, they showed sensitivity to different cues similarly to monolinguals but took longer to revise initial misinterpretations. Within heritage children, we found that presence of morphosyntactic cues facilitated performance across tasks while cross-linguistic influence was only identified in production and offline comprehension but not in online processing. Additionally, input quantity predicted their production and offline comprehension accuracy of non-canonical structures, whereas age modulated their production. Lastly, online processing was not modulated by age nor input.

Introduction

Heritage speakers (HSs) are early bilinguals who acquire their heritage (minority) language (HL) from birth and the societal dominant language either at the same time as their first language or later via immersion. Although HSs are considered native speakers of the HL and are exposed to (near-)native input to the HL from birth, HSs of different languages have been shown to differ from their age-matched monolingual peers in their HL development and use across linguistic domains, especially when tested on offline linguistic tasks (Chondrogianni & Schwartz, 2020; Daskalaki et al., 2019; Hao & Chondrogianni, 2021; Montrul & Polinsky, 2021). Additionally, HSs may have highly varied linguistic experience with both the HL and the societal dominant language (Rothman et al., 2022).

Developmentally, structures with non-canonical word order¹, e.g., passives, have been reported to be especially problematic relative to the canonical word order, e.g., actives, for child HSs to produce and comprehend (e.g., Chondrogianni & Schwartz, 2020; Hao & Chondrogianni, 2021). These non-canonical structures involve the displacement of constituents from the original position, where sentential arguments are interpreted (Chomsky, 1993). For example, in English passives, as in “The girl was kissed by the boy”, *the girl* is the patient of the verb *kiss* that takes up the position occupied by a grammatical subject in an active sentence, whereas the agent *the boy* is within a by-phrase. Recent studies suggest that both linguistic (related to the specific linguistic property or related to language) and individual difference (ID) factors, including both child-internal and child external factors (following the terminology used in Paradis, 2023) modulate child HSs' acquisition of non-canonical structures. For example, linguistically, modulating factors may include the presence or absence or the transparency of the morphosyntactic cues within the structures (Chondrogianni & Schwartz, 2020; Hao & Chondrogianni, 2021), cross-linguistic influence (CLI) from the societal dominant language (Mai et al., 2018), and the relative positions of agent and patient in the structure (Hao & Chondrogianni, 2021). For ID factors, chronological age (a child-internal factor; Hao & Chondrogianni, 2021) and HL input quantity (a [proximal] child-external factor; Daskalaki et al., 2019; Flores et al., 2017) have been shown to affect HL development although empirical results are mixed.

A recent study by Hao and Chondrogianni (2021) examined how these linguistic and ID factors modulated Mandarin–English² child HSs' production and offline comprehension of

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.



three Mandarin non-canonical structures with differing word order and the presence or absence of morphosyntactic cues, i.e., *BA*-, *BEI*-, and *OSV*-constructions. They found that child HSs performed worse relative to age-matched monolingual children across structures and tasks. Importantly, the presence or absence of morphosyntactic cues modulated child HSs' performance which showed signs of CLI from English. On the other hand, chronological age but not current language use at home (an index for input quantity) predicted their performance. Input quantity has been argued to be a trigger of HSs' performing differently from monolingual baselines (Polinsky & Scontras, 2020). As such, the observed differences between heritage and monolingual children coupled with the lack of an effect of HL input in the study by Hao and Chondrogianni (2021) merits further investigation. Specifically, the reason why no HL input effects were found in the study by Hao and Chondrogianni (2021) may be due to the fact that the child HSs in their sample were primarily first-generation migrant children with varying age of immigration. For these children, HL input may have been generally homogeneous, not giving rise to enough variability that would trigger HL input effects. Therefore, it is important to test a group of heritage children, such as that of second-generation immigrants, who may be characterised with a larger degree of variability in the amount of input they receive to ascertain the role of this particular child-external factor along with the role of chronological age.

Furthermore, the majority of existing studies on HL development has used production or offline comprehension to ascertain how child HSs produce or comprehend HL structures. How child HSs process non-canonical structures in real-time remains to be understood. To the best of our knowledge, there is no study examining child HSs' online processing of non-canonical structures, and the current study aims to fill this gap. This is important because existing studies on adult HSs, although limited in number and targeted at other linguistic properties, show that they may adopt monolingual-like online processing strategies while displaying worse offline comprehension accuracy (see Jegerski, 2018a, 2018b; Keating et al., 2016, among others). For example, although Spanish adult HSs did not show sensitivity to the marking of direct objects in Differential Object Marking constructions in HL offline (Jegerski, 2018b; Montrul & Sánchez-Walker, 2013), their performance in a Visual World Eye-tracking experiment suggested that they did use the relevant marker in real-time during sentence comprehension (Jegerski & Sekerina, 2021). Such a discrepancy between online and offline comprehension in HSs found in the literature calls for more research, including both online and offline comprehension measurements in the same HS population to ascertain their linguistic abilities in a more detailed way, especially when online processing pressure has been proposed to trigger the differences between heritage and monolingual speakers (Polinsky & Scontras, 2020). In addition, online measures are important because they afford more direct access to how language processing unfolds in real time – automatic language processing that is not subject to potential introspection and to (the same degree of) affective factors as in offline measures, which measure how languages are comprehended after linguistic information is completely available.

In the present study, we expand on Hao and Chondrogianni (2021) to examine how Mandarin–English child HSs process Mandarin non-canonical structures in real-time, and to illustrate the effect of different linguistic and child-level factors in modulating HL development and processing with a more homogenous

group of child HSs. Specifically, we test the production and offline as well as online comprehension of three Mandarin non-canonical structures with differing word order and the presence or absence of morphosyntactic cues, i.e., *BA*-, *BEI*-, and *OSV*-constructions, in five- to nine-year-old second-generation Mandarin–English child HSs.

Modulating factors for HL development

The presence, absence or the transparency of morphosyntactic cues in syntactic structure(s) has been argued to modulate HSs' development of non-canonical structures (Chondrogianni & Schwartz, 2020; Hao & Chondrogianni, 2021; Janssen, 2016; Kim et al., 2018). Hao and Chondrogianni (2021) showed that Mandarin–English child HSs performed better on the Mandarin structure with a morphosyntactic cue (i.e., *BEI*-constructions) than on the one without (i.e., *OSV*-constructions), in both offline comprehension and production. Furthermore, Kim et al. (2018) found that Korean–English child HSs' offline comprehension of Korean non-canonical structures was positively modulated by their productive abilities for the use of case markers. On the other hand, Janssen (2016) illustrated that Russian–English child HSs performed worse than Polish–English child HSs in comprehending Russian and Polish non-canonical structures respectively, which was argued to be caused by the fact that morphosyntactic cues in Russian are less complex and more transparent than those in Polish.

Another linguistic factor, more specifically language-related factor, is the influence from the societal dominant language to the HL. In contrast to monolingual speakers who deal with only one language on a daily basis, HSs' constant use of the societal dominant language may have an effect on their HL development and use, an effect commonly coined as cross-linguistic influence (CLI). CLI has been shown to take place when HSs prefer the HL structure that is shared between the HL and the ML, and especially when the HL allows more than one structure (Müller & Hulk, 2001; and see Chondrogianni, 2023; van Dijk et al., 2022 for a discussion of if structural overlap is necessary for CLI). In the context of HSs' production of Chinese non-canonical structures, CLI has been reported in the form of HSs' avoidance of non-canonical structures while opting for the structure that overlaps between the two languages (Hao & Chondrogianni, 2021; Mai et al., 2018; Polinsky et al., 2010). For example, Mandarin HSs opt for canonical SVO structures in Mandarin instead of agentive *BA*-constructions. In comprehension, Mandarin/Cantonese–English child and adult HSs' preference in interpreting non-canonical (non-agent-first) structures as canonical (agent-first) in comprehension (Hao & Chondrogianni, 2021; Kidd et al., 2015) has also been interpreted as an effect of CLI due to influence from English, an agent-first SVO language. Importantly, the canonical structure in Mandarin/Cantonese is the shared structure between Mandarin/Cantonese and English, i.e., *SVO*-constructions. Additionally, Hao and Chondrogianni (2021) further suggested that CLI may interact with another linguistic factor, i.e., word order. Specifically, relative to the shared structure between the HL and the societal dominant language, i.e., *SVO*-constructions, structures requiring thematic role reversal, i.e., *BEI*- and *OSV*-constructions, were more prone to CLI than structures having the same thematic role ordering, i.e., *BA*-constructions.

Turning to individual difference (ID) factors, e.g., chronological age, input quantity, among others, findings have been mixed on if and how these factors separately or jointly modulate

HL development (see Paradis, 2023 and commentaries for a summary of how different ID factors are operationalised and their role in HL development). For example, although some studies have identified positive correlations between chronological age and HL performance in an age range similar to the current study (Flores et al., 2017; Hao & Chondrogianni, 2021), others have found no improvement in the HL as a function of age (Chondrogianni & Schwartz, 2020; Janssen, 2016). Similarly, input quantity, often measured as current home language use, was found to modulate HL performance in some studies (e.g., Daskalaki et al., 2019; Janssen, 2016), but not in some other studies (e.g., Hao & Chondrogianni, 2021). Importantly, when the two factors were examined within the same population, Hao and Chondrogianni (2021) showed that age modulated HL development over and above input quantity, but the reversed pattern was observed by Janssen (2016). These discrepancies among studies might, however, be because of the heterogeneity that different studies had in their sampled populations and the fact that input quantity may survive as a predicting factor depending on the linguistic (sub-)domain(s) under investigation (and how it is quantified). For example, Daskalaki et al. (2019) showed that Greek child HSs' accuracy on subject realisation, but not post-verbal subject pronoun placement, was predicted by input quantity.

This study extends this line of research and examines the role of the presence or absence of morphosyntactic cues, CLI, word order, chronological age, and input quantity in Mandarin-English child HSs' production and comprehension (online and offline) of Mandarin *BA*-, *BEI*- and *OSV*-constructions. Importantly, the heritage children we focused on have a relatively homogeneous age of exposure to the societal dominant language (English) and generation status, but with varied chronological age (five to nine) and input quantity.

Mandarin non-canonical structures and their development

Word order in Mandarin and English

Like English, the canonical word order in Mandarin is Subject³-Verb-Object (SVO), as in (1). Essentially, the canonical word order (SVO-constructions henceforth) has a Noun-Verb-Noun (NVN) combination at the phrasal level.

- (1) Zhangsan ti-le Lisi yi-xia.
Zhangsan kick-asp Lisi one-cl
'Zhangsan kicked Lisi once'

On the other hand, the three non-canonical structures tested in the study, i.e., the *BA*-construction (2), the *BEI*-construction (3), and the *OSV*-construction (4), have the second Noun Phrase (NP2) in the corresponding SVO-construction moved from its post-verbal position to a position before the verb. As a result, the three non-canonical structures are all NNV in their phrasal combination.

- (2) Zhangsan ba Lisi ti-le yi-xia.
Zhangsan ba Lisi kick-asp one-cl
'Zhangsan kicked Lisi once.'
- (3) Lisi bei Zhangsan ti-le yi-xia.
Lisi bei Zhangsan kick-asp one-cl
'Lisi was kicked by Zhangsan once'
- (4) Lisi, Zhangsan ti-le yi-xia.
Lisi, Zhangsan kick-asp one-cl
'Lisi was kicked by Zhangsan once'

In *BA*-constructions, the moved NP2 still follows the subject and is separated from the subject by the morphosyntactic cue *ba*. Therefore, *BA*-constructions share the subject-object/agent-patient ordering with *SVO*-constructions and have no equivalent in English. Contrarily, in *BEI*- and *OSV*-constructions, the moved NP2 precedes the subject. As a result, *BEI*- and *OSV*-constructions share the object-subject/patient-agent ordering. *BEI*-constructions differ from *OSV*-constructions in that the two NPs in *BEI*-constructions are separated by the morphosyntactic cue *bei*, while *OSV*-constructions do not require a morphosyntactic cue in-between the two NPs. In studies of Mandarin syntax, *BEI*-constructions are considered passives (C.-T. J. Huang et al., 2009). Although English does have passives which are also non-canonical in word order, it still has an NV(-by N) phrasal combination. *OSV*-constructions, on the other hand, are sometimes analysed as object topicalisation which is also available in English. However, a pause after the NP1 in naturalistic production of *OSV*-constructions is not necessary although we separated the two NPs with a comma in example (4). This contrasts with the English object topicalization, whose frequency is also extremely rare, especially in children's input (Slabakova, 2015).

The development and processing of Mandarin non-canonical structures

Studies have shown that monolingual Mandarin-speaking children naturalistically produce *BA*- and *BEI*-constructions as early as two years of age (Deng et al., 2018). In terms of processing, monolingual Mandarin-speaking children process *BA*- and *BEI*-constructions in a qualitatively adult-like way in real-time from the age of three years (Y. T. Huang et al., 2013; Zhou & Ma, 2018) and are indistinguishable from adults in offline comprehension and production of *BA*-, *BEI*- and *OSV*-constructions from the age of five years (Hao & Chondrogianni, 2021). Nonetheless, mixed results have been observed concerning potentially differential performance across the different non-canonical structures. Specifically, Y. T. Huang et al. (2013) reported better offline comprehension performance of *BA*-constructions relative to *BEI*-constructions, whereas Deng et al. (2018) observed earlier naturalistic production of *BEI*-constructions relative to *BA*-constructions. Interestingly, Deng et al. (2018) found that the production of *BEI*-constructions developed relatively earlier than that of *BA*-constructions, even though input frequency of *BA*-constructions (2.62%) was significantly higher than *BEI*-constructions (0.13%), suggesting the development of these structures might not be modulated by input frequency at least in the monolingual context. No differential performance between *BA*- and *BEI*-constructions has also been reported in both offline comprehension and production (e.g., Hao & Chondrogianni, 2021; Zhou & Ma, 2018). Meanwhile, relatively lower offline comprehension accuracy and more production errors of *OSV*-constructions relative to *BA*- and *BEI*-constructions has been observed in monolingual children as well (Hao & Chondrogianni, 2021).

Studies examining the development of these structures in heritage bilingual contexts are few, but provide converging evidence that both Mandarin-English adult HSs (Polinsky et al., 2010) and child HSs (Hao & Chondrogianni, 2021) have different performance in all three structures compared with age-matched monolingual peers in offline comprehension and/or production, and that they show a preference for the shared structure between Mandarin and English. For example, Mandarin-English adult HSs in Polinsky et al. (2010)'s case study opted for SVO-constructions in their production while the monolingual baseline participants preferred *BA*-constructions. A similar picture was found by Mai et al. (2018) in Cantonese-

English adult HSs' production of Cantonese *ZOENG*-constructions, the Cantonese counterpart of *BA*-constructions. This pattern of preferring the shared structure between the two languages suggests CLI from the societal dominant language to the HL.

More recently, Hao and Chondrogianni (2021) examined the offline comprehension and production of all the three structures in Mandarin–English child HSs aged five to nine. Unlike the monolingual children and adults, the child HSs in their study showed a clear performance advantage in *BA*-constructions (relative to *BEI*- and *OSV*-constructions) coupled with a disadvantage in *OSV*-constructions (relative to the *BEI*- and *BA*-constructions) both in offline comprehension accuracy and production (priming magnitude). Additionally, these child HSs produced significantly more *SVO*-constructions when monolinguals produced non-canonical structures even in producing the agent-first *BA*-constructions, similar to the Mandarin–English adult HSs in Polinsky et al. (2010)'s study, reinforcing the argument for CLI.

The Present Study

The present study investigated the production and online and offline comprehension of Mandarin non-canonical structures in Mandarin–English child HSs and how their performance was modulated by language-related factors, i.e., the presence or absence of morphosyntactic cues and CLI, and ID factors, i.e., chronological age and input quantity. Specifically, we asked the following research questions:

1. How do child HSs produce and comprehend (online and offline) the three non-canonical structures, relative to monolingual children?
2. How are child HSs' production and comprehension (online and offline) modulated by the presence or absence of morphosyntactic cues? Is there evidence for CLI from English?
3. How do chronological age and input quantity individually and/or jointly modulate child HSs' production and comprehension (online and offline) of non-canonical structures?

To answer these questions, we adopted a comprehension-to-production priming task to test production, a self-paced listening task with picture verification to examine online and offline comprehension at the same time, and a language background questionnaire to collect child-level information for ID factors.

Predictions

Starting with RQ 1, we expect child HSs to show lesser priming magnitude and more variable production in the priming task and lower accuracy in the offline comprehension task across structures compared with age-matched monolingual peers (as previously observed by Hao & Chondrogianni, 2021). However, even though their offline performance might be worse than that of the monolingual children, child HSs might adopt online processing strategies also deployed by monolingual children (as found in adult HSs, e.g., Jegeriski, 2018a, 2018b). Note that unless we discuss the relative performance of specific tasks, we intentionally select words such as different/differential performance rather than worse/better to describe differences between child HSs and monolingual children (see Kupisch & Rothman, 2018; Ortega, 2020; Rothman et al., 2022 for arguments on why such terminology choices matters).

For RQ 2, firstly, we predicted that if additional linguistic cues have an assistive role (Hsu, 2018) and provided that child HSs show any sensitivities to the morphosyntactic cue (cf. Hao & Chondrogianni, 2021), child HSs will display better performance on the *BEI*-constructions relative to that of *OSV*-constructions. This would mean a larger priming effect in production and higher offline comprehension accuracy. In online processing, the assistive role of morphosyntactic cues might surface in the form of fewer processing costs associated with (re-)analysing the structure with a morphosyntactic cue, i.e., *BEI*-constructions, relative to the one without, i.e., *OSV*-constructions.

Secondly, we expected CLI from English to Mandarin to surface (cf. Hao & Chondrogianni, 2021). Specifically, child HSs would prefer the shared structure between English and Mandarin, i.e., the canonical *SVO*-construction. Additionally, CLI would also interact with word order. Therefore, in production, it would lead child HSs to produce *SVO*-constructions to a larger extent in contexts where monolingual children prefer non-canonical structures (Mai et al., 2018; Polinsky et al., 2010), especially *BEI*- and *OSV*-constructions (Hao & Chondrogianni, 2021). In comprehension, if a *BA*-advantage were to be observed, it could be indirect evidence for CLI given the partial surface overlap between *S(BA-)* *OV* and *SVO* between Mandarin and English.

Turning to child-level factors (RQ 3), and given that the sample of children in the present study belong to second-generation children who may receive more variable input, we expected chronological age and input quantity to potentially modulate child HSs performance across structures and tasks (e.g., Daskalaki et al., 2019; Flores et al., 2017; Paradis, 2023).

Methodology

Participants

A total of 64 five-to-nine-year-old children participated in the study. Thirty-two were bilingual children speaking Mandarin as the HL and English as the societal dominant language (Age in month: mean = 88.43, range = 61–116, SD = 15.53; SES⁴: mean = 16.85, range = 14–18, SD = 1.14). The remaining 32 participants were monolingual children living in China (Age in month: mean = 85.84, range = 60–111, SD = 16.74; SES: mean = 16.81, range = 12–24, SD = 2.26). The two groups were on both age ($t(58) = 0.61$, $p = .54$) and SES ($t(47) = 1.23$, $p = .09$). However, to ensure homogeneity within the heritage group, we excluded four child HSs who spoke another language other than Mandarin and English, leaving 28 child HSs in further analyses. Concerning their migration status, all child HSs were second-generation HSs born and raised in the U.K. Eighteen of them had both Mandarin-speaking parents while the remaining 10 had one Mandarin-speaking parent and another English-speaking parent. As such, they had various Age Onset of Acquisition (AoA) of the societal language English. However, the variation was not large (AoA in months: Mean = 9.32, range = 0–36, SD = 11.34)⁵. Finally, all participants had no reported history of hearing, speech, language, socioemotional or developmental disorders.

Language background questionnaire

We administered the Alberta Language Environment Questionnaire adapted to heritage speakers (Daskalaki et al., 2020) to collect child HSs' language background, e.g., their input quantity (operationalised as current home language use following the previous

studies mentioned before). We used the child's current home language use (HLU) score as a proxy for input quantity. Specifically, parents were asked to rate on a scale from 0 (Mandarin almost never/English almost always) to 4 (Mandarin almost always/English almost never) how frequently the child was spoken to in Mandarin/English by their parents, other guardians (caregivers, grandparents, etc.) and siblings (input) and how frequently the child directed speech to these family members in Mandarin/English (output). The HLU score was operationalised as the mean proportion of Mandarin input and output of the child (HLU: mean = 0.51, range = 0.08–0.84, SD = 0.26). Higher HLU score indexes more Mandarin input.

Production Task

A comprehension-to-production priming task was adopted because it facilitates the production of these structures which are infrequent in naturalistic contexts (Deng et al., 2018). Secondly, it taps into abstract syntactic knowledge in children (Branigan & Pickering, 2017). As such, a smaller priming magnitude would be interpreted as more production difficulties. In the task, participants were presented with a picture on a laptop while a pre-recorded audio clip describing the picture (who did what to whom) was played first (served as a prime), after which participants were asked to describe a new picture shown on the screen and were instructed to describe it as quickly as possible.

To construct the primes, we selected five verbs, i.e., *tui* 'push', *yao* 'bite', *ti* 'kick', *qin* 'kiss', and *ju* 'raise'. For the targets, three new verbs, i.e., *zhui* 'chase', *xi* 'clean', and *wei* 'feed' were selected in addition to *tui* 'push' and *ti* 'kick', which have been used in the primes too. All primes shared the same structure: Noun Phrase (NP) + morphosyntactic cue *ba*, *bei*, or null + NP + Adverb + Verb Phrase (VP). Furthermore, the NPs in all sentences were disyllabic, while all verbs were monosyllabic, followed by an aspectual marker and an adverb (either marking the frequency of the event, i.e., *yi-xia*, 'once' or the result of the action). As for the adverbs, we included *qingqingde* 'gently', *xiaoxinde* 'carefully' and *manmande* 'slowly', immediately after the second NPs and before the VPs. Each of the three adverbs was used three times across verbs. In the primes, each verb appeared nine times and was distributed evenly across conditions so that each condition consisted of 15 trials, making a total of 45 primes. There were no lexical overlaps between the primes and the targets. All the NPs, Adverbs and VPs were approved by the teachers at the Edinburgh Chinese schools, confirming that all participants should have been familiar with them.

In addition, to further limit the role of animacy, and world knowledge among other factors, we ensured that all sentences were semantically reversible and that the typical sizes of the two animals in each sentence were comparable (both in real-world and in the pictures). The last character of all the NPs also covered all four tones in Mandarin, e.g., *zhizhu* 'spider' (high tone), *gongniu* 'cow' (rising tone), *heme* 'hippo' (dipping tone) and *xiaolu* 'deer' (falling tone). Additionally, to avoid order bias and repetition, three separate lists were made. Each prime picture was depicted with all three structures, and the three prime sentences for the same picture appeared in one of the three lists respectively. For instance, (5), (6), and (7) are the *BA*-, the *BEI*-, and the *OSV*-prime for the prime picture (Fig. 1), and they were arranged into list A, B and C respectively (see supplementary materials for the lists).



Figure 1. Example of pictures for primes in the production task and for experimental trials in the comprehension task

- (5) *yizhi shanyang BA yizhi laolang tile yixia*
 one-cl goat ba one-cl wolf kick-perf once
 'A goat kicked a wolf.'
- (6) *yizhi laolang BEI yizhi shanyang tile yixia*
 one-cl wolf bei one-cl goat kick-perf once
 'A wolf was kicked by a goat.'
- (7) *yizhi laolang, yizhi shanyang tile yixia*
 one-cl wolf, one-cl goat kick-perf once
 'A wolf was kicked by a goat.'

In addition to the experimental trials, 20 fillers were included. Each filler consisted of a picture with two animals performing an intransitive action (e.g., *yuedu* 'reading', *shuxie* 'writing', *paobu* 'running', *kaixin* 'being happy' and *tiaoyue* 'jumping'), as in (8). Additionally, all primes were arranged in a pseudorandom order so that trials from the same condition did not appear consecutively. The trial order for each participant was the same.

- (8) *yi-zhi zhizhu kanshu, yi-zhi zhizhu xiexin*
 one-cl spider read one-cl spider write
 'A spider is reading; a spider is writing.'

Comprehension task

Based on the assumption that a mismatch between visual and linguistic stimuli would cause comprehension difficulties, we manipulated the matching between the sentences and the pictures, to examine online and offline comprehension of *BA*-, *BEI*- and *OSV*-constructions with a self-paced listening task with picture verification. The rationale was that if participants could use a particular cue, then a mismatch between the picture and content of the sentence should lead to elevated reaction times (RTs) and worse accuracy in offline comprehension. Therefore, crossing Structure and Matching, six experimental conditions (*BA*-match, *BA*-mismatch, *BEI*-match, *BEI*-mismatch, *OSV*-match, and *OSV*-mismatch) were tested in a within-subject design (see Table 1).

All experimental sentences shared the same structure: I saw + Noun Phrase (NP) + morphosyntactic cue *ba*, *bei*, or null + NP + Adverb + Verb Phrase (VP). For the verbs, we selected all eight verbs used in the production task, i.e., *tui* 'push', *zhui* 'chase', *yao* 'bite', *ti* 'kick', *qin* 'kiss', *xi* 'clean', *ju* 'raise', and *wei* 'feed'. Each of the verbs was used six times across conditions, appeared in each condition. As for the adverbs, we included *qingqingde* 'gently', *xiaoxinde* 'carefully', *kaixinde* 'happily', and *manmande* 'slowly', immediately after the second NPs and before the VPs. Each of the four adverbs was used twice across verbs. Similar to the production task, we ensured that all sentences were semantically reversible and that the typical sizes of the two animals in each sentence were comparable (both in real-world and in the pictures).

Table 1. Experimental conditions for the comprehension task, paired with Figure 3.

| Condition | Segment 1 | Segment 2 | Segment 3 | Segment 4 | Segment 5 |
|---------------------|----------------------|---------------------------------|---|------------------------|--------------------------------|
| BA-match | Wokanjian 'I saw' | yizhi shanyang 'one-cl goat' | BA yizhi laolang 'ba one-cl wolf' | qingqingde 'gently' | tile yixia 'kick-perf once' |
| BA-mismatch | Wokanjian 'I saw' | yizhi laolang 'one-cl wolf' | BA yizhi shanyang 'ba one-cl goat' | qingqingde 'gently' | tile yixia 'kick-perf once' |
| BEI-match | Wokanjian 'I saw' | yizhi laolang 'one-cl wolf' | BEI yizhi shanyang 'bei one-cl goat' | qingqingde 'gently' | tile yixia 'kick-perf once' |
| BEI-mismatch | Wokanjian 'I saw' | yizhi shanyang 'one-cl goat' | BEI yizhi laolang 'bei one-cl wolf' | qingqingde 'gently' | tile yixia 'kick-perf once' |
| OSV-match | Wokanjian 'I saw' | yizhi laolang 'one-cl wolf' | yizhi shanyang 'one-cl goat' | qingqingde 'gently' | tile yixia 'kick-perf once' |
| OSV-mismatch | Wokanjian 'I saw' | yizhi shanyang 'one-cl goat' | yizhi laolang 'one-cl wolf' | qingqingde 'gently' | tile yixia 'kick-perf once' |

The task also included 20 fillers (with half of them adapted from the production task and the remaining half newly created). Each filler consisted of a picture with two animals performing an intransitive action (e.g., *yuedu* 'reading', *shuxie* 'writing', *paobu* 'running', *kaixin* 'being happy' and *tiaoyue* 'jumping'), as in (9). Filler sentences either matched or mismatched the pictures. Specifically, it could be that the picture was about two of the same type of animal performing different actions or two different animals performing the same action. The filler trials were also broken into five segments (indicated in the example with slashes).

- (9) Wo kanjian/ yi-zhi zhizhu/ kanshu/, yi-zhi zhizhu/ xiexin
I saw/ one-cl spider/ read/ one-cl spider/ write
'I saw that a spider is reading; a spider is writing.'

Like the production task, six separate lists were created to ensure any given condition of the same item appeared once in any given list, and across all lists, all conditions of all items were represented (see supplementary materials). Participants were pseudo-randomly assigned to different lists and presented with a full list. The relative position of the agent/patient in the pictures was also counterbalanced, i.e., half trials had agents on the left and half on the right. In each experimental list, all sentences were arranged in a pseudorandom order so that trials from the same condition did not appear consecutively. Additionally, the trial order was the same for each participant.

The experimental sentences were recorded by a male monolingual speaker of Standardised Mandarin (Putonghua) at a normal rate. In segmenting the recorded sentences, we ensured that each segment sounded as natural as possible. No word boundaries were broken in segmentation, and each segment was realised fully. At the end of each sentence, a beep sound was played, and the participants were then asked to judge if the sentence they heard matched the picture. Participants did not receive any feedback throughout the experiment.

Procedure

All participants took part in the study at their homes. We implemented the experimental tasks with JsPsych (de Leeuw, 2015) on a webpage. Each participant participated in all the experimental tasks, and the entire session lasted approximately 50 - 70 minutes, depending on the participants' age. The presentation of the experimental tasks was counterbalanced to cancel out potential

carry-over effects between tasks: the production task was administered first to a random half of the participants, and the remaining participants were first tested with the comprehension task. The whole process of the experiment for each participant was audio-recorded. The language background questionnaire for child HSs was presented via Qualtrics and completed by caregivers before or after their child(ren) worked on the task battery. All the responses were later transcribed and scored by the first author of this paper. Additionally, all participants and their parents were informed of their ethical rights of participation, verbally and in written form, prior to the experiment. Before any tasks, participants (and their parents) were asked to press a button on the web to give consent for their participation. The study has been approved by the institutional ethics committee.

Coding and Scoring

In the production task, recordings were first transcribed by a machine and then checked by the first author of the paper. Because of the property of Mandarin that each word is an individual character and the fact that we are interested in the specific structure, the reliability of coding (between the machine and the transcriber) reached 99%. Disagreements were solved by another Mandarin native speaker. Transcribed sentences were then coded as "BA", "BEI", "OSV", and "SVO" if their utterances encoded correct thematic roles and were complete. On the other hand, complete utterances with a reversed thematic role configuration compared to the picture were coded as "Reversed". Incomplete utterances, code-switching utterances, and utterances failing to establish who did what to whom, e.g., separately describing intransitive actions for the two animals involved in the picture (e.g., example 10), etc., were coded as "Other" and were excluded from further analysis.

- (10) yi-zhi laoshu he yi-zhi laohu zhan zaiyiqi
one-cl mouse and one-cl tiger stand together
'A mouse and a tiger are standing together.'

Additionally, sentences having only one of the two NPs realised were coded as "BA" or "BEI" if (1) the morphosyntactic cue *ba* or *bei* was present, and (2) the realised NP carried the correct thematic role, or as "Reversed" if (1) the morphosyntactic cue *ba* or *bei* was present, and (2) the realised NP carried incorrect thematic role.

In the comprehension task, we measured children's offline comprehension accuracy based on how they responded to the question at the end of each trial. If the participant gave a correct response to the question of whether the sentence matched the picture asked at the end of each trial, the response was scored as "1". Otherwise, it was scored as "0". We then included trials with correct responses for the RT data analysis. In analysing the RT data, we firstly excluded extreme values that are below 500ms or above 5000ms after checking the distribution of the data, as well as outliers that were below or above 2 standard deviations of the mean calculated for each structure per participant and per condition. Then, we converted raw RTs to residual RTs, which were the differences between raw RTs, and predicted RTs calculated for each participant and trial based on the duration of each segment. This allows us to control for the differences in length across trails and segments and individual differences in responding to different items and conditions. Residual RTs were used in further analyses and visualisations of RT data.

Results

Statistical analyses were carried out with the *lme4* package and the *multinom* package in R (R Core Team, 2018). Multinomial logistic regressions, binomial logistic regressions, and generalised linear mixed-effect regressions were adopted to respectively analyse the production data, accuracy data, and RT data (see also Hao & Chondrogianni, 2021). We included the maximal random effects justified by the design where possible (Barr et al., 2013). Specifically, the maximal random effects included both by-subject and by-items random intercepts, as well as by-subject random slopes for Structure and Condition, and by-item random slopes for Group, Structure, and Condition. When the maximal model failed to converge, we tried different optimisers first if possible, using the *afex* package, and then iteratively simplified the random effect structures until convergence was achieved, i.e., removing random effect(s) accounting for the least variance. To identify the optimal model, we adopted the stepwise backward selection approach (unless stated otherwise) starting from the maximal model, using likelihood ratio tests. Variance Inflation Factor was calculated for the optimal models to check for multicollinearity.

Unlike previous priming studies (e.g., Messenger et al., 2012), we analysed both priming effects and overall production patterns when primed. Specifically, we measured (1) if the production of a specific structure was more likely after a prime of the same structure type compared with after a prime of a different structure type (conventional priming effect), and (2) the distribution of different response types produced following different prime types (production patterns). For the production pattern analysis, we compared the degree to which, for example, structure A was produced after prime type A vs. when structure B/C was produced after prime type A. This allowed us to gauge not only the underlying syntactic representations of different groups but also the production patterns/syntactic choices when participants were primed vs. not primed (not producing the prime structure). Because of this, we adopted multinomial logistic regressions as these can incorporate categorical dependent variables with more than two unordered levels.

For the post hoc analyses, pairwise comparisons with Bonferroni-corrected *p*-values were conducted for binomial logistic and generalised linear mixed-effect regressions. For multinomial logistic regressions, we firstly exhausted all possible combinations of

reference levels for all variables and then conducted analyses with reduced models when any significant interactions were attested.

Production task

Because of recording issues, e.g., extreme noise, etc., data from 3 monolingual and 2 child HSs were excluded. We checked again that the two groups were still matched on age and SES after the exclusion of these five children. Overall, the monolingual group produced 1305 responses and the heritage group 1170 responses. We then excluded responses coded as "Other". This resulted in an exclusion of 272 (20.8%) and 332 (28.4%) responses from the monolingual and heritage group respectively.

Figure (2) shows the proportions of different response types across prime types. To answer our research questions statistically, multinomial logistic regressions were fitted with Response Type (BA, BEI⁶, OSV, SVO, and Reversed) as the dependent variable. As fixed effects, Group (Monolingual and Heritage; RQ 1) and Prime Type (BA, BEI, and OSV; RQ 2) were entered. The optimal model included the interaction between the two independent variables. However, given the complexity of a multinomial logistic model and the fact that all our independent variables are categorical in nature, we only included significant individual simple effects in Table (2). This is because any interactions in the optimal model were compared to the reference level, i.e., the use of BEI-constructions after BEI-primed by the monolingual group, which would not be meaningful for our interests, e.g., consider comparing the reference level to the use of BA-constructions after BEI-primed by the heritage group.

The optimal model suggested that after BEI-primed, (1) both groups were more likely to produce BEI-constructions than after other prime types; (2) priming was stronger in the monolingual group than in the heritage group; (3) the monolingual group produced more BA-constructions than the heritage group did; and (4) the heritage group produced more SVO-constructions and reversal errors than the monolingual group did.

Post hoc analyses suggested that firstly, for priming, the likelihood of producing a specific structure was the highest when the participants were firstly exposed to a prime of the same structure for both groups, i.e., a priming effect was observed across structures for both groups. Additionally, both groups were less primed by OSV-constructions than by BEI-constructions. However, when the monolingual group was primed to the same extent by BA- and BEI-primed, the priming magnitude was strongest after BA-primed than after BEI-primed for the heritage group. Secondly, for production patterns, the analyses showed that when not producing the prime structures, the heritage group produced significantly more reversal errors than the monolingual group did across structures. After BEI- and OSV-primed, the monolingual group mostly produced the other two non-canonical structures when not primed, but the heritage group resorted predominantly to SVO-constructions (to an extent that SVO-constructions were the most used structure even compared to the primed).

We then fitted models to examine how individual-level factors, i.e., age and HLU, modulated the heritage group's priming and syntactic choices after different primes. The optimal model (see supplementary materials) included Prime Type (BA, BEI, and OSV), Age (Scaled), and HLU (Scaled) without any interactions as fixed effects. Specifically, priming increased with the increase in age and of HLU. For production patterns, age negatively modulated child HSs' likelihood of producing reversal errors (Estimate

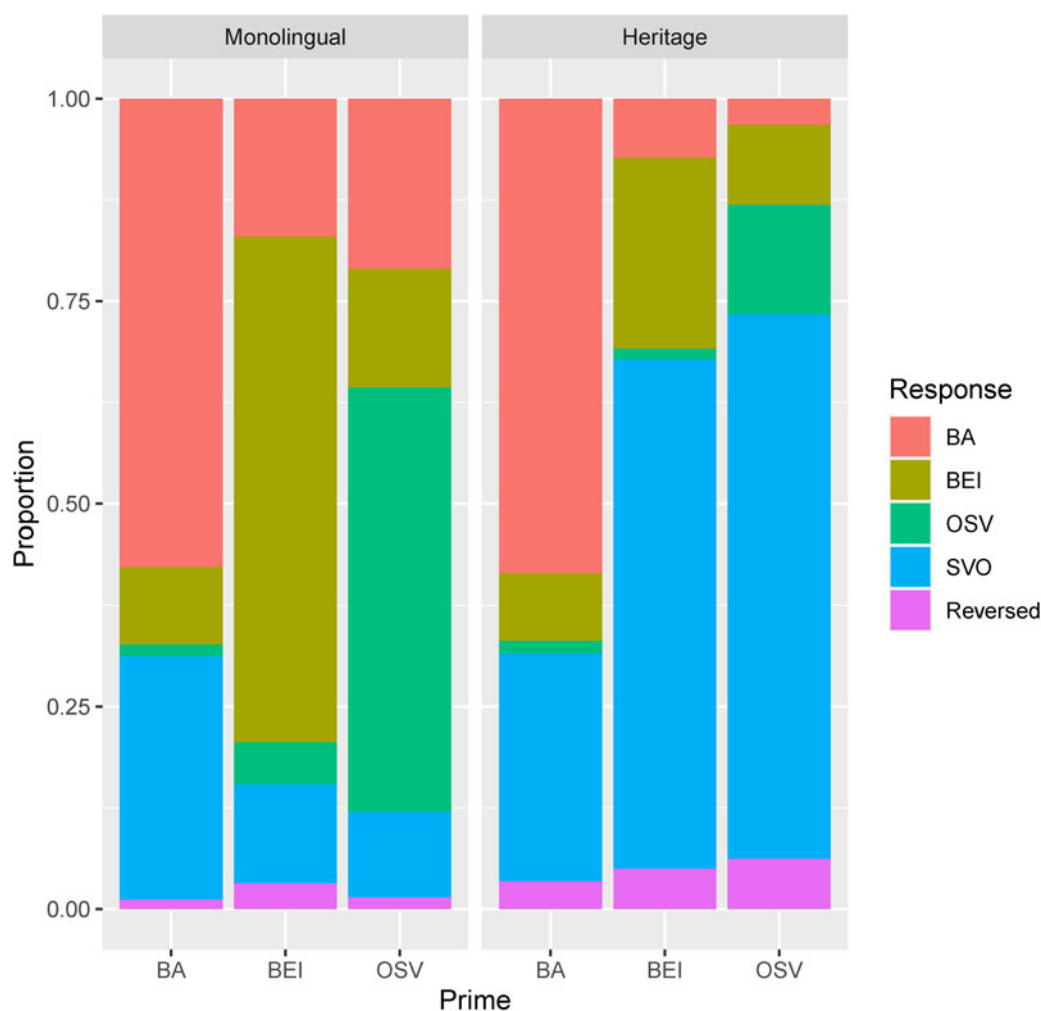


Figure 2. Proportion of response types following different prime types in the monolingual and heritage groups

$= -1.14$, $SE = .36$, $b = -3.09^{**}$, $p < .01$). HLU negatively predicted their probability of producing SVO-constructions (Estimate = -0.31 , $SE = .11$, $b = -2.76^{**}$, $p < .01$). Recall that the lower HLU, the more English input the child had.

Comprehension task

Accuracy data

Figure (3) shows the offline accuracy for both groups across structures and conditions. To statistically compare the two groups (RQ 1) and see if structure type modulates any potential group differences/similarities (RQ 2), we fitted binomial logistic regressions with Group (**Monolingual** and Heritage), Structure (BA, **BEI**, and OSV) and Condition (**Match** and Mismatch) as fixed effects. The optimal model (Table 3) included a two-way interaction between Group and Structure.

The optimal model together with the post hoc analyses revealed that (1) the monolingual group outperformed the heritage group across structures and conditions; (2) both groups had more errors in mismatched conditions than in matched conditions; (3) for both groups, OSV-constructions induced more errors than BEI- and BA-constructions; (4) the two groups differed in whether BA- and BEI-constructions received similar accuracy. Specifically, across conditions, BA- and BEI-constructions were comprehended

equally well by the monolingual group, whereas the heritage group showed higher accuracy in BA- than in BEI-constructions across conditions.

To understand the role of age and input quantity (RQ 3), we included only the heritage group and ran binomial logistic regressions using a forward stepwise approach. We firstly established a base model with Structure and Condition as fixed effects (the optimal model did not include the interaction term). Then, we added each child-level factor and ran likelihood ratio tests against the base model. We then built the final model with all main factors added stepwise until we identified the optimal model. The results revealed that HLU (input quantity) but not age modulated the heritage group's accuracy across structures (see figure 4 for a visualisation). Specifically, HLU positively predicted accuracy across structures for the heritage group (Estimate = 0.35 , $SE = 0.13$, $t = 2.64^{**}$, $p < .01$).

Reaction Times

For RT analyses, we included the trials where participants gave correct offline comprehension responses. Figure (5) illustrates how listening times (represented by residual RTs; RTs henceforth) contrast between the monolingual and the heritage group across segments, conditions, and structure types.

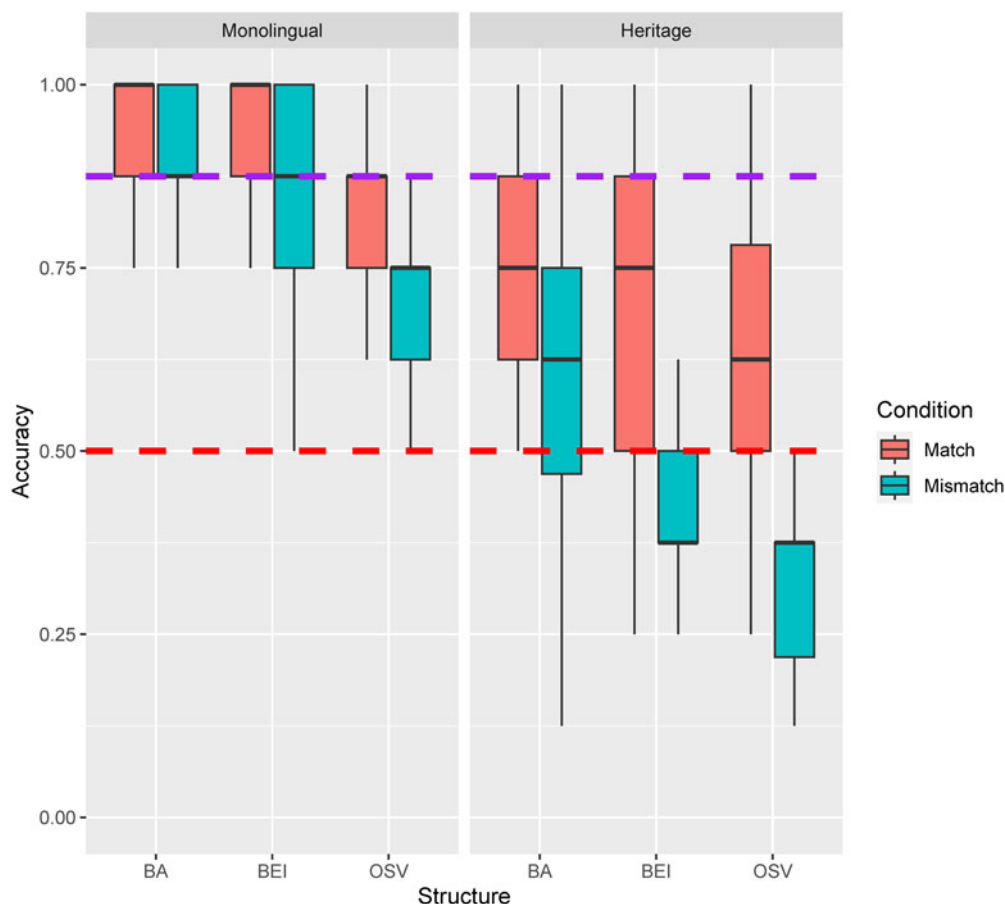


Figure 3. Offline comprehension accuracy across conditions and structures in the monolingual and heritage groups

Linear mixed-effect models were fitted with residual RTs (scaled) as the dependent variable. For the independent variables, we firstly entered Group (**Monolingual** and Heritage; RQ 1), Structure (BA, **BEI**, and OSV; RQ 2), Condition (**Match** and Mismatch) as well as Segment (Segment 1, Segment 2, **Segment 3**, Segment 4, and Segment 5). However, adding Segment into the model led to convergence issues and/or singular fit with all possible random effect structures. Therefore, we then ran models separately for each segment. The models for Segments 1 and 2 failed to show any significant effects. Thus, only models for Segment 3, 4, and 5 were reported.

Table (4) shows the optimal model for Segment 3 which included only Structure, Condition and their interaction (but not Group) as fixed effects. The model suggested that both groups took a longer time listening to the mismatched conditions across structures. The interaction term was driven by the fact that the matching effect was stronger in OSV-constructions. Specifically, participants spent more time in the OSV-mismatch condition as opposed to the BEI-mismatch condition.

For Segment 4, the optimal model again included only Structure, Condition and their interaction but not Group as fixed effects. However, Group was included in the optimal model as a fixed effect for Segment 5 along with Condition (Structure was not selected in the optimal model for Segment 5) as well as their interaction. See table (5) for significant simple effects in the optimal models for Segment 4 and 5.

The analyses showed that the matching effect lingered to Segment 4 for both groups and further lingered to Segment 5

but only for the heritage group. Similar to Segment 3, the matching effect was more prominent for OSV-constructions in Segment 4 but such a difference among structures was no longer present in Segment 5.

Finally, to examine how child-level factors modulate child HSs' online processing (RQ 3), we included only the heritage group and ran another set of analyses. Similarly, we firstly included Segment as a fixed effect in addition to Condition and Structure, which failed to converge. We then ran models for each segment. However, neither age nor HLU was selected in any optimal models across segments.

Discussion

The present study compared Mandarin–English child HSs to their age-matched monolingual peers in their production and comprehension of Mandarin non-canonical structures (RQ1). To do so, the current study also increased the methodological granularity within HL research. For the elicitation of non-canonical structures' production, a novel priming task was adopted. Furthermore, this is also one of the few studies to examine how child HSs comprehend these non-canonical sentences in real-time using an online comprehension task and how they comprehend these structures when they have had the time to reflect at the end of the sentence, which we deemed to indicate offline comprehension accuracy. We also examined if and how linguistic factors (RQ 2) and ID factors (RQ 3) influence the production and

Table 2. Optimal model with Group (**Monolingual** and Heritage) and Prime Type (BA, **BEI**, and OSV) as fixed effects for all valid responses (BA **BEI**, OSV and SVO and Reversed) in the production task.

| Response Type | | Estimate | Standard Error | b | p |
|------------------------|---------------------------------|----------|----------------|-----------|-------|
| BEI vs BA | Intercept | -1.29 | .14 | -8.79*** | <.001 |
| | Prime Type | | | | |
| | BEI vs BA | 3.09 | .23 | 12.97*** | <.001 |
| | BEI vs OSV | 1.66 | .24 | 7.03*** | <.001 |
| | Group:Prime Type | | | | |
| | Monolingual:BEI vs Heritage:OSV | -1.58 | .251 | -3.08* | <.01 |
| BEI vs OSV | Intercept | -2.48 | .25 | -10.10*** | <.001 |
| | Prime Type | | | | |
| | BEI vs OSV | 3.76 | .29 | 12.82*** | <.001 |
| BEI vs SVO | Intercept | -1.63 | .17 | -9.68*** | <.001 |
| | Group | | | | |
| | Monolingual vs Heritage | 2.61 | .22 | 11.94*** | <.001 |
| | Prime Type | | | | |
| | BEI vs BA | 2.78 | .26 | 10.63*** | <.001 |
| | BEI vs OSV | 1.30 | .28 | 4.72*** | <.001 |
| | Group:Prime Type | | | | |
| | Monolingual:BEI vs Heritage:OSV | -1.58 | .51 | -3.08** | <.01 |
| BEI vs Reversed | Intercept | -2.97 | .31 | -9.61*** | <.001 |
| | Group | | | | |
| | Monolingual vs Heritage | 1.42 | .42 | 3.37*** | <.001 |

Notes. *p < .05, **p < .01, ***p < .001

comprehension of non-canonical structures in these child HSs. To this end, we extended a previous study by Hao and Chondrogianni (2021) to a more diverse group of second-generation children. Specifically, we included an online comprehension measure alongside production and offline comprehension measures in a group of second-generation child HSs with varied chronological age and input quantity (operationalised as current home language use). Additionally, the structures we tested, i.e., BA-, BEI-, and OSV-constructions, differ from each

other in the presence or absence of morphosyntactic cues and word order.

We hypothesised that child HSs might be less likely to be syntactically primed in production because they have relatively less experience with the HL and the specific structures tested in the study (see Contemori, 2022 for a discussion of how and why priming magnitude varies as a function language experience). In the case of the online processing task, we measured both children’s comprehension of non-canonical structures in real-time, and their offline accuracy on these constructions at the end of the sentence. We predicted that child HSs may have lower offline comprehension accuracy of these structures compared with their monolingual peers (see also Hao & Chondrogianni, 2021). However, the two groups might show similar online processing patterns (as seen in adult HSs in Jegerski, 2018b). For linguistic factors, we expected child HSs’ performance to be modulated by the presence or absence of morphosyntactic cues, across task modalities, so that BEI-constructions should receive better performance relative to OSV-constructions. We also predicted that CLI may surface in production, in the form of child HSs’ avoidance of the non-canonical structures and preference for the structure that is shared between the two languages (cf. Mai et al., 2018). Additionally, we postulated that CLI would interact with word order of the structures, i.e., structures sharing word order with the shared structures between the two languages would be less affected by CLI, at least in production. On the other hand, we predicted age to positively modulate child HSs’ performance (Hao & Chondrogianni, 2021), and the role of input quantity to be more

Table 3. Optimal model with Group (**Monolingual** and Heritage), Structure (BA, **BEI**, and OSV) and Condition (**Match** and Mismatch) as fixed effects for the accuracy data in the comprehension task.

| | Estimate | Standard Error | t | p |
|----------------------|----------|----------------|----------|-------|
| (Intercept) | 2.90 | .21 | 13.81*** | <.001 |
| Heritage | -1.99 | .24 | -8.35*** | <.001 |
| BA | 0.33 | .21 | 1.52 | .13 |
| OSV | -1.09 | .18 | -6.04*** | <.001 |
| Mismatch | -1.08 | .14 | -7.71*** | <.001 |
| Heritage: BA | 0.17 | .26 | 0.66 | .51 |
| Heritage: OSV | 0.63 | .23 | 2.73** | <.01 |

Notes. *p < .05, **p < .01, ***p < .001

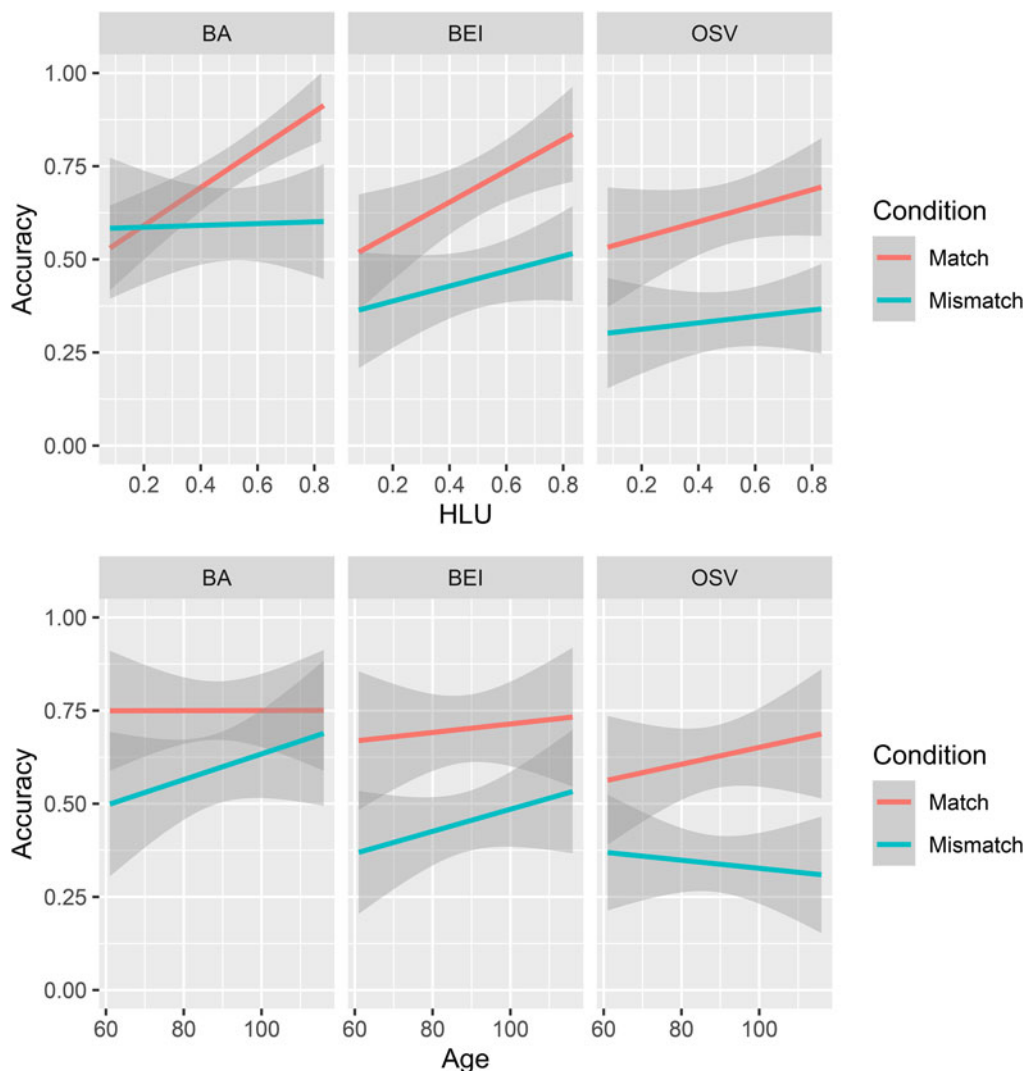


Figure 4. The relationship between offline comprehension accuracy and child-level factors across structures and conditions for the heritage group

prominent in this study compared to the previous study of Hao and Chondrogianni (2021) with first-generation children and in-line with other HL studies (e.g., Daskalaki et al., 2019; Janssen, 2016).

Child HSs vs. monolingual children

Overall, as we predicted for RQ 1, the heritage group had significantly weaker priming magnitude and lower offline comprehension accuracy across the three non-canonical structures, relative to the monolingual group. Contrary to this, the heritage group adopted qualitatively similar online processing strategies in processing non-canonical structures, compared with the monolingual group. In addition, the similarities and differences between the heritage and the monolingual groups manifested themselves differently not only in different tasks but also in different structures. Starting with production, we found that while both monolingual and heritage groups showed less priming by OSV-primers than by BEI-primers, the heritage group was also more likely to be primed by BA-constructions than by BEI-constructions, which was not observed in the monolingual group. Group differences in production patterns were also observed, i.e., the heritage group produced

more SVO-constructions after BEI- and OSV-primers but not after BA-primers and more reversal errors after all prime types, relative to the monolingual group. Following Hao and Chondrogianni (2021) we postulate that the lesser priming magnitude of OSV-constructions was caused by the lack of morphosyntactic cues and the stronger priming of BA-constructions resulted from the fact that there is an agreement between its word order and the dominant agent-patient ordering (an effect of CLI).

As for offline comprehension, like our production data, we found the same hierarchy within the non-canonical structures for the heritage and monolingual groups respectively. Specifically, OSV-constructions received the worst performance relative to the other two non-canonical structures, for both the heritage and the monolingual groups. Nonetheless, while the monolingual group comprehended BA- and BEI constructions equally well, better comprehension accuracy in BA-constructions than in BEI-constructions was observed in the heritage group. The fact that OSV-constructions induced more comprehension errors relative to BEI-constructions was not surprising, because the latter has a free-standing morphosyntactic cue assisting interpretation, and because this OSV-disadvantage was observed in both the heritage and the monolingual group. On the other hand, we attributed the

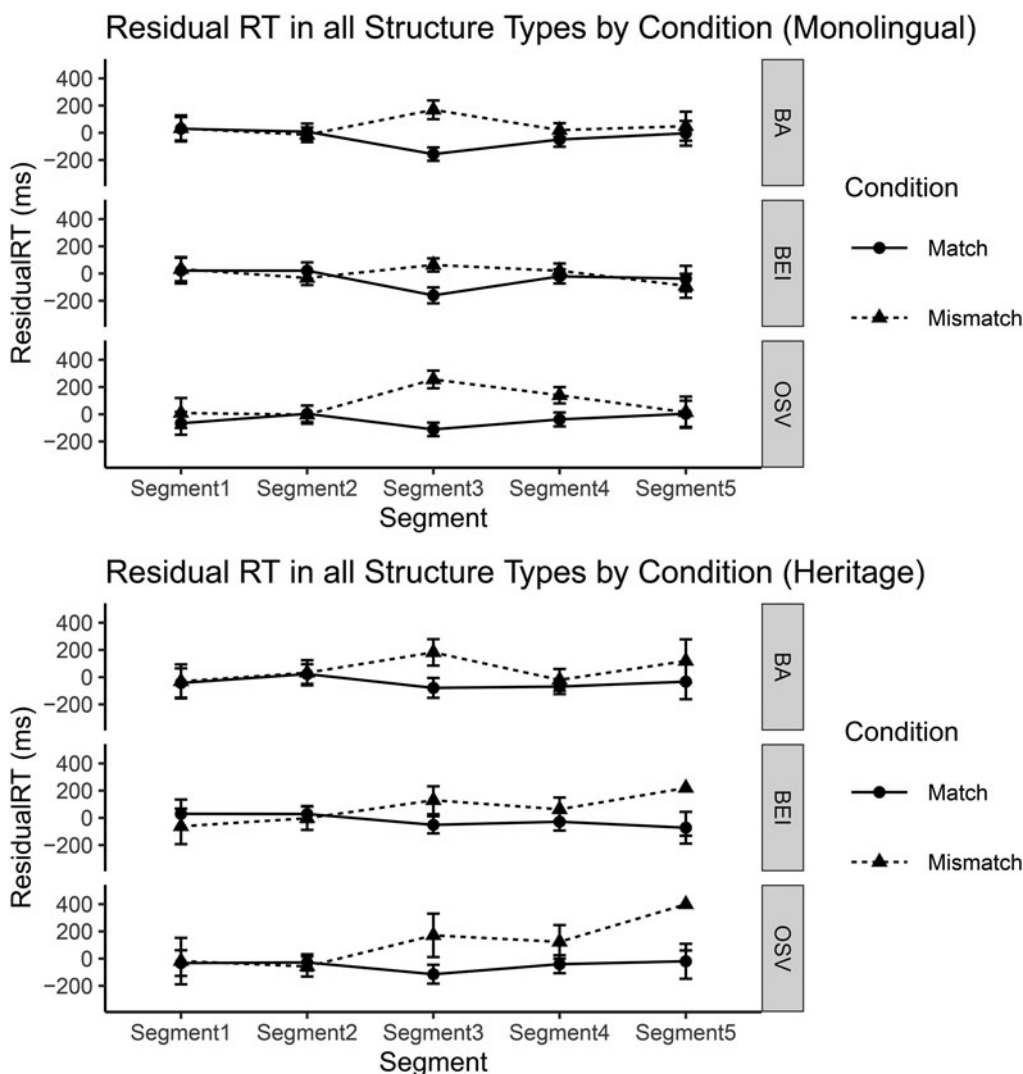


Figure 5. Residual RTs for the monolingual and heritage groups crossed with Condition and Structure Type

BA-advantage, found only in the heritage group, to CLI, as the word order in BA-constructions agrees with the dominant agent-patient ordering found in both Mandarin and English (Hao & Chondrogiani, 2021). However, as we will discuss later, our current experimental design does not allow us to tell for sure if such a BA-advantage in offline comprehension is a result of CLI or a

reflection of the heritage group’s overall preference for interpreting NPIs as agents, which we leave for future research.

Turning to online comprehension, both the heritage and the monolingual groups used the morphosyntactic cue *ba* and *bei* and word order information (two NPs) immediately when they were available (Segment 3), as manifested by the effect of matching. On the other hand, although the effect of matching lasted to the post-critical segment (Segment 4) for both groups, it further lingered to the final segment (Segment 5), which contained the verbal information, only for the heritage but not the monolingual group. The lack of matching effect for the monolingual group in the final segment suggested complete reanalysis processes before the VP. That is to say, the monolingual group had finished reanalysing their initial (mis-)interpretations before they had the verbal information. In contrast, for the heritage group, the effect of matching lasted further to Segment 5, indicating continued reanalysis processes for this group. Here, we postulate that it was the fact that the heritage group took longer to revise their initial (mis-)interpretations, if they could do this at all, that led to their worse offline comprehension accuracy.

Both groups also displayed an OSV-disadvantage in online processing. Specifically, the matching effect was found to be more

Table 4. Optimal model with Structure (BA, BEI, and OSV) and Condition (Match and Mismatch) as fixed effects for the RTs in Segment 3.

| | Estimate | Standard Error | t | p |
|--------------|----------|----------------|----------|-------|
| (Intercept) | -0.29 | .08 | -3.62*** | <.001 |
| BA | -0.01 | .07 | -0.08 | .93 |
| OSV | 0.03 | .08 | 0.33 | .74 |
| Mismatch | 0.49 | .10 | 4.96*** | <.001 |
| BA:Mismatch | 0.21 | .10 | 2.19* | .03 |
| OSV:Mismatch | 0.32 | .10 | 3.04** | <.01 |

Notes. *p < .05, **p < .01, ***p < .001

Table 5. Optimal models for the RT data in Segment 4 and 5

| | | Estimate | Standard Error | T | p |
|------------------|-------------------|----------|----------------|--------|------|
| Segment 4 | Mismatch | 0.13 | .07 | 1.69* | .05 |
| | OSV:Mismatch | 0.32 | .11 | 2.80** | <.01 |
| Segment 5 | Heritage:Mismatch | 0.35 | .12 | 2.77** | <.01 |

Notes. *p < .05, **p < .01, ***p < .001

prominent in OSV-constructions than in BEI-constructions in Segments 3 and 4. This suggested the assistive role of morphosyntactic cues in sentence processing. However, a BA-advantage, found in both production and offline comprehension, was not observed here in the heritage group's online processing of these structures. So, why didn't a BA-advantage surface in online processing as well, such that BA-constructions induced fewer processing costs associated with reanalysis processes, i.e., a less prominent matching effect? Although the current design cannot provide direct evidence, we postulate that the lack of BA-advantage might be caused by cue validity differences. More precisely, the cue validity of *bei* has been empirically tested to be the strongest syntactic cue for Mandarin monolingual adults in sentence comprehension (Li et al., 1992). Therefore, from a performance level, the strong cue validity favouring BEI-constructions might have cancelled out any possible BA-advantage, masking any potential differences. We leave it for future research to test if an agreement of word order between (non-canonical) structures in the HL and the structure shared between the HL and the dominant language (agent-patient ordering in the current study, i.e., BA-constructions and SVO-constructions) would lead to online processing advantage of that HL structure for child HSs.

To sum up, child HSs were less likely to be primed in production and showed worse performance in their offline comprehension of non-canonical structures than their age-matched monolingual peers. However, when they made correct end-of-sentence interpretations of these structures, they deployed the same processing strategies as monolingual children did, albeit with prolonged reanalysis processes, i.e., they took longer to recover from initial (mis)interpretations (as in mismatched conditions), which might be caused by their limited processing budget (used to balance and inhibit the relevant grammars) – see Polinsky and Scontras (2020) for more related arguments.

Linguistic factors

Our second research question was about the role of linguistic factors in modulating child HSs' production and comprehension of non-canonical structures. As we have discussed in the previous section, we found that heritage children's production and comprehension were affected by linguistic factors also applicable to monolingual children, i.e., the presence or absence of morphosyntactic cues, as well as factors unique to themselves, i.e., CLI from the societal dominant language to the HL, which also interacted with another linguistic factor, i.e., word order. It is worth noting here that although we term CLI as a linguistic factor, it is less so as a pure linguistic factor compared to the presence or absence of morphosyntactic cue in the sense that it also varies across individual as a function of language dominance for example (see Chondrogianni, 2023; Van Dijk et al., 2021).

For the role of the presence or absence of morphosyntactic cue, we have evidence across task modalities (and groups) that the

presence of a morphosyntactic cue assisted the production and comprehension (offline and online) of non-canonical structures (see also Hao & Chondrogianni, 2021). Specifically, OSV-constructions induced weaker overall priming magnitude, worse offline comprehension accuracy and a more prominent matching effect indexing greater reanalysis costs in real-time compared with BEI-constructions (for both the heritage and the monolingual group). This is in line with previous research showing that heritage children's performance on non-canonical structures is modulated by the presence or absence of morphosyntactic cue or their transparency (e.g., Janssen, 2016) and their sensitivity to relevant morphology (Kim et al., 2018). However, an effect of frequency might also lead to worse performance in the OSV-construction relative to the BEI-construction, especially when this was observed across groups. For now, we interpreted the results assuming any frequency effect, if there is any, to be secondary. This is because (i) previous studies showed that the development of Mandarin non-canonical structures might not be driven by input frequency (Deng et al., 2018); and (ii) the performance on the relatively more frequent BA-construction is not always better even in the heritage group. However, as the OSV-construction was not included in the study by Deng et al. (2018), future studies looking at properties with similar frequency but differing in the presence or absence of morphosyntactic cues would provide more direct insights.

Turning to the effect of CLI from the dominant language to the HL, we followed previous research and hypothesised that it would lead the heritage group to prefer the structure in the HL that overlaps on the surface with the societal dominant language, i.e., the canonical SVO-construction (e.g., Chondrogianni & Schwartz, 2020; Polinsky et al., 2010). As we hypothesised, we found a relatively more direct indication of CLI in the heritage group's production. Specifically, the heritage group was more likely to produce SVO-constructions than the monolingual group even when primed, which we interpret as an effect of CLI. Additionally, we also found that CLI interacted with word order (see also Hao & Chondrogianni, 2021). And this was observed in both priming magnitude and production patterns. Specifically, it led to stronger priming of BA-constructions relative to BEI- and OSV-constructions in the heritage group and overproduction of the SVO-construction especially when the structures require thematic role reversal, i.e., BEI- and OSV-constructions.

In offline comprehension, a possible indication of CLI would be the heritage group's preference for interpreting non-canonical structures as canonical. However, because our current design only allows us to tell if participants' interpretation of non-canonical structures were correct, we did not have direct evidence for what exact meanings the participants drew in the end. Nonetheless, if CLI interacts with word order (Hao & Chondrogianni, 2021), the heritage group should have higher accuracy in comprehending BA-constructions relative to BEI- and OSV-constructions. Indeed, this is what we observed in the

current study. That being said, this observation could only be indirect evidence for CLI based on the assumption that it also interacted with word order. This is because a better comprehension accuracy of *BA*-constructions could also be a result of the heritage group's reliance on interpreting NPs as agents. Further research would benefit from testing Mandarin-speaking child HSs whose societal dominant language's canonical word order is not agent-patient.

Turning to online processing, following our previous postulation that a *BA*-advantage would be indirect evidence for CLI, we found no evidence in online comprehension. However, this might be an illustration of different cue validities as we argued above. Additionally, it might also be the case that to have CLI to surface in online processing, partial overlap is required and a more time-sensitive measure would be needed (cf. van Dijk et al., 2022).

In sum, the current study revealed the assistive role of morphosyntactic cues in children's production and comprehension of non-canonical structures and CLI in child HSs at least directly in production which was also modulated by word order.

Individual differences factors

If and how ID factors, i.e., chronological age (child-internal) and input quantity (proximal child-external), affect child HSs' performance on non-canonical structures differed across task modalities (RQ 3). Based on the findings of Hao and Chondrogianni (2021), we expected an influence of chronological age and potentially input quantity on heritage children's production and comprehension of non-canonical structures. However, an age effect was only observed in the production data. On the other hand, input quantity predicted the heritage group's production performance and offline comprehension accuracy. Additionally, the heritage group's online processing of these structures was not modulated by chronological age and/or input quantity.

For the effect of input quantity, child HSs with more input in the HL (higher HLU score) had a larger priming magnitude and better comprehension accuracy across structures, regardless of their age. Although this contrasts with Hao and Chondrogianni (2021) who found no effect of input on 5- to 9-year-old Mandarin-English heritage children's comprehension and production of the exact same three non-canonical structures, other recent studies have found that input quantity in the HL can override chronological age effects in HL comprehension and production (e.g., Chondrogianni & Schwartz, 2020; Daskalaki et al., 2019) and proposed that HL input quantity might be the trigger for the observed heritage and monolingual differences in performance (Polinsky & Scontras, 2020).

On the other hand, an age effect was observed in the production data, such that children's ability to get primed by the different structures improved as they got older. So, what is driving the fact that the age effect surfaced in production but not also in offline comprehension? Apart from the fact that chronological age is a child-internal factor and input quantity is a proximal/external factor to the child, we postulate that age and input quantity might actually modulate different aspects of bilingualism. Specifically, HL input quantity, or more generally how much experience the child has had with the HL, might have an impact on how likely the child experiences CLI, while age determines the degree of development of HL meta-linguistic awareness. Some of the current findings lent evidence to this postulation. Firstly, the

(over-)production of *SVO*-constructions, an index for CLI from the dominant language to the HL, was modulated by input quantity but not age. Secondly, because most reversal errors did show syntactic priming, they were more likely to reflect uncertainty in using the HL and/or less developed meta-linguistic awareness in the HL. And importantly, age but not input quantity predicted how many reversal errors the heritage group made.

Overall, the current study found that different ID factors surfaced differently across task modalities. Input quantity modulated both child HSs' performance in both production and offline comprehension, while age only predicted their production. Additionally, when child HSs interpret non-canonical structures accurately and adopt monolingual-like processing strategies in real-time, their online processing was no longer modulated by these ID factors. Although we cannot tell for sure the underlying reasons for the differential effect ID factors have on different task modalities, it merits further research to understand how and, importantly, why different ID factors affect individual HL development (see Paradis, 2023 for a summary of current such attempts).

Conclusion and limitations

Three main conclusions can be drawn from the current results. Firstly, although child HSs are different in the production and offline comprehension of non-canonical structures compared to age-matched monolingual children, they are qualitatively similar in online processing. Secondly, child HSs' production and comprehension of non-canonical structures are modulated by the presence or absence of morphosyntactic cues, and their production shows a (direct) effect of CLI, which further interacts with word order. Thirdly, input quantity predicts their production and offline comprehension accuracy, whereas chronological age only modulates their productive ability.

The observed overall group-level similarities and differences (between monolingual and child HSs) fit in the existing literature. Within child HSs, individual differences were well attested. However, given the current sample, analysis, and research questions, there are several important aspects the current study could not investigate, suggesting limitations of the current study. Firstly, within the current heritage sample, certain participants have two Mandarin-speaking parents while others have one English-speaking parent. This has implications for Age onset of Acquisition of the majority language, i.e., HSs can be simultaneous or sequential bilinguals, the former typically obtaining when only one parent is a speaker of the HL and/or both parents are themselves (second generation) bilinguals of the HL, dominant in the societal majority language. Secondly, as we do not have a fit-for-purpose objective proficiency measure, matching and/or understanding of its effect could not be conducted. Thirdly, although the current study assumed structural frequency not to matter based on available monolingual literature and argued indications of frequency effects to be secondary, such effects would be more readily examined when corpus data, preferably bilingual ones, are available. We leave these questions for future research and caution readers not to over-generalise results. Meanwhile, because all testing was conducted online at participants home, replications both in lab and at home are encouraged.

Acknowledgements. This research was supported in part by research funds from the School of Philosophy, Psychology and Language Sciences to the first author. We would like to thank Professor Elma Blom and Dr Hannah Rohde

for their valuable comments and suggestions. We would also like to thank ISBPAC2022 for the opportunity to present the research and the discussions/comments/feedback we received there.

Data availability. Supplementary materials, including data and statistical analyses, to this article can be found online at <https://osf.io/zndka/>.

Notes

- 1 In this paper, we term structures with non-canonical word order and the canonical word order non-canonical structures and the canonical structure respectively, unless stated otherwise.
- 2 Throughout the study, the first language mentioned in a language pair as such is the HL while the second the societal dominant language.
- 3 In this paper, when we refer to subject or object, unless otherwise specified, we refer to notional subject or object, i.e., the actual doer of the action (corresponding to the semantic notion of agent) and the receiver of the action (corresponding to the semantic notion of patient) respectively.
- 4 SES = socioeconomic status, measured by maternal education level in years.
- 5 As suggested by an anonymous reviewer, we also ran analyses to ascertain if AoA played a role. Because it was never selected in the optimal models, we did not report related analyses. But see the R scripts for more information.
- 6 Unless specified, bold-faced levels are chosen as the reference level for the variable.

References

- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Branigan, H. P., & Pickering, M. J. (2017). An experimental approach to linguistic representation. *Behavioral and Brain Sciences*, 40, e282. <https://doi.org/10.1017/S0140525X16002028>
- Chomsky, N. (1993). *Lectures on government and binding: The Pisa lectures* (7th ed). Mouton de Gruyter.
- Chondrogianni, V. (2023). Chapter 12. Cross-linguistic influences in bilingual morphosyntactic acquisition. In I. Elgort, A. Siyanova-Chanturia, & M. Brysbaert (Eds.), *Bilingual Processing and Acquisition* (Vol. 16, pp. 294–315). John Benjamins Publishing Company. <https://doi.org/10.1075/bpa.16.12cho>
- Chondrogianni, V., & Schwartz, R. G. (2020). Case marking and word order in Greek heritage children. *Journal of Child Language*, 1–30. <https://doi.org/10.1017/S0305000919000849>
- Contemori, C. (2022). Chapter 2. Syntactic priming and child language competence: A critical review and methodological considerations for future research. In K. Messenger (Ed.), *Trends in Language Acquisition Research* (Vol. 31, pp. 15–34). John Benjamins Publishing Company. <https://doi.org/10.1075/tilar.31.02con>
- Daskalaki, E., Blom, E., Chondrogianni, V., & Paradis, J. (2020). Effects of parental input quality in child heritage language acquisition. *Journal of Child Language*, 47(4), 709–736. <https://doi.org/10.1017/S0305000919000850>
- Daskalaki, E., Chondrogianni, V., Blom, E., Argyri, F., & Paradis, J. (2019). Input effects across domains: The case of Greek subjects in child heritage language. *Second Language Research*, 35(3), 421–445. <https://doi.org/10.1177/0267658318787231>
- de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, 47(1), 1–12. <https://doi.org/10.3758/s13428-014-0458-y>
- Deng, X., Mai, Z., & Yip, V. (2018). An aspectual account of *ba* and *bei* constructions in child Mandarin. *First Language*, 38(3), 243–262. <https://doi.org/10.1177/0142723717743363>
- Flores, C., Santos, A. L., Jesus, A., & Marques, R. (2017). Age and input effects in the acquisition of mood in Heritage Portuguese. *Journal of Child Language*, 44(4), 795–828. <https://doi.org/10.1017/S0305000916000222>
- Hao, J., & Chondrogianni, V. (2021). Comprehension and production of non-canonical word orders in Mandarin-speaking child heritage speakers. *Linguistic Approaches to Bilingualism*. <https://doi.org/10.1075/lab.20096.hao>
- Hsu, D.-B. (2018). Children’s syntactic representation of the transitive constructions in Mandarin Chinese. *PLOS ONE*, 13(11), e0206788. <https://doi.org/10.1371/journal.pone.0206788>
- Huang, C.-T. J., Li, Y. A., & Li, Y. (2009). *The syntax of Chinese*. Cambridge University Press.
- Huang, Y. T., Zheng, X., Meng, X., & Snedeker, J. (2013). Children’s assignment of grammatical roles in the online processing of Mandarin passive sentences. *Journal of Memory and Language*, 69(4), 589–606. <https://doi.org/10.1016/j.jml.2013.08.002>
- Janssen, B. (2016). *The acquisition of gender and case in Polish and Russian: A study of monolingual and bilingual children* [PhD Dissertation]. University of Amsterdam.
- Jegerski, J. (2018a). Sentence Processing in Spanish as a Heritage Language: A Self-Paced Reading Study of Relative Clause Attachment: Sentence Processing in Spanish as a Heritage Language. *Language Learning*, 68(3), 598–634. <https://doi.org/10.1111/lang.12289>
- Jegerski, J. (2018b). The processing of the object marker *a* by heritage Spanish speakers. *International Journal of Bilingualism*, 22(6), 585–602. <https://doi.org/10.1177/1367006916681083>
- Jegerski, J., & Sekerina, I. A. (2021). The Psycholinguistics of Heritage Languages. In S. Montrul & M. Polinsky (Eds.), *The Cambridge Handbook of Heritage Languages and Linguistics* (1st ed., pp. 449–470). Cambridge University Press. <https://doi.org/10.1017/9781108766340.021>
- Keating, G. D., Jegerski, J., & Vanpatten, B. (2016). Online processing of subject pronouns in monolingual and heritage bilingual speakers of Mexican Spanish. *Bilingualism: Language and Cognition*, 19(1), 36–49. <https://doi.org/10.1017/S1366728914000418>
- Kidd, E., Chan, A., & Chiu, J. (2015). Cross-linguistic influence in simultaneous Cantonese–English bilingual children’s comprehension of relative clauses. *Bilingualism: Language and Cognition*, 18(3), 438–452. <https://doi.org/10.1017/S1366728914000649>
- Kim, K., O’Grady, W., & Schwartz, B. D. (2018). Case in Heritage Korean. *Linguistic Approaches to Bilingualism*, 8(2), 252–282. <https://doi.org/10.1075/lab.16001.kim>
- Kupisch, T., & Rothman, J. (2018). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 22(5), 564–582. <https://doi.org/10.1177/1367006916654355>
- Li, P., Bates, E., Liu, H., & MacWhinney, B. (1992). Cues as Functional Constraints on Sentence Processing in Chinese. In *Advances in Psychology* (Vol. 90, pp. 207–234). Elsevier. [https://doi.org/10.1016/S0166-4115\(08\)61893-2](https://doi.org/10.1016/S0166-4115(08)61893-2)
- Mai, Z., Kwan, C., & Yip, V. (2018). Expressing displacement in heritage Cantonese: Cross-linguistic influence and structural vulnerability. *International Journal of Bilingualism*, 22(6), 603–618. <https://doi.org/10.1177/1367006916681084>
- Messenger, K., Branigan, H. P., McLean, J. F., & Sorace, A. (2012). Is young children’s passive syntax semantically constrained? Evidence from syntactic priming. *Journal of Memory and Language*, 66(4), 568–587. <https://doi.org/10.1016/j.jml.2012.03.008>
- Montrul, S., & Polinsky, M. (Eds.). (2021). Grammatical Aspects of Heritage Languages. In *The Cambridge Handbook of Heritage Languages and Linguistics* (1st ed., pp. 579–758). Cambridge University Press. <https://doi.org/10.1017/9781108766340.026>
- Montrul, S., & Sánchez-Walker, N. (2013). Differential Object Marking in Child and Adult Spanish Heritage Speakers. *Language Acquisition*, 20(2), 109–132. <https://doi.org/10.1080/10489223.2013.766741>
- Müller, N., & Hulk, A. (2001). Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition*, 4(1), 1–21. <https://doi.org/10.1017/S1366728901000116>
- Ortega, L. (2020). The Study of Heritage Language Development From a Bilingualism and Social Justice Perspective. *Language Learning*, 70(S1), 15–53. <https://doi.org/10.1111/lang.12347>
- Paradis, J. (2023). Sources of individual differences in the dual language development of heritage bilinguals. *Journal of Child Language*, 1–25. <https://doi.org/10.1017/S0305000922000708>

- Polinsky, M., & Scontras, G. (2020). Understanding heritage languages. *Bilingualism: Language and Cognition*, 23(1), 4–20. <https://doi.org/10.1017/S1366728919000245>
- Polinsky, M., Zhang, B., & Gallo, C. G. (2010). *Heritage Chinese: A new view from production*. Heritage Language Insititute IV, University of California, Los Angeles.
- R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rothman, J., Bayram, F., DeLuca, V., Di Pisa, G., Duñabeitia, J. A., Gharibi, K., Hao, J., Kolb, N., Kubota, M., Kupisch, T., Laméris, T., Luque, A., van Osch, B., Pereira Soares, S. M., Prystauka, Y., Tat, D., Tomić, A., Voits, T., & Wulff, S. (2022). Monolingual comparative normativity in bilingualism research is out of “control”: Arguments and alternatives. *Applied Psycholinguistics*, 1–14. <https://doi.org/10.1017/S0142716422000315>
- Slabakova, R. (2015). The effect of construction frequency and native transfer on second language knowledge of the syntax–discourse interface. *Applied Psycholinguistics*, 36(3), 671–699. <https://doi.org/10.1017/S0142716413000386>
- van Dijk, C., Dijkstra, T., & Unsworth, S. (2022). Cross-linguistic influence during online sentence processing in bilingual children. *Bilingualism: Language and Cognition*, 25(4), 691–704. <https://doi.org/10.1017/S1366728922000050>
- Van Dijk, C., Van Wonderen, E., Koutamanis, E., Kootstra, G. J., Dijkstra, T., & Unsworth, S. (2021). Cross-linguistic influence in simultaneous and early sequential bilingual children: A meta-analysis. *Journal of Child Language*, 1–33. <https://doi.org/10.1017/S0305000921000337>
- Zhou, P., & Ma, W. (2018). Children’s Use of Morphological Cues in Real-Time Event Representation. *Journal of Psycholinguistic Research*, 47(1), 241–260. <https://doi.org/10.1007/s10936-017-9530-y>