



ARTICLE

Sources of children's difficulties with non-canonical sentence structures: Insights from Mandarin

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

¹School of Philosophy, Psychology and Language Sciences, University of Edinburgh, United Kingdom and

²Department of Language and Culture (ISK), UiT The Arctic University of Norway, Norway

Corresponding author: Jiuzhou Hao; Email: jiuzhou.hao@uit.no

(Received 24 January 2023; revised 01 May 2024; accepted 04 July 2024)

Abstract

The present study investigated whether children's difficulty with non-canonical structures is due to their non-adult-like use of linguistic cues or their inability to revise misinterpretations using late-arriving cues. We adopted a priming production task and a self-paced listening task with picture verification, and included three Mandarin non-canonical structures with differing word orders and the presence or absence of morphosyntactic cues. Forty five-to-ten-year-old Mandarin-speaking children were tested and compared to adults. Results showed that children were indistinguishable from adults in how they used different cues in real-time, although their performance in offline comprehension and production was more prone to errors but improved given the increase of age. These results suggest that the current child sample has adult-like cue-use patterns and use late-arriving cues to revise misinterpretations. The observed worse offline accuracy and production difficulties relative to adults result from their less developed domain-general abilities in performing tasks.

Keywords: Mandarin; non-canonical structures; processing; production; priming

Introduction

The CANONICAL WORD ORDER in English is realised as Subject¹-Verb-Object as in (1). Structures that differ from this canonical word order are considered NON-CANONICAL, e.g., passives in English (2). Across typologically different languages, non-canonical structures, as opposed to canonical ones, have been consistently shown to take a significantly longer time for children to acquire (e.g., Borer & Wexler, 1987; Demuth, 1989; Dittmar et al., 2008). This is manifested in both production and comprehension (Borer & Wexler, 1992; Brooks & Tomasello, 1999; Huang et al., 2013; Messenger et al., 2012). However, the reasons for this remain subject to theoretical debate.

¹In this article, subjects/objects refer to logical subjects/objects.

- (1) The police officer was chasing the thief.
- (2) The thief was chased by the police officer.

Previous research, attempting to understand if and why non-canonical structures are difficult for children, has mainly compared the production and/or the comprehension of one non-canonical structure to that of the canonical structure, e.g., passives vs. actives in English. Although informative, such a design cannot directly tease apart the role of word order and the presence or absence of (late-arriving) MORPHOSYNTACTIC CUES². This is because the two structures differ from each other in both regards. Specifically, in canonical structures, e.g., English actives, agents are realised before patients. Additionally, there is no morphosyntactic cue assisting the mapping of syntactic functions to thematic roles; instead, the canonical word order cue that the first Noun Phrase (NP1) is the agent dictates such mappings. In contrast, in non-canonical structures, e.g., English passives, patients are displaced to a position before agents. In these structures, morphosyntactic cues, e.g., *-ed*, *-by*, are typically available while the canonical word order cue that NP1s are agents gives rise to erroneous interpretations of non-canonical structures. Crucially, the canonical word order cue appears early on relative to the morphosyntactic cues within non-canonical structures.

The current study attempts to shed light on the role of word order and the presence and absence of morphosyntactic cues by examining how Mandarin-learning children produce and comprehend three Mandarin non-canonical structures, i.e., the BA-CONSTRUCTION, the BEI-CONSTRUCTION, and the OSV-CONSTRUCTION. These three structures have different word orders and are different in the presence vs. absence of morphosyntactic cues. Furthermore, we adopt both production and online and offline comprehension tasks to test the same child population. The inclusion of an online comprehension task is critical because it allows us to understand how children use different linguistic cues in real-time when processing non-canonical structures.

Acquisition accounts of non-canonical structures

The Competition Model

The (Unified) COMPETITION MODEL (MacWhinney, 1987, 2018) attributes children's difficulties with non-canonical structures to their non-adult like reliance on various (non-)linguistic cues and distinguishes between CUE AVAILABILITY and CUE RELIABILITY. The former refers to how often a cue occurs and the latter to how likely the cue leads to correct interpretations of a sentence it occurs in. According to this model, children rely heavily on the most available cue(s) regardless of the reliability of the cue(s). Instead, adults rely on the more reliable ones. As children get older (around the age of eight to ten), they converge on adult strategies, relying on the most reliable cue. In the case of English passives vs. actives, for example, the canonical word order cue is highly available because of the high frequency of actives in general. In contrast, the morphosyntactic cues in passives, i.e., *-ed* and *-by*, are less frequent, although they are more reliable in

²There are also theories attributing children's difficulties with non-canonical word order to the syntactic operations through which non-canonical structures are derived. Due to the age of the children in the study, the lack of clear description of the related structures in Mandarin and concerns about the suitability and robustness of the predictions of such accounts, we decided not to entertain this set of accounts.

indicating that the sentence is a passive. Therefore, according to the model, children prefer the canonical word order cue even when comprehending non-canonical structures, which typically have non-agent subjects. In contrast, the less available yet more reliable cues, such as the *-ed* and the *by-phrase* in English passives, are ignored during comprehension (Bever, 2013) or dropped during production by children (Slobin, 1973). Additionally, the model also indicates that children's reaction times gradually get faster and converge with adults eventually.

The Incremental Processing account

The INCREMENTAL PROCESSING ACCOUNT postulates that it is children's inability to reanalyse or inhibit the (initial) (mis-)interpretations drawn from cues that come early in the sentences that causes their difficulties with non-canonical structures (Huang et al., 2013; Trueswell & Gleitman, 2004). For example, in processing English passives, while both children and adults initially interpret NP1s as agents using the canonical word order cue (an early-arriving cue), adults do revise this interpretation immediately when the passive-related morphosyntactic cues are available later in the input. In contrast, children commit to their initial interpretation. However, the question of when children converge on adult-like reanalysis ability is not clear. Previous empirical evidence suggests that children have great difficulties revising their initial interpretations until the age of seven (see Choi & Mazuka, 2003; Clackson et al., 2011; Trueswell & Gleitman, 2004, among others). Additionally, as this account focuses on how children use different cues during real-time sentence processing, it only (directly) applies to comprehension. However, it highlights the importance of understanding the use of cues in real time.

Mandarin non-canonical structures and their acquisition and processing

As in English, the canonical word order in Mandarin is also Subject-Verb-Object (the SVO-construction; see example 3). The three Mandarin non-canonical structures tested in the study, i.e., the *BA*-construction (4), the *BEI*-construction (5), and the *OSV*-construction (6), differ from the *SVO*-construction because they all have the second NP (NP2) pre-posed from the post-verbal position in the *SVO*-construction. In terms of phrasal combination, they are all Noun-Noun-Verb (NNV) contrasting to the NVN combination in the *SVO*-construction.

- (3) Zhangsan ti-le Lisi yi-xia.
 Zhangsan kick-ASP Lisi one-CL
 'Zhangsan kicked Lisi once'
- (4) Zhangsan ba Lisi ti-le yi-xia.
 Zhangsan BA Lisi kick-ASP one-CL
 'Zhangsan kicked Lisi once.'
- (5) Lisi bei Zhangsan ti-le yi-xia.
 Lisi BEI Zhangsan kick-ASP one-CL
 'Lisi was kicked by Zhangsan once'

- (6) Lisi, Zhangsan ti-le yi-xia.
 Lisi, Zhangsan kick-ASP one-CL
 'Lisi was kicked by Zhangsan once'

Critically for the purpose of this study, the three non-canonical structures differ from each other in word order, the presence or absence of morphosyntactic cue, or both. While the pre-posed object still follows the subject in *BA*-constructions, it precedes the subject in *BEI*- and *OSV*-constructions. Therefore, the word order in *BA*-constructions is essentially SOV, but it is OSV in *BEI*- and *OSV*-constructions. This has implications for the applicability of the agent-first canonical word order cue. While the canonical word order cue alone could give correct interpretations of who the doer of the action is for the *BA*-construction, it gives reversed interpretations for the *BEI*- and the *OSV*-constructions. Additionally, while the two NPs in *BA*- and *BEI*-constructions are separated by the morphosyntactic cues *ba* and *bei* respectively, *OSV*-constructions do not require a morphosyntactic cue in-between the two NPs. Note that although we separated the two NPs in example (6), a pause after the NP1 in naturalistic production is not necessary. Li et al. (1992) postulated that when two NPs co-occur without any other cues (as in *OSV*-constructions), the word order cue that NP2s function as the agent (NP2 cue/*OSV* cue) is at play to a certain degree (60% of the time for adults). As for the availabilities of these cues (see Li et al., 1992, for a summary), because the canonical word order in Mandarin is SVO, the agent-first canonical word order cue is more frequent than the *ba* and *bei* cues. Comparing the two morphosyntactic cues, given the relatively higher frequency of *BA*-constructions overall, the *ba* cue is of higher availability than the *bei* cue. *OSV*-constructions do not bear any morphosyntactic cue and are of the lowest frequency among the three non-canonical structures – NP2 cue has the lowest availability. In terms of cue reliabilities, the results from Li et al., (1992) with Mandarin-speaking adults suggest that the *bei* cue has the highest reliability followed by word order cues, and then the *ba* cue. It is worth noting here that word order cues in Li et al., (1992) were collapsed between the canonical word order cue and the NP2 cue. However, the fact that the *bei* cue has higher reliability than the *ba* cue allows us to make predictions for the Competition Model, as we detail below in the relevant section.

The acquisition and processing of Mandarin non-canonical structures

In terms of the acquisition and/or processing of these three structures, empirical studies are very limited in number and have received mixed results, crucially in terms of whether the three structures develop symmetrically. For instance, both Zhou and Ma (2018) and Huang et al. (2013) tested the online comprehension of *BA*- and *BEI*-constructions using the visual world paradigm in children aged three to five and five respectively; Hu et al. (2018) investigated the offline comprehension of *OSV*-constructions with and without a topicalisation cue (the *a* cue) in children aged three to five, and Deng et al. (2018) examined the naturalistic production of *BA*- and *BEI*-constructions in children around the age of two. Hao and Chondrogianni (2021) is the only previous study that has tested the offline comprehension and production of all the three structures in children aged above five (to nine).

In more detail, Zhou and Ma (2018) auditorily presented *BA*-constructions (7) and *BEI*-constructions (8) with omitted NP1s to three- and five-year-old children. For online processing, the authors measured children's eye-movements while listening; for offline

comprehension, children were asked to select the picture matching the sentence they heard. The results showed that both three- and five-year-old children directed their eye-gaze to the target, i.e., the patient and the agent in *BA*- and *BEI*-constructions, respectively, after hearing the morphosyntactic cues with the five-year-olds performing adult-like while the three-year-olds showed a quantitative delay caused by, as argued by the authors, immature cognitive ability. As for offline comprehension, children, regardless of their age, used both *ba* and *bei* to derive correct interpretations to the same degree.

(7) Ba shizi qingqingde bao-le qilai
 BA lion gently hold-ASP up
 'Something gently held the lion.'

(8) Bei shizi qingqingde bao-le qilai
 BEI lion gently hold-ASP up
 'Something was gently held by the lion.'

Huang et al. (2013) manipulated the forms of NP1s in *BA*- and *BEI*-constructions to test five-year-olds' online processing of these structures using eye-tracking, and offline comprehension with an act-out task. Participants were shown three thematically related objects, e.g., a seal (an expressed item), a shark (a likely agent) and a fish (a likely theme) while listening to one of the following four sentences, after which they were asked to act out what they had just heard. In example (9) and (10), NP1s were expressed while NP2s were pronouns. The rationale was that if participants could use the morphosyntactic cues in processing, they should look at the likely theme (fish) or the likely agent (shark) respectively in *BA*- and *BEI*-constructions after hearing the pronoun NP2. In pronoun NP1 conditions (example 11 and 12), however, eye-movements directed to the likely agent in *BA*-constructions or theme in *BEI*-constructions would be evidence for the successful use of the *ba* and *bei* cues.

(9) Haibao ba ta henkuai-jiu chidiao-le.
 Seal BA it quickly eat-ASP
 'The seal is quickly eating it.'

(10) Haibao bei ta henkuai-jiu chidiao-le.
 Seal BEI it quickly eat-ASP
 'The seal is quickly eaten by it.'

(11) Ta ba haibao henkuai-jiu chidiao-le.
 It BA seal quickly eat-ASP
 'It is quickly eating the seal.'

(12) Ta bei haibao henkuai-jiu chidiao-le.
 It BEI seal quickly eat-ASP
 'It is quickly eaten by the seal.'

The results showed that for the five-year-olds in the study, the passive-related morphology *bei* was overall more vulnerable and prone to interference than *ba* in offline comprehension. However, eye-movements suggested successful use of both *ba* and *bei*, contradicting to what Zhou and Ma (2018) found. The differential results might only

reflect the differences in experimental design. Specifically, the task adopted by Huang et al. (2013) might be too complex and require more than syntactic processing. For example, the task adopted the use of pronominal forms which might not be fully developed yet in these children. Additionally, more than two referents were presented visually in the scene when the sentences were only dealing with two entities. This might have been problematic because children do not have fully developed executive functions, e.g., inhibitory ability in particular (Novick et al., 2008; Trueswell et al., 1999), and children are more prone to interference in cue-based retrieval than adults are (Omaki et al., 2014). Meanwhile, because Zhou and Ma (2018) adopted stimuli without NP1s, the sentences might be easier to comprehend simply because there was no apparent need to establish the syntactic and the thematic relations of the omitted NP1s.

Interestingly, differential performance between *BA*- and *BEI*-constructions was found in naturalistic production as well by Deng et al. (2018), however in a different direction, i.e., better performance on *BEI*-constructions compared to *BA*-constructions. Deng et al. (2018) analysed the naturalistic production of *BA*- and *BEI*-constructions in children aged around two, showing that these structures developed quite early and were not modulated by input frequency; at the very least, these children's produced *BA*- and *BEI*-constructions following adult-like constraints. Overall, the naturalistic production of *BA*- and *BEI*-constructions happened around the age of two years, with *BEI*-constructions (0.02%) produced two months earlier than *BA*-constructions (1.27%), even though the input frequency of *BA*-constructions (2.62%) was significantly higher than *BEI*-constructions (0.13%).

Turning to *OSV*-constructions, Hu et al. (2018) tested the offline comprehension of *OSV*-structures with and without a topic cue (not the *bei* cue) in children aged three to five. Using a picture selection task, they found that the comprehension of object topicalisation without a topic cue (*OSV*-constructions in the current study) was well above chance at the age of three and reached ceiling at the age of five to six. Nonetheless, it is unclear if *OSV*-constructions would induce better/worse performance than *BEI*- or *BA*-constructions. Hao and Chondrogianni (2021) found a numerical disadvantage (although not statistically significant) in offline comprehension (picture selection) and production (priming) of *OSV*-constructions as opposed to *BA*- and *BEI*-constructions in children aged five to nine. Furthermore, the same population showed indistinguishable performance between *BA*- and *BEI*-constructions. This supports the findings of Zhou and Ma (2018) but contradicts the findings of Huang et al. (2013) and Deng et al. (2018). Another important finding in Hao and Chondrogianni (2021) was that although children reached ceiling in both comprehension and production and were adult-like as a group, they did observe a developmental effect for all the three structures where older children perform better.

In sum, limited empirical studies have shown that children produce *BA*- and *BEI*-constructions in spontaneous speech as early as two years of age (Deng et al., 2018), that they show qualitatively adult-like online processing of *BA*- and *BEI*-constructions from the age of three (Huang et al., 2013; Zhou & Ma, 2018), and that they are indistinguishable from adults in production and offline comprehension of *BA*-, *BEI*- and *OSV*-constructions from the age of five (Hao & Chondrogianni, 2021). However, there have been mixed results on whether the three structures develop symmetrically. Indeed, children's potentially differential performance as a function of structural properties provides critical evidence for/against different theoretical accounts as we will elaborate below. Furthermore, if and how children make use of different cues in processing non-canonical structures with both NPs fully realised remains to be understood, especially when they process the *OSV*-construction.

The present study

This study set out to examine the development and processing of non-canonical structures in Mandarin-speaking children aged from five to ten. Importantly, we tested the same children's production, as well as online and offline comprehension of three Mandarin non-canonical structures, i.e., *BA*-, *BEI*-, and *OSV*-constructions, that differ from each other in word order and the presence or absence of morphosyntactic cues. We targeted children between the ages of five- and ten-years-old. This covers the range when children are said to converge with adults on similar constructions in other languages, which includes 8- to 10-year-olds (the Competition Model) and 7-year-olds and older (the Incremental Processing Account). A comprehension-to-production priming task was adopted to test children's production of these structures, which would otherwise be infrequent in naturalistic production (Deng et al., 2018). Apart from examining production accuracy and error types, the priming paradigm also allows us to understand and make predictions about the abstract structural representations of these structures. This is because we take priming to index the short-term (re)activation of abstract structural representations (e.g., Branigan & Pickering, 2017); consequently, the priming magnitude indexes the ease of accessing/activating the representation – smaller priming indexes more production difficulties (see also Hao et al., 2023). Priming magnitude is also linked to accuracy, in that structures that are primed more are also more accurate. Finally, a self-paced listening task with picture verification (see also Marinis & Saddy, 2013) was adopted to examine both online and offline comprehension. An adult control group was tested as well because, to our knowledge there is no previous study that tested the production and comprehension of these structures in adults.

We asked the following research questions (RQs):

RQ1: Are the three non-canonical structures equally difficult or easy for children (and adults) to produce or be primed?

Looking at accuracy first, the Competition Model (MacWhinney, 1987, 2018) makes direct predictions about what structures would be more accurate for children to produce, and the types of errors that children will make. Specifically, the Competition Model predicts that children will rely mostly on cues that are higher in AVAILABILITY regardless of their reliability. Therefore, *BEI*- and *OSV*-constructions will be harder than *BA*-constructions because NPs in *BA*-constructions denote the agent which follows the canonical word order cue – a cue with higher availability. At the same time, the *BEI*-construction will induce higher production rates than the *OSV*-construction because of the presence of the *bei* cue and its relatively higher frequency than the *OSV*-construction (Li et al., 1992). When children do not produce the target structures, the Competition Model predicts a higher likelihood of producing the *SVO*-construction overall and reversal errors especially when the *BEI*- and *OSV*-constructions are the target structures. The latter is because NPs can be misinterpreted as the agent if the relevant (morphosyntactic) cues are not considered in *BEI*- and *OSV*-constructions. In addition, the Competition Model also predicts the omission of morphosyntactic cues, i.e., *ba* and, especially, *bei* (see Slobin, 1973). Note that the Incremental Processing Account does not make any predictions about production. In terms of the priming magnitude, we assume that structures that are easier to be (re-)activated should give rise to higher priming effects. The relative ease of reactivation may depend on the relative word order and the morphosyntactic cues, although it should be noted that these accounts do not make any predictions about priming.

Q2: Are the different (non-)canonical structures equally difficult or easy to comprehend offline and process in real-time for children (and adults)?

Both accounts predict children to have better offline comprehension of the *BA*-construction relative to the *BEI*- and *OSV*-constructions, due to the high availability of the canonical word order cue and the relatively higher availability of *ba* over *bei* (Competition Model), or the early position of the canonical word order cue (Incremental Processing Account). As for differences between *BEI*- and *OSV*-constructions, the Competition Model predicts better comprehension accuracy for *BEI*-constructions over *OSV*-constructions because of the presence of the *bei* cue (and its relatively higher frequency compared to *OSV*-constructions). In contrast, the Incremental Processing Account would predict similar offline comprehension accuracy between the two. This is because the early-arriving canonical word order cue would lead to erroneous interpretations for both *BEI*- and *OSV*-constructions that children cannot revise. Turning to online processing, the Competition Model predicts that children will use these cues to different degrees given their different availability ($ba > bei > OSV$). In contrast, it is not clear if the Incremental Processing Account would predict differential use of (late-arriving) cues. Although previous online processing studies showed that children do make use of different cues interactively to revise misinterpretations (see Snedeker & Huang, 2015 for a summary of evidence), these studies focused on non-(morpho)syntactic cues, e.g., prosody, verb semantics, etc., it is not clear if and how children make use of different cues within the morphosyntactic proper, a contribution of the present study.

RQ3: What is the role of chronological age in children's production and comprehension of non-canonical structures? And is any developmental effect modulated by structure type?

We expect developmental effects to surface across structures in production and offline comprehension (Hao & Chondrogianni, 2021) and in online comprehension (Zhou & Ma, 2018). Although both the Competition Model and the Incremental Processing Account would expect developmental effects, their predictions diverge in that they predict potentially differential developmental rates for the three structures. Specifically, according to the Competition Model, the cue that is relatively more RELIABLE (as opposed to more available) might develop faster, i.e., *bei* faster than *ba*. On the other hand, as the Incremental Processing Account attributes children's difficulties with reanalysis to their limited processing resources, which are expected to develop with age, it predicts similar developmental rates for structures that require reanalysis. These would be the *BEI*- and the *OSV*-constructions.

Methodology

Participants

A total of 40 children participated in the study. Five were excluded either because of withdrawal (1), partial data loss (1), incomplete procedure (2), or below-chance filler comprehension (1), leaving 35 participants for further analysis as a group (typically-developing monolingual children; CHI) (Mean age = 85.20 months, SD = 16.63, Range = 60 - 111). Thirty adult Mandarin speakers with a mean age of 25.73 years (SD = 4.20, Range = 19 - 36) participated in the study as well and were all included in further analysis,

constituting the adult group (ADT). Among them, 23 were monolingual Mandarin speakers residing in China, while the remaining seven participants were Mandarin-dominant bilingual speakers studying in Edinburgh with less than eight months of naturalistic exposure to English. None of the participants in either group had a history of speech and/or language delay or impairment or other developmental disorders.

Production task

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 (prime), after which, participants were asked to describe a new picture shown on the screen (target) and were instructed to describe it as quickly as possible.

Each of the three non-canonical structures served as a prime fifteen times, making a total of 45 prime sentences. For the primes, we selected five transitive verbs, i.e., *tui* 'push', *yao* 'bite', *ti* 'kick', *qin* 'kiss', and *ju* 'raise', with each appearing three times per prime type. As for the targets, three new (transitive) verbs, i.e., *zhui* 'chase', *xi* 'clean', and *wei* 'feed', were selected along with *tui* 'push' and *ti* 'kick' which were also used in the primes. Importantly, we manipulated the distributions of the NPs and VPs so that there was no overlap in lexical items between the primes and their respective targets to level out item-based priming effects (Tomasello, 2000).

All primes shared the same structure: Noun Phrase (NP) + morphosyntactic cue *ba*, *bei*, or null + NP + Adverb + Verb Phrase (VP). Furthermore, the NPs (N = 45) 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). Additionally, all NPs were selected to be frequent and familiar for Mandarin-learning children aged 5 to 10 (see also Hao & Chondrogianni, 2021). 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.

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). Additionally, to avoid any repetition or order effects, we made three separate lists (see the OSF page for the lists), such that each prime picture was depicted with all three structures and each depiction for the same prime picture appeared in only one list. For instance, (13), (14), and (15) are the *BA*-, *BEI*- and *OSV*-primes for the prime picture (Figure 1), and they were arranged into list A, B and C respectively.

- (13) yi-zhi laoshu ba yi-zhi wugui tui-le yixia.
 one-CL mouse BA one-CL turtle push-PERF once
 'A mouse pushed a turtle.'
- (14) yi-zhi wugui bei yi-zhi laoshu tui-le yixia.
 one-CL turtle BEI one-CL mouse push-PERF once
 'A turtle was pushed by a mouse.'

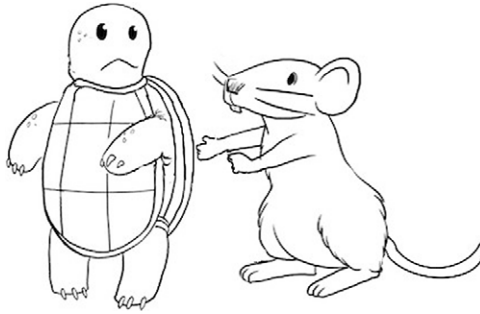


Figure 1. Example of a prime picture (experimental trials).

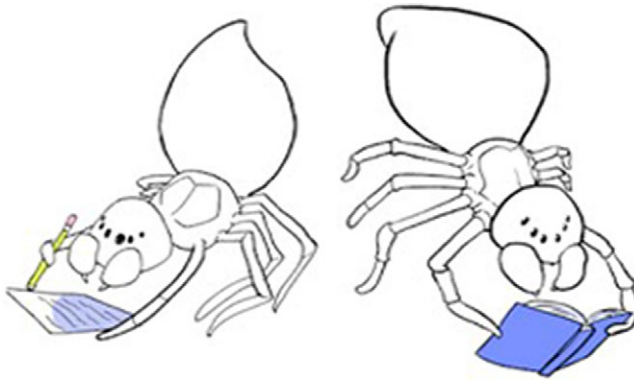


Figure 2. Example of a prime picture (filler trials).

- (15) yi-zhi wugui, yi-zhi laoshu tui-le yixia.
one-CL turtle, one-CL mouse push-PERF once
'A turtle, a mouse pushed.'

The task also included 20 fillers. 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 (16) (Figure 2). Targets for fillers also involved pictures with two animals performing an intransitive action.

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

Additionally, all trials were arranged in a pseudorandom order where trials from the same experimental condition did not appear consecutively.

Comprehension task

A self-paced listening task with picture verification was administered. 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

Table 1. Experimental conditions for the comprehension task

Condition	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
BA-match	<i>Wokanjia</i> 'I saw'	<i>yizhi shanyang</i> 'one-CL goat'	<i>BA yizhi laolang</i> 'BA one-CL wolf'	<i>qingqingde</i> 'gently'	<i>tile yixia</i> 'kick-PERF once'
BA-mismatch	<i>Wokanjia</i> 'I saw'	<i>yizhi laolang</i> 'one-CL wolf'	<i>BA yizhi shanyang</i> 'BA one-CL goat'	<i>qingqingde</i> 'gently'	<i>tile yixia</i> 'kick-PERF once'
BEI-match	<i>Wokanjia</i> 'I saw'	<i>yizhi laolang</i> 'one-CL wolf'	<i>BEI yizhi shanyang</i> 'BEI one-CL goat'	<i>qingqingde</i> 'gently'	<i>tile yixia</i> 'kick-PERF once'
BEI-mismatch	<i>Wokanjia</i> 'I saw'	<i>yizhi shanyang</i> 'one-CL goat'	<i>BEI yizhi laolang</i> 'BEI one-CL wolf'	<i>qingqingde</i> 'gently'	<i>tile yixia</i> 'kick-PERF once'
OSV-match	<i>Wokanjia</i> 'I saw'	<i>yizhi laolang</i> 'one-CL wolf'	<i>yizhi shanyang</i> 'one-CL goat'	<i>qingqingde</i> 'gently'	<i>tile yixia</i> 'kick-PERF once'
OSV-mismatch	<i>Wokanjia</i> 'I saw'	<i>yizhi shanyang</i> 'one-CL goat'	<i>yizhi laolang</i> 'one-CL wolf'	<i>qingqingde</i> 'gently'	<i>tile yixia</i> 'kick-PERF once'

to examine the online processing of *BA*-, *BEI*- and *OSV*-constructions. The rationale was that if participants could use a particular cue, elevated reaction times (RTs) reflecting reanalysis processes in online processing and worse accuracy performance in offline comprehension should be observed when there was such a mismatch. 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 used 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 and was ensured to appear in each condition. This gave rise to 48 experimental trials. 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. In addition, to further limit participant's potential use of information other than those we were interested in, e.g., animacy, world knowledge, etc., 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).

All six conditions for any given item were presented with the same picture. For instance, Figure 3, showing the agent, a sheep, kicking the patient, a wolf, is depicted

**Figure 3.** Example of pictures for experimental trials in the comprehension task.

six times and each of the descriptions was sorted into one of the six conditions, as in Table 1.

Similar to the production task, there were also 20 fillers included in the comprehension task, with half adapted from the production task and 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 (17). And 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).

- (17) Wo kanjian/ yi-zhi zhizhu/ kanshu/, yi-zhi zhizhu/ xiexin
 I see/ one-CL spider/ read/ one-CL spider/ write
 ‘I saw that a spider is reading; a spider is writing.’

Six separate lists (see the OSF page for the lists) were constructed to make sure that any given condition of the same item appeared only once in any given list, and across all lists, all conditions of all items were represented. Participants were pseudorandomly assigned to different lists and presented with a full list in a within-subject design. The relative position of the agent/patient in the pictures was also counterbalanced to ensure that half of the 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. Standardised Mandarin was chosen as it is the educational language used in Mainland China and is perfectly intelligible to speakers across different regions. 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 would be played, and the participants were then asked to judge if the sentence they heard matched the picture as fast as they could. This could ensure that the participants were actively comprehending the sentences instead of pressing buttons only for finishing the task and gave an off-line comprehension accuracy measurement. Participants did not receive any accuracy feedback throughout the experiment.

Procedure

All participants took part in the study at their homes over the web. We implemented the experimental tasks with JsPsych (de Leeuw, 2015) in a webpage. Each participant participated in all the experimental tasks, and the entire session lasted approximately 50 to 70 minutes depending on the participants’ age. The presentation of the experimental tasks was counterbalanced to cancel out potential carry-over effects between tasks, such that the comprehension task was administered first to a random half of the participants and the rest were firstly tested with the production task. The whole process of the experiment for each participant was audio-recorded. All the responses in the production task 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 page to give consent for their participation. The study has been approved by the institutional ethics committee.

Coding and scoring

For the production task, participants' responses were coded as either "BA", "BEI", "OSV" or "SVO" when the produced utterance was complete, described the targeted action, and carried correct thematic roles. If the utterance was complete and depicted correct actions but carried reversed thematic roles, a coding of "Reversed" would be given. Incomplete or unidentifiable utterances, utterances with incorrect action depicted, utterances failing to establish who did what to whom, e.g., separately describing intransitive actions for the two animals involved in the picture (see example 18), etc., and responses with NP omissions when no morphosyntactic cue is used (see example 19) were coded as "Other" and were excluded from further analysis.

(18) 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.'

(19) yi-zhi laoshu zhuigan-zhe
 one-CL mouse chase-CONT
 'A mouse is chasing after something.' (Omission of the object: omitted NP1 in OSV-constructions or omitted NP2 in SVO-constructions)

'Something is chasing after a mouse.' (Omission of the subject: omitted NP2 in OSV-constructions)

Note that although the same utterance as Example (19) would be coded as a passive in Leonard et al. (2006)'s study with Cantonese children, we did not follow that approach, nor did we code such instances as "OSV". This is because both subject and object could be omitted freely in Mandarin, especially when the entity depicted is highly recoverable from the context or is in the common ground between interlocutors. Therefore, without the presence of a morphosyntactic cue, the realised NP in example (19) could be interpreted as the subject in SVO- or OSV-constructions as well as the object in OSV-constructions. If we had followed Leonard et al. (2006), we might have artificially overestimated children's ability in producing OSV-constructions when in fact they are still in the phase of preferring SVO-constructions.

For the comprehension task, all responses were firstly coded as "1" or "0" respectively, when the participant gave a correct or incorrect response to the judgement of whether the sentence matched the picture. We then selected participants with an above chance level (50%) accuracy in the filler trials for further analyses of the experimental items. Secondly, for the RT analyses, we only included the trials where participants gave a correct response to the picture verification task. In trimming the RT data, we excluded extreme values below 500ms or above 5000ms after checking the distribution of the data and outliers that were below or above two standard deviations of the mean calculated for each structure per participant and condition. Then, we converted raw RTs to residual RTs (the differences between the predicted values and the actual value) to control for the differences in length

across trials and segments, as well as individual differences in responding to different stimuli. Specifically, the predicted values were calculated for each participant and trial based on the length of each segment using linear mixed effect regressions. Residual RTs were used in further analyses and visualisations of RT data.

Data analysis

Statistical analyses were carried out with the *lme4* package and the *mlogit* 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.

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, if it was a logit model, we firstly ran an optimiser selection process with the *afex* package, and if the model still did not converge or it was not a logit model, we iteratively removed the random effect accounting for the least variance until convergence was achieved. To identify the optimal model, we adopted the stepwise backward selection approach starting from the maximal model.

For production data analyses, we opted for multinomial regressions because they would give a detailed picture of how different structures were used (production patterns) across prime types instead of focusing on the use of one particular structure only. For the post hoc analyses, given the nature of multinomial models, we first exhausted all possible combinations of reference levels for all variables (dependent and independent) and then conducted analyses with reduced models when any significant interactions were attested. Specifically, to unpack any interactions, we followed Mangiafico (2016) to break the full model into reduced models.

Results

Production task

A total of 1575 utterances were produced by the CHI group. Within all the responses, 94 responses (6%) were lost due to recording issues or noises and 268 (17%) were scored as “Others”, and hence were all excluded for further analysis. On the other hand, the ADT group produced 1350 responses altogether, with 196 (15%) coded as “Others” and excluded.

Figure (4) shows the response structure distribution across prime types in the ADT and CHI groups. Note that “Other” responses were excluded. Overall, both groups showed the tendency to reuse the prime structures in their own production, e.g., the likelihood of producing *BEI*-constructions was the highest after *BEI*-primes as opposed to after *BA*- and/or *OSV*-primes. However, the magnitude of priming was weaker in the CHI group for all structures. As for production patterns, when not using the prime structures, both groups were more likely to produce *SVO*-constructions after *BA*-primes than after *BEI*- and/or *OSV*-primes, with the CHI group doing so to a larger extent. On the other hand, the CHI group also opted for *BA*-constructions after *BEI*- and *OSV*-primes more often than the ADT group did.

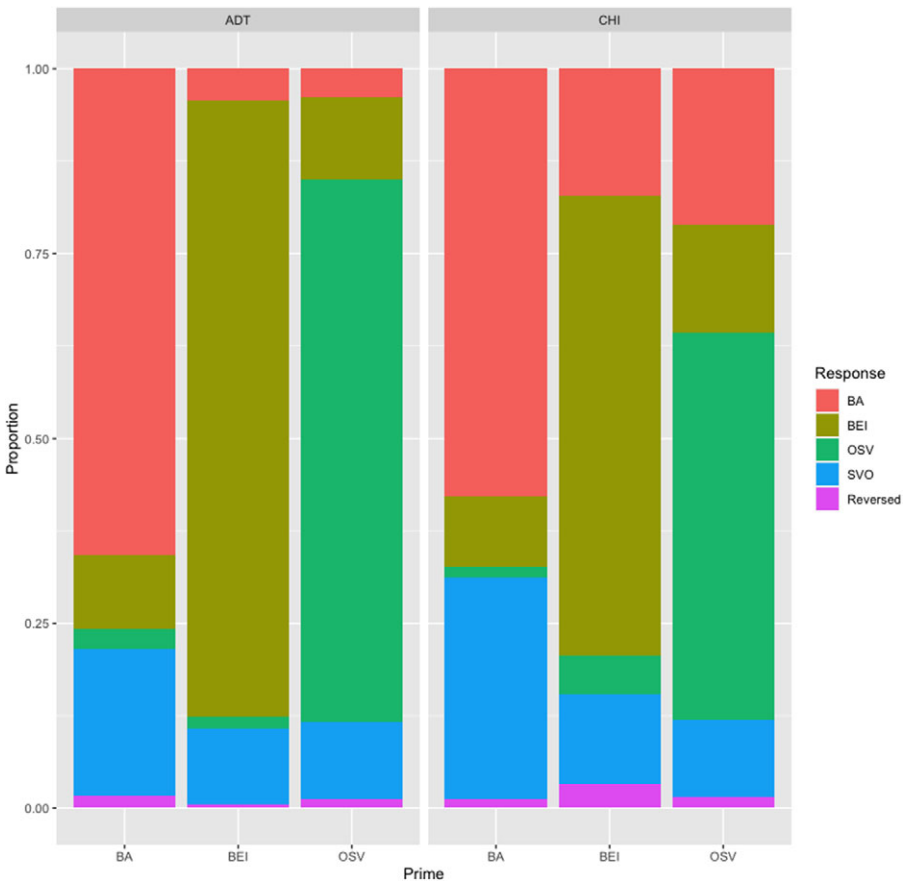


Figure 4. Proportion of response types following different prime types in the CHI group and the ADT group.

To statistically account for group differences/similarities and how they were modulated by structure types (RQ 1), we ran multinomial logistic regression models with all valid responses (BA, **BEI**³, OSV, SVO and Reversed) included in the dependant variable. As fixed effects, Group (ADT and CHI) and Prime Type (BA, **BEI**, and OSV) were entered. *BEI*-constructions were chosen as the reference level for both the dependant and independent variables to allow direct calculation of priming, i.e., the production of *BEI*-constructions after *BEI*-primes. The optimal model included the interaction terms between the two independent variables (see Table (2)). However, interaction terms between the two independent variables were not all presented in the table although some of them were significant and were selected in the optimal model. This is because all our independent variables are categorical in nature, any comparisons with the reference level, i.e., the use of *BEI*-construction after a *BEI*-prime in the ADT group, aren't meaningful for

³Unless specified, bold-faced levels are chosen as the reference level for the variable.

Table 2. Optimal model with Group (ADT and CHI) and Prime Type (BA, BEI, and OSV) as fixed effects for all valid responses in the production task

Response Type		Estimate	Standard Error	b	p
BEI vs BA	Intercept	-2.97	.25	-11.94***	<.001
	Group ADT vs CHI	1.68	.29	5.81***	<.001
	Prime Type BEI vs BA	4.85	.30	15.90***	<.001
	BEI vs OSV	1.92	.39	4.93***	<.001
BEI vs OSV	Intercept	-4.01	.41	-9.74***	<.001
	Group ADT vs CHI	1.53	.48	3.19**	<.01
	Prime Type BEI vs BA	2.71	.55	4.97***	<.001
	BEI vs OSV	5.90	.44	13.32***	<.001
BEI vs SVO	Intercept	-2.09	.17	-12.64***	<.001
	Group ADT vs CHI	0.46	.23	1.94	.05
	Prime Type BEI vs BA	2.77	.26	10.62***	<.001
	BEI vs OSV	2.02	.27	7.34***	<.001
BEI vs Reversed	Intercept	-5.11	.71	-7.21***	<.001
	Group ADT vs CHI	2.14	.77	2.76**	<.01
	Prime Type BEI vs BA	3.29	.83	3.94***	<.001
	BEI vs OSV	2.96	.85	3.47***	<.001

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

our interest, e.g., consider comparing the reference level to the use of *BEI*-construction after an *OSV*-prime in the CHI group.

The optimal model showed that for the priming effect: (1) both groups produced more *BEI*-constructions after *BEI*-primes than after *BA*- and/or *OSV*-primes; (2) the ADT group showed stronger priming than the CHI group after *BEI*-primes; for production patterns: when not primed and compared with the ADT group, the CHI group produced more (1) *BA*-responses (2) *OSV*-responses and (3) reversal errors. Nonetheless, the two groups produced a similar number of *SVO*-responses after *BEI*-primes.

Post hoc analyses with re-levelled reference levels and reduced models further suggested that (1) priming surfaced in both groups but was stronger in the ADT group for all structures; (2) for both groups, priming after *BEI*-primes was stronger than after *OSV*- and/or *BA*-primes which did not differ; (3) group differences in production patterns were modulated by structure types. Specifically, after *BA*-primes, the two groups only differed in that the CHI produced more *SVO*-responses. On the other hand and similar to *BEI*-prime conditions, after *OSV*-primes, the CHI group produced (1) more *BA*-responses and (2) more *BEI*-responses but (3) a similar number of *SVO*-responses or (4) reversal errors, compared with the ADT group. Additionally, children's likelihood of producing *BA*-constructions after *BEI*- or *OSV*-primes was higher than that of producing other non-primed structures.

Finally, to answer RQ 3 about any potential developmental effects in production, we ran another multinomial logistic regression with Age (in months as a continuous variable;

Table 3. Model summary with (BA, BEI, and OSV) and Age (scaled) as fixed effects for all valid responses in the production task for the CHI group

Response Type		Estimate	Standard Error	b	p
BEI vs BA	Intercept	-1.29	.15	-8.67***	<.001
	Prime Type BEI vs BA	3.09	.24	12.89***	<.001
	BEI vs OSV	3.90	.33	11.97***	<.001
	Age (Scaled)	-0.32	.15	-2.13*	.03
BEI vs OSV	Intercept	-2.62	.28	-9.31***	<.001
	Prime Type BEI vs OSV	3.74	.29	12.77**	<.001
	Age (Scaled)	-0.74	.27	-2.75**	<.01
	Prime Type*Age (Scaled) BEI*Age vs OSV*Age	1.20	.32	3.79***	<.001
BEI vs Reversed	Intercept	-3.86	.62	-6.28***	<.001
	Prime Type BEI vs OSV	1.59	.77	2.04*	.04
	Age (Scaled)	-1.64	.51	-3.22**	<.01
	Prime Type*Age (Scaled) BEI*Age vs OSV*Age	1.97	.70	2.82**	<.01
BEI vs SVO	Intercept	-1.66	.18	-9.48***	<.001
	Prime Type BEI vs BA	2.81	.27	10.56***	<.001
	BEI vs OSV	1.35	.28	4.77***	<.001
	Age (Scaled)	-0.47	.17	-2.72**	<.01
	Prime Type*Age (Scaled) BEI*Age vs OSV*Age	0.57	.28	2.01*	.04

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

scaled) in addition to Prime Type (BA, BEI, and OSV) as fixed effects for the CHI group only.

Table (3) shows the model where BEI was chosen as the reference level for the dependent variable (Response Type). Similarly, only significant factors of theoretical interest were included in this table. Overall, age significantly predicted priming across structures where older children were more likely to be primed. However, the effect of age was different across structures in terms of production patterns. Specifically, after BEI-primed, older children were less likely to produce BA-, OSV-, SVO-responses and reversal errors. Similarly, after OSV-primed, the production of BA-, and BEI-responses became less likely in older children. However, the production patterns after BA-primed were not modulated by age.

Comprehension task

Accuracy data

Figure (5) shows the offline comprehension accuracy of the two groups across different conditions and structures. The ADT group reached ceiling (above 90 per cent accuracy)

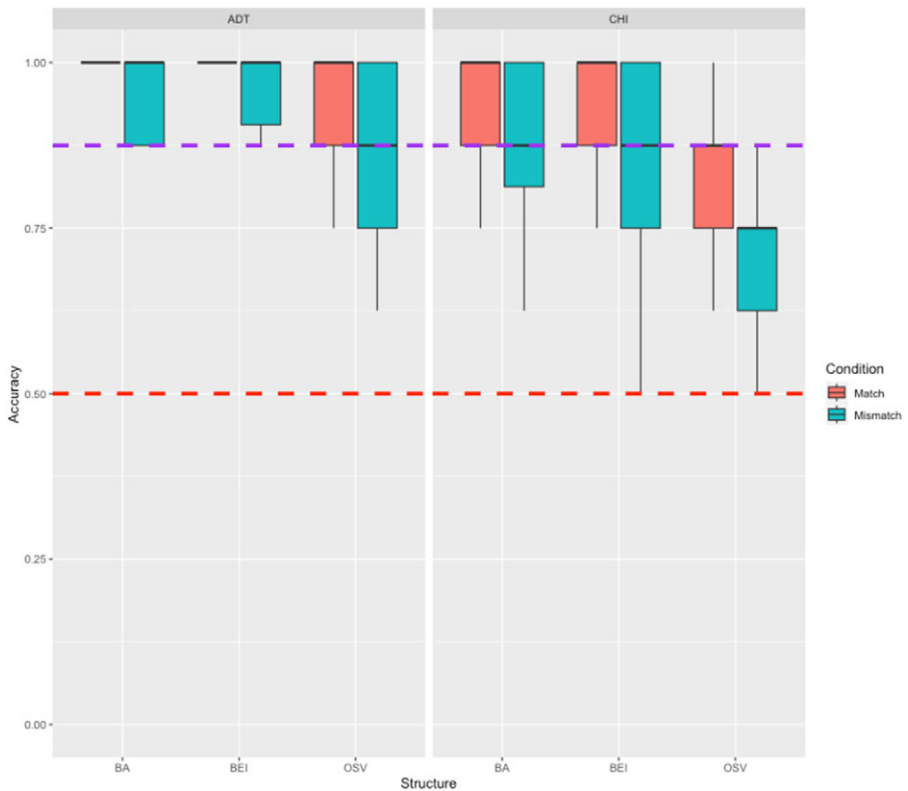


Figure 5. Offline comprehension accuracy across Conditions and Structures in the ADT and CHI groups.

across structures and conditions, although lower accuracy could be observed in the mismatched conditions. On the other hand, the CHI group also reached ceiling across structures in the matched conditions, with larger individual variabilities compared with the ADT group. Whereas, in the mismatched conditions, the CHI group made more errors, especially in the OSV-construction, making the mean accuracy across structures below 90%.

To statistically account for the data and answer RQ 2, we adopted generalised linear mixed effect analyses with a logistic link function (GLML) as we coded the accuracy data as a binary outcome (correct vs incorrect). As fixed effects, Group (ADT and CHI), Condition (Match and Mismatch) and Structure (BA, BEI, and OSV) were entered. The optimal model included all three variables without any of their interaction terms (Table 4)

As revealed by the model and post hoc analyses with re-levelled reference levels, the CHI group only differed from the ADT group in terms of overall accuracy, i.e., the CHI group performed significantly worse than the ADT group across structures and conditions. On the other hand, performance in the CHI group was modulated by Condition and Structure in the same way as it was in the ADT group: (1) matching effect across structures; (2) more errors in OSV-constructions across conditions; (3) no differential performance between BA- and BEI-constructions across conditions.

Table 4. Optimal model with Group (**ADT** and **CHI**), Condition (**Match** and **Mismatch**) and Structure (**BA**, **BEI**, and **OSV**) as fixed effects for the accuracy data in the comprehension task

	Estimate	Standard Error	<i>t</i>	<i>p</i>
(Intercept)	3.86	0.22	17.42***	<.001
GroupCHI	-1.23	0.19	-6.60***	<.001
ConditionMismatch	-0.93	0.16	-5.89***	<.001
StructureBA	0.32	0.17	1.88	0.06
StructureOSV	-0.99	0.14	-6.94***	<.001

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

We then excluded the **ADT** group and ran another GLML to examine the effect of Age in the **CHI** group (RQ 3). The optimal model included Condition (**Match** and **Mismatch**), Structure (**BA**, **BEI**, and **OSV**) and Age (Scaled) as fixed effects. Similarly, no interaction term was selected in the model. As also evident in figure (6), Age positively predicted comprehension accuracy in the **CHI** group across structures and conditions (Estimate = 0.51, SE = 0.09, *t* = 5.44***, *p* <.001).

Reaction times

After the analysis of accuracy data, we only included the trials where participants gave correct offline comprehension responses for reaction time analysis. The exclusion of RT extreme values and outliers resulted in a deletion of 316 (4.88%) data points in the **ADT** group and 368 (5.63%) in the **CHI** group.

Figure (7) illustrates how listening times (represented by residual RTs; RTs henceforth) contrast between the **ADT** and the **CHI** groups across segments, conditions, and structure types. As seen in the plots, the **ADT** and the **CHI** group showed a qualitatively similar pattern across structures, such that a matching effect surfaced in the critical segment (Segment 3) where the morphosyntactic and/or word order cue was available.

To statistically determine if the **CHI** group differed from the **ADT** group in their processing of the three structures and the role of structure type (RQ 2), residual RTs (scaled) were entered as the dependent variable in several liner mixed-effect analyses (LMs). Firstly, we included Group (**ADT** and **CHI**), Structure (**BA**, **BEI**, and **OSV**), Condition (**Match** and **Mismatch**) and Segment (Segment 1, Segment 2, **Segment 3**, Segment 4, and Segment 5) as fixed effects. However, adding the optimal model identified including Segment as a fixed effect failed to converge. Therefore, we ran separate models for each segment.

For Segment 1, only Group (**ADT** and **CHI**) was included in the optimal model suggesting that the **ADT** group took a shorter time listening to the first segment (Estimate = 0.26, SE = 0.07, *t* = 3.57***, *p* <.001). For Segment 2, however, no significant effect was found. As for the critical segment (Segment 3), Group (**ADT** and **CHI**) was not included in the optimal model which had Structure (**BA**, **BEI**, and **OSV**), Condition (**Match** and **Mismatch**) and their interaction as fixed effects (Table 5).

As revealed by the model and post hoc analyses disentangling interaction terms, the results showed that for both the **ADT** group and the **CHI** group, a matching effect was found in the critical segment where the mismatched trials took a longer time to listen to

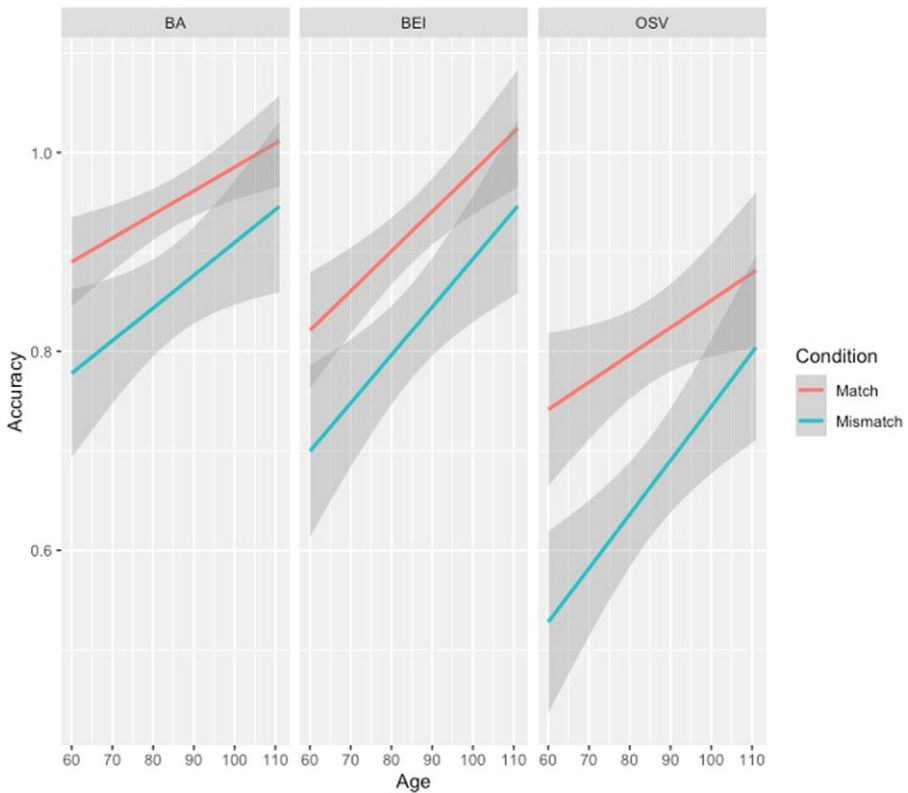


Figure 6. Offline comprehension accuracy as a function of Age across Conditions and Structures in the CHI group.

compared with the matched trials in both groups. Importantly, such a matching effect was more prominent for *OSV*-constructions relative to the other two constructions which were not significantly different from each other.

For the post-critical segment (Segment 4), the optimal model included Structure (BA, BEI, and OSV), Condition (Match and Mismatch) and their interaction as fixed effects (Table 6). Again, Group (ADT and CHI) was not included.

Similarly to the critical segment, as suggested by the analyses, the effect of matching lingered to the post-critical segment as well. Importantly, such a lingered matching effect did not differ significantly between the CHI group and the ADT group and was again more prominent for *OSV*-constructions relative to the other two constructions.

Lastly, for Segment 5, an effect of Structure (BA, BEI, and OSV) was identified. Specifically, both the ADT group and the CHI group took longer listening to *OSV*-constructions than to *BEI*-constructions ($Estimate = 0.16, SE = 0.05, t = 3.13^{**}, p < .01$) and to *BA*-constructions ($Estimate = 0.10, SE = 0.05, t = 2.09^*, p < .05$).

Then we fitted models to examine the effect of age (RQ 3) in the CHI group. For a more targeted identification of the age effect, we adopted the forward model selection approach here, i.e., comparing the models with Age (Scaled) to the null models (firstly for models including Segment as a fixed effect and secondly for models run separately for each segment). However, models with Age as a fixed effect did not improve model fits from the null models, i.e., Age did not predict the RTs across structures, conditions, or segments.

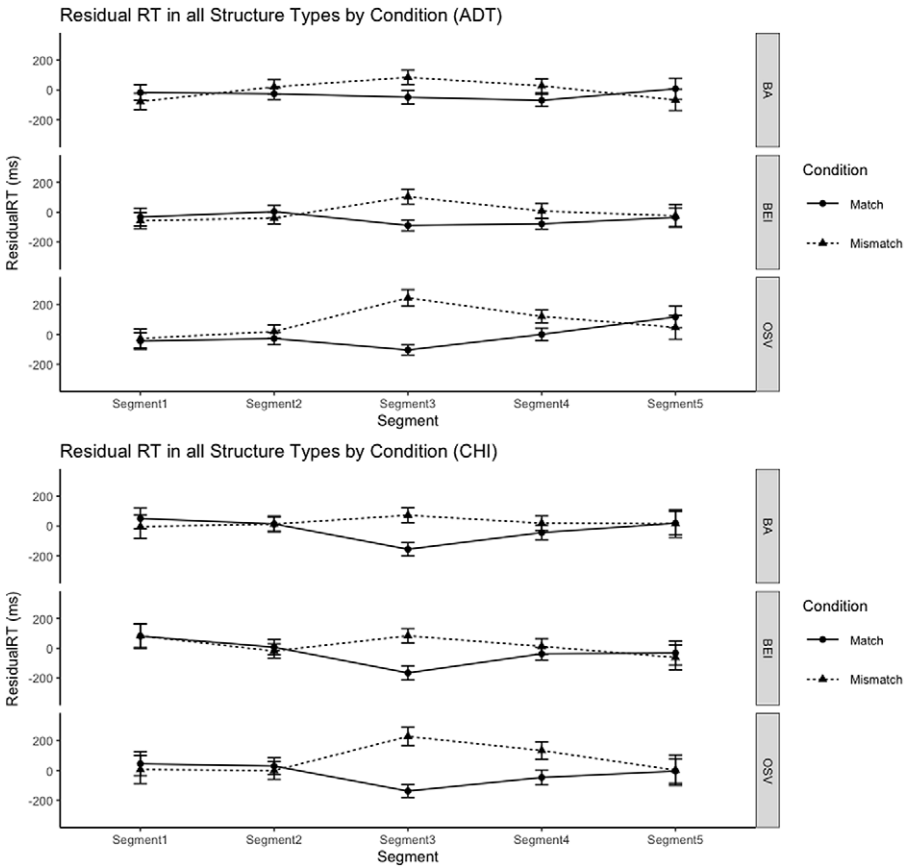


Figure 7. Residual RTs across the ADT and CHI groups crossed with Condition or Structure Type.

Table 5. Optimal model with Structure (BA, BEI, and OSV) and Condition (Match and Mismatch) as fixed effects for the RTs in Segment 3 (critical segment) for the ADT and CHI groups

	Estimate	Standard Error	<i>t</i>	<i>p</i>
(Intercept)	-0.32	0.07	-4.61	<.001
StructureBA	0.00	0.06	0.07	0.94
StructureOSV	-0.03	0.06	-0.40	0.69
ConditionMismatch	0.53	0.08	7.02***	<.001
StructureBA:ConditionMismatch	0.05	0.08	0.64	0.52
StructureOSV:ConditionMismatch	0.40	0.08	4.72***	<.001

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

Table 6. Optimal model with Structure (BA, BEI, and OSV) and Condition (Match and Mismatch) as fixed effects for the RTs in Segment 4 (post-critical segment) for the ADT and CHI groups

	Estimate	Standard Error	<i>t</i>	<i>p</i>
(Intercept)	-0.16	0.05	-3.17**	<.01
StructureBA	0.01	0.06	0.09	0.93
StructureOSV	0.08	0.06	1.32	0.19
ConditionMismatch	0.18	0.06	2.77**	<.01
StructureBA:ConditionMismatch	0.04	0.09	0.46	0.65
StructureOSV:ConditionMismatch	0.25	0.09	2.73**	<.01

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

Discussion

The present study examined the production and comprehension of non-canonical structures with differing linguistic properties (i.e., word order and the presence or absence of morphosyntactic cues) in typically developing Mandarin-speaking children aged five to ten and an adult control group. Three research questions were addressed: how children produce (RQ1) and comprehend (online and offline; RQ2) non-canonical structures, and importantly whether the production and comprehension are modulated by structure type, and whether they develop with age (RQ3).

Production (RQ1)

The production task was a comprehension-to-production priming task. Unlike other priming studies (e.g., Messenger et al., 2012), we analysed production patterns when participants were not primed as well as syntactic priming. Overall, we found that children were less likely to be primed across structures, but their performance was modulated by structure type in a similar way compared to that of adults.

For RQ 1, the results showed that children were less accurate/likely in producing all three structures compared to adults. However, both the ADT and CHI groups produced more BEI-constructions than BA- and OSV-constructions, which did not differ from each other. When the primed structures were not produced, children produced (1) more SVO-constructions after BA-primers but not after BEI- or OSV-primers, (2) more of the other two non-canonical structures, especially BA-constructions after BEI- or OSV-primers, (3) more reversal errors only after BEI-primers, relative to adults. The differential production patterns within children and compared to adults as a function of prime type are critical, suggesting children's difficulties with these structures cannot be attributed to their differential preference for cues. Although children did produce more SVO-constructions than adults did, this was only after BA-primers. Additionally, the production of SVO-constructions was more likely after BA-primers than after BEI- or OSV-primers for both children and adults. Therefore, the result that children produced more SVO-constructions after BA-primers is less likely to suggest something developmental and/or specific to children. We postulate here that this might be caused by the nature of the priming paradigm. In the case of BA-constructions because BA-constructions also share the thematic role ordering with SVO-constructions, the priming of the specific syntax alone might be less prominent. Additionally, different populations might be differentially

primed at different levels, e.g., emphasis on the subject/object, constituent order, syntax, etc. For example, unlike adults who were primed syntactically more often and opted for the exact same syntax, children might be more sensitive to priming of constituent order, e.g., children's higher likelihood of producing *BA*-constructions and reversal errors after *BEI*- and *OSV*-primes might be caused by the fact that children were more primed on constituent order.

Although children's increased production of *OSV*-constructions after *BEI*-primes could be explained by their different sensitivities to different levels of priming as we postulated above, it could also be the case that they dropped the *bei* cue as predicted by the Competition Model. However, the fact that children did not drop the *ba* cue after *BA*-primes (causing reversal errors) as much as adults did, and they produced more *BEI*-constructions after *OSV*-primes than adults did (and after *BA*-primes) speaks against an omission explanation. Future research using naturalistic production measurements or tasks without priming would be of great importance. Priming also differed in magnitude as a function of structure type in a similar way for children and adults.

For RQ 3, we found age to modulate children's production, with older children showing stronger priming across structures. For production patterns, age seemed to have differential effects across structures. Specifically, the likelihood of producing *SVO*-constructions when not primed by *BA*-constructions did not significantly decrease with age. However, with the increase of age, children did produce less often the other two non-canonical structures when not primed by *BEI*- and *OSV*-constructions. This is interesting because if age positively predicted overall priming, it means that older children should converge with adults in their sensitivity to syntactic priming over priming of other levels, i.e., increase in syntactic priming and decrease in priming of other levels. However, a decrease in the priming of other levels was only found in *BEI*- and *OSV*-constructions but not in *BA*-constructions. Here, we argue that sensitivity to syntactic priming is developmental but sensitivity to different levels of priming, e.g., information structure, constituent order, etc., is further modulated by the (dis-)similarity among different levels of features of a structure. We leave this for future research to scrutinise and encourage researchers to examine both priming and production pattern when not primed at the same time.

Comprehension

The comprehension experiment, i.e., a self-paced listening task with picture verification, investigated how children make use of different linguistic cues to process in real-time and interpret offline three Mandarin non-canonical structures. Overall, the results suggested that children's offline comprehensions were more prone to errors (Huang et al., 2013), but children used different cues in an adult-like way (Zhou & Ma, 2018).

Starting with offline comprehension, for RQ2, we found that children had lower accuracy across structures and conditions compared to adults, even though they also performed at ceiling across structures in the matched conditions. This is in contrast to Hao and Chondrogianni (2021) and Zhou and Ma (2018) who found indistinguishable offline comprehension accuracy between children and adults. This discrepancy might reflect differences in tasks and participants between the studies, but not a reflection in differences of syntactic representations, especially when children performed at ceiling as well in the matched conditions. Specifically, the experimental sentences in the current study were longer in terms of the number of words compared with both studies; and

participants also listened to the sentence in real time, rather than just deciding at the end of the sentence. Therefore, the present study taps into different abilities (real-time processing) and offline comprehension using more demanding materials; and the accuracy differences between children and adults might be a suggestion that children were less accurate in performing a more complex task. Additionally, the current study had a larger sample size both in children and adults and more items, i.e., it had more power to detect any differences.

Differential performances among structures were observed in offline comprehension such that the *BA*- and *BEI*-constructions were comprehended equally well (Zhou & Ma, 2018) and had higher accuracy than *OSV*-constructions (a numerical tendency in Hao & Chondrogianni, 2021). However, again, because differences in comprehension accuracy across structures were not limited to children but were attested in adults as well, it is more likely to reflect sentence comprehension mechanisms in general rather than something only developmental. We argue that morphosyntactic cues in non-canonical structures assisted comprehension, and that children as young as five converged with adults on how they use these cues (see also Huang et al., 2013; Özge et al., 2019; Zhou & Ma, 2018, among others). Importantly, this was further corroborated by our online processing results.

In terms of an effect of age in offline comprehension (RQ 3), we found that age positively modulated children's offline comprehension accuracy across structures and conditions (Zhou & Ma, 2018). Importantly, such a developmental effect did not differ across structures (see also Hao & Chondrogianni, 2021). Although this would lend support to the Incremental Processing Account, we postulate, based on results from online processing which we will discuss later, that age might have modulated other cognitive processes involved in making correct judgements in the task and not just specifically the comprehension of the three specific syntactic structures. For example, older children might have better executive functions to perform better meta-linguistically, and this may in turn have affected their overall performance across conditions.

For online processing, however, the only difference between children and adults was that children took longer to listen to the first segment across structures and conditions. Otherwise, children were indistinguishable from adults both quantitatively and qualitatively in the other four segments across structures and conditions (RQ 2). For both children and adults, a mismatching effect was found in Segment 3 where morphosyntactic cues and word order information were available; this effect lingered to Segment 4 for all structures. Furthermore, listening time was also modulated by structure type, such that longer listening times in Segment 3, 4 and 5 were found in *OSV*-constructions compared to *BA*- and/or *BEI*-constructions, constituting evidence for an *OSV*-disadvantage, a pattern also found in offline comprehension (and production). Importantly, it was found in both children and adults and in both the Match and Mismatch conditions. Meanwhile, *BA*- and *BEI*- were processed in a similar way in both groups. As we argued, these reflected the assistive role of morphosyntactic cues in non-canonical structure processing for both children and adults. Additionally, the fact that *OSV*-constructions induced longer listening times even in the last segment could explain the lower offline accuracy of this structure across conditions. On the other hand, the reason for a longer listening time for children in Segment 1 might be that children need more time to get themselves prepared and familiarised with switching back and forth between tasks, i.e., from picture verification (from the last trial) to self-paced listening. Future research would benefit from adapting tasks with fewer demands in task switching.

Contrary to our offline comprehension results, we did not observe any age effects within the child group in the RT data across structures, conditions, and segments (RQ 3). The lack of age effects in online processing is surprising given the recent literature that shows domain-general processing ability predicts children's online processing (e.g., Woodard et al., 2016), if age is taken as a proxy for the development of executive functions. The lack of such an effect in the online processing component of our study could be related to the specific age range and the number of participants. It might be the case that more children at the lower and higher end of the distribution were needed for age effects to emerge in the online processing task. In our study, age effects were only observed as longer RTs between the children and the adults, but not within the child group itself.

Theoretical implications of present findings

Evidence for the Competition Model comes from the production and the offline comprehension results such that children were less primed and had lower comprehension accuracy across conditions and structures compared to adults. In online processing, the fact that children actively interpreted NPIs as agents (as evident in their engaging in syntactic reanalysis) also provides evidence for the Competition Model. However, children did not perform better in *BA*-constructions relative to *BEI*-constructions (the *ba* cue has higher availability than the *bei* cue) and showed adult-like use of less available but highly reliable morphosyntactic cues during real-time processing, challenging the Competition Model. In other words, cue reliability trumps cue availability in children's production and offline comprehension. However, the results from online processing showed that children integrate highly reliable and available cues in a similar way. This may indicate that the automaticity of integrating cues with differing reliability and availability in real-time differs between online processing and children's offline performance or production (see also Özge et al., 2015, 2019; Zhou & Ma, 2018, among others). Future research could scrutinise the relationship between these different modalities from a developmental perspective.

For the Incremental Processing Account, the most direct evidence is that children incrementally processed non-canonical structure using different cues interactively. However, the children in our study were able to revise their initial (mis-)interpretations and reached ceiling in offline comprehension, contrary to previous empirical findings, e.g., Choi and Mazuka (2003), Clackson et al. (2011), Trueswell and Gleitman (2004), among others. We note here that this might be because previous studies mainly examined the use of different cues from different (non-)linguistic (sub-)domains, e.g., discourse vs. syntactic cues, prosodic vs. syntactic cues, whereas we examined the use of different cues within the (morpho)syntactic domain. This also has implications for the other two findings that cannot be straightforwardly explained by the Incremental Processing Account, i.e., the lack of better performance on *BA*-constructions and the better performance on *BEI*-constructions relative to *OSV*-constructions. We propose that although the Incremental Processing Account assumes the use of different cues to be a dynamic process, it would benefit from further incorporating cue availability and/or reliability. For example, if it is indeed the case that the *bei* cue is more reliable than the *ba* cue (Li et al., 1992), it is not surprising to find equal performance between *BA*- and *BEI*-constructions, as the high reliability of the *bei* cue cancels out potential advantage of the word order cue available in *BA*-constructions (*SVO*). Future research on different cues (morphosyntactic, prosodic, etc.) would shed light on this question.

Conclusion

The present study found that, for the production and comprehension of non-canonical structures, Mandarin-speaking children (1) were syntactically primed to a lesser degree compared to adults, and showed production patterns (when not primed syntactically) indicative of priming of other linguistic levels; (2) were indistinguishable from adults in how they made use of different linguistic cues in real time; and (3) were more subject to errors than adults in offline comprehension. We interpreted these results as showing that children have less developed abilities in performing the tasks rather than their having different linguistic representations and processing strategies. Additionally, children's priming magnitude in production and offline comprehension accuracy were positively correlated with age. We argued that such an age effect was not a mere reflection of better syntactic knowledge and processing ability, but more of better task performance and increased sensitivity to syntactic priming over and above priming of other levels. We took these to argue that available first language acquisition accounts in their current formats cannot explain children's difficulties with non-canonical structures. Postulations on how the current theories could be further developed and points to methods were discussed.

Data availability statement. Supplementary materials, including data and statistical analyses, to this article can be found online at: <https://osf.io/h7suz/>.

Competing interest. The author(s) declare none.

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>
- Bever, T. G. (2013). The cognitive basis for linguistic structures I. In M. Sanz, I. Laka, & M. K. Tanenhaus (Eds.), *Language Down the Garden Path* (pp. 1–80). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199677139.003.0001>
- Borer, H., & Wexler, K. (1987). The Maturation of Syntax. In T. Roeper & E. Williams (Eds.), *Parameter Setting* (Vol. 4, pp. 123–172). Springer Netherlands. https://doi.org/10.1007/978-94-009-3727-7_6
- Borer, H., & Wexler, K. (1992). Bi-unique relations and the maturation of grammatical principles. *Natural Language and Linguistic Theory*, *10*(2), 147–189. <https://doi.org/10.1007/BF00133811>
- Branigan, H. P., & Pickering, M. J. (2017). Structural priming and the representation of language. *Behavioral and Brain Sciences*, *40*, e313. <https://doi.org/10.1017/S0140525X17001212>
- Brooks, P. J., & Tomasello, M. (1999). Young children learn to produce passives with nonce verbs. *Developmental Psychology*, *35*(1), 29–44. <https://doi.org/10.1037/0012-1649.35.1.29>
- Choi, Y., & Mazuka, R. (2003). Young Children's Use of Prosody in Sentence Parsing. *Journal of Psycholinguistic Research*, *32*(2), 197–217. <https://doi.org/10.1023/A:1022400424874>
- Clackson, K., Felser, C., & Clahsen, H. (2011). Children's processing of reflexives and pronouns in English: Evidence from eye-movements during listening. *Journal of Memory and Language*, *65*(2), 128–144. <https://doi.org/10.1016/j.jml.2011.04.007>
- 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>
- Demuth, K. (1989). Maturation and the Acquisition of the Sesotho Passive. *Language*, *65*(1), 56. <https://doi.org/10.2307/414842>
- 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>
- Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German Children's Comprehension of Word Order and Case Marking in Causative Sentences. *Child Development*, *79*(4), 1152–1167. <https://doi.org/10.1111/j.1467-8624.2008.01181.x>

- 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>
- Hao, J., Chondrogianni, V., & Sturt, P. (2023). Heritage language development and processing: Non-canonical word orders in Mandarin–English child heritage speakers. *Bilingualism: Language and Cognition*, 1–16. <https://doi.org/10.1017/S1366728923000639>
- Hu, S., Guasti, M. T., & Gavarró, A. (2018). Chinese Children's Knowledge of Topicalization: Experimental Evidence from a Comprehension Study. *Journal of Psycholinguistic Research*, 47(6), 1279–1300. <https://doi.org/10.1007/s10936-018-9575-6>
- 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>
- Leonard, L. B., Wong, A. M.-Y., Deevy, P., Stokes, S. F., & Fletcher, P. (2006). The production of passives by children with specific language impairment: Acquiring English or Cantonese. *Applied Psycholinguistics*, 27(2), 267–299. <https://doi.org/10.1017/S0142716406060280>
- 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)
- MacWhinney, B. (1987). The competition model. In B. MacWhinney, National Science Foundation (U.S.), & Carnegie-Mellon University (Eds.), *Mechanisms of language acquisition [sic]*. L. Erlbaum Associates.
- MacWhinney, B. (2018). A unified model of first and second language learning. In M. Hickmann, E. Veneziano, & H. Jisa (Eds.), *Trends in Language Acquisition Research* (Vol. 22, pp. 287–312). John Benjamins Publishing Company. <https://doi.org/10.1075/tilar.22.15mac>
- Mangiafico, S. S. (2016). *Summary and analysis of extension program evaluation in R* (1.18.8). rcompanion.org/handbook/.
- Marinis, T., & Saddy, D. (2013). Parsing the Passive: Comparing Children with Specific Language Impairment to Sequential Bilingual Children. *Language Acquisition*, 20(2), 155–179. <https://doi.org/10.1080/10489223.2013.766743>
- 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>
- Novick, J. M., Thompson-Schill, S. L., & Trueswell, J. (2008). Putting lexical constraints in context into the visual-world paradigm. *Cognition*, 107(3), 850–903. <https://doi.org/10.1016/j.cognition.2007.12.011>
- Omaki, A., Davidson White, I., Goro, T., Lidz, J., & Phillips, C. (2014). No Fear of Commitment: Children's Incremental Interpretation in English and Japanese Wh-Questions. *Language Learning and Development*, 10(3), 206–233. <https://doi.org/10.1080/15475441.2013.844048>
- Özge, D., Küntay, A., & Snedeker, J. (2019). Why wait for the verb? Turkish speaking children use case markers for incremental language comprehension. *Cognition*, 183, 152–180. <https://doi.org/10.1016/j.cognition.2018.10.026>
- Özge, D., Marinis, T., & Zeyrek, D. (2015). Incremental processing in head-final child language: Online comprehension of relative clauses in Turkish-speaking children and adults. *Language, Cognition and Neuroscience*, 30(9), 1230–1243. <https://doi.org/10.1080/23273798.2014.995108>
- R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Slobin, D. (1973). Prerequisites for the development of grammar. *Studies of Child Language Development*. New York, Holt, Rinehart and Winston.
- Snedeker, J., & Huang, Y. T. (2015). Sentence processing. In E. L. Bavin & L. R. Naigles (Eds.), *The Cambridge Handbook of Child Language* (2nd ed., pp. 409–437). Cambridge University Press. <https://doi.org/10.1017/CBO9781316095829.019>
- Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, 74(3), 209–253. [https://doi.org/10.1016/S0010-0277\(99\)00069-4](https://doi.org/10.1016/S0010-0277(99)00069-4)
- Trueswell, J., & Gleitman, L. (2004). Children's Eye Movements during Listening: Developmental Evidence for a Constraint-Based Theory of Sentence Processing. In H. J. & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (p. 28).

- Trueswell, J., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, *73*(2), 89–134. [https://doi.org/10.1016/S0010-0277\(99\)00032-3](https://doi.org/10.1016/S0010-0277(99)00032-3)
- Woodard, K., Pozzan, L., & Trueswell, J. (2016). Taking your own path: Individual differences in executive function and language processing skills in child learners. *Journal of Experimental Child Psychology*, *141*, 187–209. <https://doi.org/10.1016/j.jecp.2015.08.005>
- 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>

Cite this article: Hao, J., Chondrogianni, V., & Sturt, P. (2024). Sources of children's difficulties with non-canonical sentence structures: Insights from Mandarin. *Journal of Child Language* 1–28, <https://doi.org/10.1017/S0305000924000424>