Journal of Child Language (2026), 1–27
doi:10.1017/S0305000926100750

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Noun-Modifying Clause Construction in Mandarin-Speaking Children

Mengyao Shang[1], Ziyin Mai[1], Stephen Matthews[2] and Virginia Yip[1]

[1]The Chinese University of Hong Kong, Hong Kong and [2]The University of Hong Kong, Hong Kong
Corresponding author: Ziyin Mai; Email: maggiezymai@cuhk.edu.hk

(Received 04 April 2025 UTC; revised 21 March 2026 UTC; accepted 17 May 2026 UTC)

Abstract

In Mandarin, the head noun in [modifier clause-de-N] encodes not only the arguments or adjuncts of the clausal verbs (e.g. "John/biscuits/plate" in the event of "John eating biscuits") but also non-argument, non-adjunct participants extended from the clausal event (e.g. "crumbs" resulting from the event of "John eating biscuits"). This article adopts a typological, multi-factorial analytic framework of noun-modifying clause constructions (NMCCs) and proposes a developmental hierarchy of three types of NMCCs: argument > adjunct > extended. We tested the hierarchy in two studies: an analysis of 728 tokens of NMCC produced by 369 children (1;2–6;9) in naturalistic corpora and a contextualised sentence repetition task conducted with 121 children (3;0–4;11) in China. Our findings show that NMCC emerges surprisingly early around 2;0 in Mandarin children. Argument NMCCs show developmental primacy in frequency and accuracy. Adjunct and extended NMCCs are acquired at a later stage, consistent with the proposed hierarchy.

Keywords: Mandarin Chinese; language acquisition; grammatical development; child language corpus; elicited imitation

摘要

汉语普通话名词修饰子句结构以[修饰子句-的-名词]的形式出现。其中被修饰的名词既可以是修饰子句中动词的论元或者附加语（如"约翰吃饼干"事件中的约翰、饼干、盘子），还可延伸至其他事件元素，如"约翰吃饼干"所产生的饼干碎屑。本文用类型学框架分析名词修饰子句，同时考虑影响其习得的多种因素，提出名词修饰子句的发展优先级假设：论元型 > 附加语型 > 延伸型。我们通过两项研究验证该假设:一，我们分析了自然语料库中369 名普通话儿童 (1;2–6;9) 产出的 728 个名词修饰子句。二，我们通过句子重复实验测得 121 名儿童 (3;0–4;11) 在特定语境下产出三种名词修饰子句的准确率和偏误。结果显示，普通话儿童早在 2;0 岁左右开始产出名词修饰子句。论元型名词修饰子句在自然语境中的使用频率和重复实验中的准确率均呈早期发展优势，而附加语型与延伸型的习得相对滞后。这一系列结果均符合我们提出的发展优先级假设。

© The Author(s), 2026. 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.

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

1. Introduction

Nouns are among the very first lexical items that young children comprehend and spontaneously produce by the end of the first year or soon after that (Tardif et al., 2008). Modifying and restricting the nouns with lexical elements (e.g. adjective, noun) or clauses (e.g. relative clause) to form complex nominal structures is a natural next step, which can be a lengthy process for young learners aged 2–6 cross-linguistically (e.g. Diessel & Tomasello, 2000, 2005; Mintz & Gleitman, 2002). In English, the post-nominal *relative clause (RC)* restricts the referent of the modified head noun such that the resulting noun phrase (NP) denotes a smaller set of referents than the otherwise bare head noun. A filler–gap dependency between the head noun and a syntactic gap in the clause is commonly observed (Hawkins, 1999). In (1), for instance, the head noun is co-indexed with a gapped argument position in the RC.

(1) English relative clause (RC)
the phone$_i$ [which he bought __$_i$].

However, such a syntactic gap is not necessary in Mandarin RCs. In Mandarin, RC is linked to the nominal head by particle *de* and appears in the prenominal position, on a par with other noun-modifying structures such as adjectival and nominal modifiers in the language. A distinct feature of such *nounmodifying clause constructions* in Mandarin (hereafter *NMCCs,* following Matsumoto et al., 2017), as opposed to RCs in English and many European languages, is that they do not necessarily involve a syntactic gap, and they appear to allow a looser and more "relaxed" relation with the head noun than would be possible in English-type RCs. That is, not only "gapped" NMCCs in which arguments are co-indexed with a syntactic gap in the modifying clause (e.g. *shouji* "mobile phone" in (2a)) can naturally appear; "gapless" NMCCs in which adjuncts (e.g. *shangdian* "shop" in (2b)) or nouns extended from the events depicted by the NMCC (e.g. *fapiao* "receipt" in (2c)) are also common (Tsai, 1997; Comrie, 1998; Cheng & Sybesma, 2006; Wu, 2022; Pan, 2022; cf. Huang, 2016).

(2) Mandarin noun-modifying clause construction (NMCC): [modifying clause] *de* head.
a. argument NMCC: [ta mai __$_i$] de shouji$_i$
3SG buy *de* mobile phone
"the mobile phone that he bought"

b. adjunct NMCC: [ta mai shouji] de shangdian
3SG buy mobile phone *de* shop
"the shop where he bought a mobile phone"

c. extended NMCC: [ta mai shouji] de fapiao
3SG buy cell phone *de* receipt
lit: "the receipt that (is related to) him buying a mobile phone"
or "the receipt for the mobile phone that he bought"

The uniform modifier-noun sequencing in Mandarin noun-modifying structures and the wider range of head–clause relations in Mandarin NMCC provide valuable opportunities to understand how young children develop complex structures to modify the noun. Will children acquire extended NMCC first and then progress to

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

adjunct and argument types, because extended NMCC allows a more "relaxed" relation between the head and the modifying clause? Despite many studies on argument NMCCs (e.g. Chen & Shirai, 2015; Yang et al., 2020) and general prenominal modification (e.g. Liu & Chan, 2012; Packard, 1988) in child Mandarin, the full range of NMCCs in child Mandarin has not been studied systematically and rigorously. To fill this gap, this study comprehensively describes and analyses Mandarin NMCCs produced by young children in naturalistic and elicited settings, following a typological analytic framework which groups together languages with a general NMCC (Matsumoto et al., 2017).

2. Mandarin NMCC: a typological analytic framework

As noted in Comrie (1998), a number of areal-related Asian languages, such as Japanese, Korean, and Mandarin, share a general NMCC that covers a substantially broader range of uses than RCs as traditionally defined. Mandarin NMCCs are prenominal noun-modifying clauses linked to the nominal head through the particle *de*. Depending on the thematic relations between the head and the lexical verb in the modifying clause, NMCCs can be analysed as argument, adjunct, and extended types.

In *argument NMCCs* (or RCs as traditionally defined), the head is coreferential with an argument in the modifying clause. Studies on RCs posit that relativisation leaves a syntactic gap in the clause, for example, subject and object RCs should have a gap_i in the subject or object position co-indexed with the head, as in (3a&b). However, as Mandarin allows null arguments, argument NMCCs like (3c), which are typically defined as RCs, can be interpreted as a subject RC containing a subject gap_i as the THEME, or an object RC with an object gap_j as PATIENT. This well-known ambiguity within Mandarin argument NMCCs is not the focus of this study.

(3) Argument NMCC (subject and object RC)

a. [jiaoshou fabiao ___$_i$] de lunwen$_i$ (head noun = PATIENT)
professor publish *de* paper
"the paper that the professor published"

b. [___$_i$ fabiao lunwen] de jiaoshou$_i$ (head noun = AGENT)
publish paper *de* professor
"the professor that published papers"

c. [___$_i$ fabiao ___$_j$] de lunwen$_{i/j}$
publish *de* paper
"the papers that were published," or
"the papers that (someone) published"

In *adjunct NMCCs*, the head is conceptually coreferential with an adjunct of the modifying clause. Whether this entails a syntactic gap–filler relation between the modifying clause and the head is under debate (Comrie, 1998; see LaPolla, 2017 for Chinese; Yoon, 1993 for Korean; Kempson & Kurosawa, 2009 for Japanese). English allows adjunct heads to be relativised into RCs with relative pronouns encoding specific thematic relations between the head noun and the clausal event (e.g. *where* for LOCATION and *when* for TIME). Such adjunct RCs have an argument RC equivalent in which the head and

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

the RC are linked by [preposition-*which*] instead, as in (4a–b). Unlike argument and adjunct RCs in English which are relativised by two distinctive sets of relative pronouns entailing specific thematic relations between the head and the clause, one single linking particle *de* is shared by all Mandarin NMCCs. Crucially, full adjunct NMCCs can appear in [S-V-O-*de*-N] sequence, in which the specific thematic role of the head noun (e.g. LOCATION, INSTRUMENT) is not overtly marked and can thus only be inferred, see (5).

(4) Adjunct NMCC in English (from Radford, 2016: 394)
a. There are [places ***where*** [they sell counterfeit watches]].
a' … places ***at which*** [they sell counterfeit watches]
b. They lived in [times ***when*** [money was tight]].
b' … times ***at which*** [money was tight]

(5) Adjunct NMCC in Mandarin.
a. [ta xie shufa] de shufang (head noun = LOCATION)
3SG write calligraphy *de* study
"the study in which he practices calligraphy"

b. [ta daibu fanren] de shoukao (head noun = INSTRUMENT)
3SG arrest prisoner *de* handcuffs
"the handcuff with which he arrests prisoners"

In *extended NMCCs*, an extended array of semantic relations between the head noun and the modifying clause is possible. Like adjunct NMCCs, full extended NMCCs can appear in [S-V-O-*de*-N] sequence, in which the specific thematic role of the head noun is not overtly marked. In (6a), the head *weidao* "smell" is related to the clausal event "fish being burnt," forming a head–clause association on the basis of the pragmatic encyclopedic knowledge that burnt food produces a distinct smell, a by-product of fish-grilling. Similarly, in (6b), the head noun *yinyue* "music" is connected with the clausal event *wo paobu* "I run," as an accessory of the event, by the pragmatic plausibility of listening to music while running. In (6c), the head *diianying* "movie" has a causal relation with the clausal event *shui-zhao* "fall asleep." Extended NMCCs are not unique to Mandarin or Chinese languages; they are also observed in Korean (Lee & Lee, 2012; Yoon, 1993) and Japanese (Kempson & Kurosawa, 2009; Kurosawa, 2003).

(6) Extended NMCC in Mandarin.
a. [ta kao-hu yurou] de weidao (head noun = BY-PRODUCT)
3SG grill-burnt fish *de* smell
Lit. "the smell that (is related to) he burnt the fish"

b. [wo paobu] de yinyue (head noun = ACCESSORY)
1SG run *de* music
Lit. " the music that (is related to) I run"

c. [ni shui-zhao] de na-bu dianying (head noun = EXTERNAL CAUSE)
2SG fall asleep *de* that-CL movie
Lit. "the movie that (is related to) you fall asleep"
" the movie that made you fall asleep"

(a–c adapted from Matsumoto, 2017: 36, LaPolla, 2017: 97, and Matthews & Yip, 2017: 114, respectively)

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

3. Relative clause and prenominal modifications in child Mandarin

Although acquisition studies adopting the three-way split typological classification of NMCC as in our study are scarce, there has been systematic documentation of argument NMCC (or "RC" in its narrow sense), as well as prenominal modification in children. Studies on the acquisition of NMCC found a robust developmental pattern in English and many European languages: subject RCs, where the head is relativised from a subject position in the clause, are acquired earlier and less problematic than object RCs, which are relativised from an object position (see Lau & Tanaka, 2021, for a review). Research on the acquisition of argument NMCC in Mandarin, however, has mixed findings. Since the goal of our study is not to compare subject and object RCs within argument NMCC, below we will review findings on the acquisition of RC in general, rather than delving into the subject advantage debate.

Using naturalistic data, Chen and Shirai (2015) analysed NMCCs (termed RCs) in spontaneous production of four children (0;11–3;5) and their caregivers in Beijing. The first use was documented at 1;4 and 1;7 in the two youngest children. Argument NMCCs (subject and object RCs) accounted for most of the RCs in both the children (80.1%) and their caregivers (76.2%). The remaining NMCCs (19.9% and 23.8%) were either "oblique RCs" or "gapless RCs," which correspond to the categories of adjunct and extended NMCCs in our study. Longitudinal development shows that the relative proportions of oblique and gapless RCs remain low across MLU stages from 2.10 to 3.88. Both the caregiver input and the child output demonstrate a higher frequency for argument NMCCs than adjunct and extended NMCCs. Additionally, early Mandarin RCs (0;11–3;5) are mostly stand-alone NPs (52.7%) with the head noun functioning as a direct object of the modifying clause (61.5%, e.g. *baba mai de ban* "the board that Daddy bought"), appearing in the [S-V-*de*-head N] sequence consistent with the canonical SVO word order.

Among experimental approaches to argument NMCCs, Hsu et al. (2009) reported high accuracy (83.5%) in producing simple argument NMCCs in the [V-O-*de*-head N] sequence to differentiate pictures contrasting actions (e.g. [_ *ti/reng zuqiu*] *de nanhai* "the boy who is kicking/throwing the football") by 4-year-olds. In a picture selection task, He et al. (2017) found that older children (4;11–6;6) performed significantly more accurately than younger children (3;6–6;6) in matching pictures to auditorily presented argument NMCCs embedded either in the matrix subject [[NMCC]$_S$-V-O] or in the matrix object [S-V-[NMCC]$_O$]. Depending on the type of embedding, accuracy ranged from 50% to 80% in the two groups, capturing rapid development in comprehension of argument NMCC from age 3 onwards. Yang et al. (2020) tested comprehension of NMCCs through a referent selection task with concurrent eye movements recorded. The 4-year-olds performed well above chance levels (>60%) in selecting the correct animal referent from an array of four candidates on the basis of an auditorily presented argument NMCC (e.g. [*laoshu qin _] de gongji* "the rooster that the mouse kissed"). Their fixation time at the correct referent corroborated their offline judgements. Overall, although NMCCs emerge very early before age 2 and all types are attested in spontaneous production, Mandarin children display gradual and measurable development in both production and comprehension of argument NMCCs between ages 2 and 6 in controlled experimental settings.

Regarding prenominal modification as a broad category, Packard (1988) analysed 141 tokens of prenominal modification in [modifier-*de*-head N] sequences produced by 2-year-olds in naturalistic play, showing that adjectival and nominal noun modifiers such as *gui-de-shu* "expensive-*de*-book" and *yufa-de-shu* "grammar-*de*-book" emerge earlier and are more productive than NMCCs in the toddlers. Liu and Chan (2012), who

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

analysed 1034 tokens of [modifier-*de*-head N] sequences produced by preschoolers (3;0–6;0), reported similar findings. Specifically, they found that 22.8% of the clausal modifiers fell outside the scope of RCs as traditionally defined, and these instances were rarely studied for their own sake.

The first and only acquisition study including a full range of NMCCs is Lai et al. (2023), who collected 60 tokens of NMCCs produced by Cantonese monolingual children (1;7–5;6) from longitudinal and cross-sectional corpora. Cantonese shares many structural features with Mandarin including canonical SVO word order and prenominal NMCCs without explicit marking of the thematic relation between the head noun and the clausal verb. In the longitudinal data, argument NMCCs were produced earlier and more frequently than the other types. Nevertheless, their cross-sectional data showed that argument NMCCs were the most productive type in the 4-year-olds, but not the 2-, 3-, and 5-year-olds. The seemingly contradictory results in the longitudinal and cross-sectional data were attributed to the small number of attested NMCCs.

4. Developmental hierarchy of Mandarin NMCC: a multi-factorial proposal

What are the factors shaping the development of NMCCs in Mandarin? We draw from two theoretical accounts in the RC literature. First, syntactic gap-based accounts posit that NPs that are higher in the Noun Phrase Accessibility Hierarchy are easier to relativise than those lower in the Hierarchy (e.g. subject > direct object > indirect object, Keenan & Comrie, 1977; O'Grady, 2011). Along the same line, RCs with the head nouns farther away from the syntactic gaps are harder to acquire than those with head nouns closer to the relevant gaps (Gibson, 1998; Hawkins, 1999; O'Grady, 1997). These approaches explain acquisition differences across NMCC types on the basis of grammatical roles of the head noun and the syntactic distance between the head and its corresponding clausal gap. Although they make testable predictions about argument NMCCs, they do not predict or explain the development of gapless adjunct and extended NMCCs in Mandarin, where syntactic gaps are absent (see (2), (5), (6)).

A second tradition invokes frequency in the input (Brandt et al., 2009; Kidd et al., 2007; Ozeki & Shirai, 2007), predicting that the type of NMCC that is most frequent in the input will be acquired earlier. In Mandarin, for example, argument NMCCs dominate NMCCs in naturalistic child-directed speech (75–78%, Chen & Shirai, 2015) and adult-to-adult speech (71%, Lancaster Corpus of Mandarin Chinese, Ming & Chen, 2010). Since, in those samples, the ratio between adjunct and extended NMCCs is not reported, we conducted our own corpus search in the child-directed speech by parents of two monolingual Mandarin children, Tong (Deng & Yip, 2018) and Xue (Zhang & Zhou, 2009) in Child Language Data Exchange System (CHILDES; see 5.2 for coding scheme). Our results showed that argument NMCCs accounted for 74.73% of total NMCCs, very similar to that reported in Chen and Shirai (2015) and Ming and Chen (2010). Interestingly, adjunct and extended NMCCs accounted for similar proportions of NMCCs in the input (12.09% and 13.19%, respectively, more details in Supplementary Table S1). Given this evidence, input frequency would predict an advantage of argument NMCCs and similar developmental trajectories between adjunct and extended NMCCs in Mandarin children.

We hypothesise a third factor that shapes the developmental trajectory of NMCC: the conceptual importance or prominence of the head noun's role in clausal events, ordered on a scale shown in (7). The scale is motivated by independent evidence from event

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

perception and memory research: located furthest to the left, AGENT and PATIENT (especially the AGENT) are most prominent in human event cognition. They are identified and configured into mental schemas most rapidly during processing of visual scenes and activated by verbs most consistently during lexical processing (Hafri et al., 2018; see Zacks, 2020 for a review). After them, other event elements such as INSTRUMENT, LOCATION, and TIME are also prominent: lexical processing research shows that when the verb entails specific instruments (e.g. *cut, dig*), the instrument is part of the generalised situation scheme and automatically activated by the verb, whereas information about the LOCATION of the action is less activated (Ferretti et al., 2001). Nevertheless, LOCATION and TIME are two powerful labels and organisers of events in memory, affecting memory retrieval and processing speed (Radvansky et al., 1998, 2017; Swallow et al., 2009). Finally, ACCESSORY, BY-PRODUCT, and EXTERNAL CAUSE of the event are the lowest in prominence. BY-PRODUCTS usually emerge at subsequent stages of the event (e.g. smell of burnt fish in (6a)). ACCESSORIES are optional elements that elaborate upon the core event structure by providing secondary instrumental details (e.g. music in (6b)). EXTERNAL CAUSES are initiating factors originating outside the relevant verbs' event model (e.g. movie in (6c)). The three clusters of event roles in (7) align with the three NMCC types (i.e. argument, adjunct, extended) identified by Matsumoto et al. (2017) in a non-arbitrary way. The alignment reflects universal tendencies in natural languages to map agents and patients to subject and object positions, and instruments, locations, and times to adjuncts.

(7) Scale of conceptual prominence of event roles (more prominent > less prominent)

				ACCESSORY		
			INSTRUMENT	BY-PRODUCT		
	AGENT	PATIENT	LOCATION	TIME	EXTERNAL CAUSE	
most prominent	|-*	----------*	----------		----------*-|	least prominent

Given the syntactic, input frequency and conceptual factors above, we propose a developmental hierarchy for the three types of NMCCs in Mandarin, shown in (8), which predicts that NMCCs further to the left should develop earlier than those further to the right. The dominance of argument NMCC over adjunct and extended NMCCs in early Mandarin (and Cantonese) has already received strong empirical support (e.g. Chen & Shirai, 2015; Lai et al., 2023; Liu & Chan, 2012), consistent with input frequency and conceptual prominence effects. The proposed superiority of adjunct NMCC over extended NMCC is, on the other hand, an uncharted territory and hence the main focus of our study.

(8) Developmental hierarchy of NMCC in Mandarin (earlier > later)
Argument NMCC > Adjunct NMCC > Extended NMCC

This study addresses two overarching questions: to what extent does the development of the three NMCC types in Mandarin-speaking children align with the proposed hierarchy, and to what extent can conceptual prominence explain the differences between adjunct and extended NMCCs? To address these overarching questions, we adopted a two-step approach: we first established a foundational profile of NMCC production through systematic analysis of spontaneous child speech, using a rigorous, theoretically informed coding scheme (Study 1). We then tested specific NMCC structures via an

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

elicited production task with 3- and 4-year-olds, controlling for NMCC-level (word order, animacy, length) and child-level factors (age, working memory (WM), grammatical proficiency) (Study 2).

5. Study 1: Corpus-based study

5.1. Research questions

The first study sought to provide developmental evidence on the emergence, frequency, and quality of three types of Mandarin NMCCs in naturalistic child speech. Specifically, we asked three questions: (i) At what age do monolingual Mandarin-speaking children start to produce argument, adjunct, and extended NMCCs, respectively? (ii) Among the three types of NMCCs, which type predominates in naturalistic child speech? (iii) What are the structural and semantic characteristics of early NMCCs in naturalistic child speech?

5.2. Methods

Data. As NMCCs are infrequent in naturalistic speech (Lai et al., 2023; Matthews & Yip, 2002), we exhausted available data of monolingual Mandarin children in the CHILDES (MacWhinney, 2000) in order to maximise the sample size. We also incorporated the newly constructed Beijing Child Mandarin Corpus (BJCMC, Mai et al., 2024). Table 1 presents details of the corpora used. All corpora recorded child speech during toy-play activities. According to the documentation of individual corpora, all the children were raised in Mandarin-dominant families and communities and received Mandarin input from native speakers. There are no indications that the children were auditorily or verbally impaired or had developed language atypically.

Table 1. Corpus data selected for this study

Corpus	Age	$N_{children}$	$N_{utterances}$
Cross-sectional			
LiZhou (Li & Zhou, 2008)	3;0–6;3	80	12169
Zhou 1&2 (Li & Zhou, 2004; Zhou, 2001)	1;2–6;0	185	17149
Chang 1&2 (Chang, 1998, 2000)	3;4–6;5	40	3860
BJCMC (Mai et al., 2024)	3;0–6;9	50	18117
Total	**1;2–6;9**	**355**	**51295**
Longitudinal			
Tong (Deng & Yip, 2018)	1;7–3;4	1	9250
Zhou 3 (Zhang & Zhou, 2009)	1;8–5;5	1	8350
Erbaugh (Erbaugh, 1992)	1;9–3;9	2	19879
TCCM (Cheung et al., 2011)	1;5–4;3	10	30990
Total	**1;5–5;5**	**14**	**68469**

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

Coding. We searched for utterances containing *de*[1] using the KWAL command in CLAN (MacWhinney, 2000); the first author manually checked the results and extracted the target NMCCs appearing in this sequence: [modifying clause]-*de*-head N. More details about the inclusion criteria are in Supplementary Table S2. In this study, a prenominal modifier is considered a "clause" only when it is predicated by a verb, defined as a word modifiable by the negator *bu* "not" and aspectual markers like *le*, excluding adjectives which are modifiable by degree adverbs like *hen* "very" (following Huang et al., 2009). The intended meaning of the NMCCs was determined based on the surrounding context. For corpora with accompanying audio or video files (i.e. *Tong* and *Erbaugh*), multimedia files were additionally consulted to ascertain the meaning. The NMCCs were categorised into ARGUMENT, ADJUNCT, or EXTENDED depending on the grammatical relations between the head noun and the modifying clause:

i) If the head noun falls into the broad categories of AGENT or PATIENT (including THEME, EXPERIENCER, etc.) of the event denoted by the verb phrase in the modifying clause, as in (9a), it is an *argument NMCC;*
ii) Failing (i), if the head noun expresses LOCATION, TIME, or INSTRUMENT of the event depicted in the modifying clause, as in (9b), then it is an *adjunct NMCC*. We delineate LOCATION and INSTRUMENT based on their functional significance in the event and preclude blanket inclusion of all entities related to the event. Specifically, for INSTRUMENT, we refer to entities that are considered necessary, predictable tools in the event. By LOCATION, we refer to "containment" (including loosely "encirclement") and "support" relations, but not other spatial relations such as occlusion, proximity, and attachment. This is because containment and support are two core spatial relations: Conceptually, event entities that contain or support the AGENT/PATIENT can have direct interactions with them and lead to physical consequences (e.g. moving the box containing the apple will move the apples as well). Empirically, containment and support emerge and mature the earliest in cognitive development (Casasola et al., 2003; Rigney & Wang, 2015); they are also labelled most successfully across languages by young children (Choi & Bowerman, 1991; Johnston & Slobin, 1979) and are provided most extensively in infant-directed speech (Casasola et al., 2017);
iii) Failing both (i) and (ii), if the head noun forms a relation with the modifying clause such as ACCESSORY, BY-PRODUCT, and EXTERNAL CAUSE, as in (9c), it is an *extended NMCC*.

(9) Examples of three types of NMCC in child-produced speech in our dataset.
a. Argument: [___$_i$ mai shuzi] de lao-taipo$_i$ (AGENT)
sell comb *de* old woman
"the old woman who sells combs" (XUE 3;9, *Zhou3*)

b. Adjunct: [__ chi fan] de zhuozi (LOCATION)
eat meal *de* table
"the table at which (people) eat meals" (4;10, *LiZhou*)

[1]The target Chinese character in the search is "的," excluding homophones sharing the same phonemic transcription *de*, such as "地," "得," "德," etc.

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

c. Context: in a role-play game, the child had received some change from the "cab driver" and turned to the investigator to offer the money to her.

Extended: [wo zuo jichengche] de qian (BY-PRODUCT)
1SG take taxi *de* money
Lit. "the change that (is related to) me taking the taxi"
"the change that I received after I took the taxi." (KANG 3;2, *Erbaugh*)

The first author coded the data independently. Then 20% of the extracted NMCCs were randomly selected and independently coded by the corresponding author. The agreement rate between the two authors was 91.43% (Cohen's $K = .832$, $p < .001$). When the two coders disagreed, they reviewed the coding and settled the disagreement with reference to the conversational contexts and the video recording. We extracted 728 tokens of NMCC, including 481 argument, 148 adjunct, and 99 extended NMCCs. Given the nature of the data, we present the cross-sectional and longitudinal data separately.

5.3. Cross-sectional data

A total of 525 NMCCs were extracted from the cross-sectional corpora and placed into six age groups, as shown in Table 2. The results show that NMCC is produced at a low frequency in child Mandarin, with fewer than 15 tokens per thousand utterances. An increase in productivity is observed between 1;2 and 4;11. Figure 1 illustrates the token number and proportion of each type of NMCC. Argument NMCCs are more frequent than adjunct and extended NMCCs across age groups. Argument and adjunct NMCCs were observed in the production of 2-year-olds, whereas extended NMCCs were only found in children aged 3;0 and above, suggesting later emergence of extended NMCC than the other two types.

5.4. Longitudinal data

In the longitudinal corpora with 14 children's speech data, 203 NMCCs were extracted and coded. Table 3 presents the raw token number, first emergence, and the percentage of each type, as recorded in respective corpora. As expected, there is wide variation across the corpora. Only five children produced more than 15 tokens of NMCC during the

Table 2. NMCC in the cross-sectional corpora

Group	Age	N_{NMCC}	$N_{utterances}$	N_{NMCC} per 1000 utterances
1 yr	1;2–1;11	0	1160	0
2 yrs	2;0–2;11	5	1585	3.15
3 yrs	3;0–3;11	58	11103	5.22
4 yrs	4;0–4;11	149	12267	12.15
5 yrs	5;0–5;11	158	14103	11.20
6 yrs	6;0–6;9	155	11077	13.99
Total		**525**	**51295**	

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

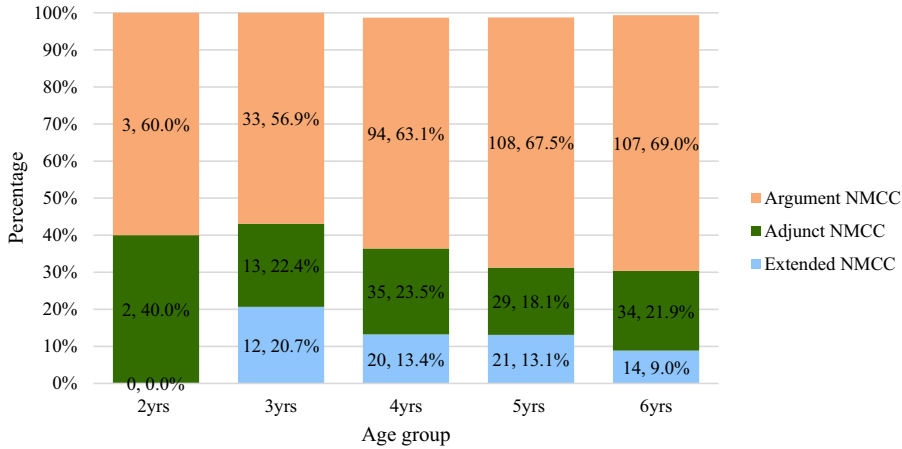


Figure 1. NMCC types in the cross-sectional corpora ("token number, percentage" in bars).

Table 3. NMCC produced in the longitudinal corpora

Child	Age	Raw token (token per 1000 utterances)	Age of 1st attested token (proportion)		
			Argument	Adjunct	Extended
Tong					
TONG	1;7–3;4	44 (4.76)	1;11 (40.9%)	2;2 (29.5%)	2;0 (29.5%)
Zhou 3					
XUE	1;8–5;5	30 (3.59)	2;0 (83.3%)	1;8 (10.0%)	3;8 (6.7%)
Erbaugh					
PANG	1;9–2;11	19 (2.24)	2;1 (78.9%)	2;6 (5.3%)	2;7 (15.8%)
KANG	2;10–3;9	40 (3.51)	2;10 (62.5%)	2;10 (30.0%)	3;2 (7.5%)
TCCM					
YANG	1;5–2;9	2 (1.04)	2;9 (100%)	–	–
XU	1;6–2;5	1 (.38)	–	2;5 (100%)	–
WU	1;7–2;10	3 (1.08)	1;9 (66.7%)	–	2;6 (33.3%)
PAN	1;7–3;9	8 (2.77)	2;6 (50.0%)	2;6 (25.0%)	2;0 (25.0%)
JC	2;2–3;5	10 (2.04)	2;6 (60.0%)	–	2;5 (40.0%)
WANG	2;5–3;4	13 (4.11)	2;5 (84.6%)	2;9 (15.4%)	–
CHOU	2;6–3;4	7 (1.37)	2;6 (71.4%)	–	3;2 (28.6%)
WUYS	2;7–3;10	8 (5.21)	2;7 (87.5%)	2;10 (12.5%)	–
CHENG	3;1–3;11	14 (4.11)	3;1 (92.9%)	–	3;4 (7.1%)
CHW	3;6–4;3	4 (1.53)	3;8 (75.0%)	–	3;10 (25.0%)

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

recording periods, all of them from the three "larger" corpora with greater numbers of total utterances per child (i.e. Tong, Zhou3, Erbaugh, see Table 1). This is consistent with previous observations and our own cross-sectional finding that NMCC is a low-frequency structure in naturalistic speech. Nevertheless, as a limitation of naturalistic samples, the earliest tokens may or may not be captured. Additionally, given the generally small numbers of analysable NMCC recorded for each child, patterns generalised from this small sample may be highly instable and thus lack generalisability. Hence, whether the order of emergence or frequency count as recorded in the longitudinal data is indicative of a developmental hierarchy is inconclusive. Notwithstanding, our longitudinal data clearly show that all individual children from the larger corpora spontaneously produced all three types of NMCCs at the age of 3, which is consistent with the patterns in the cross-sectional data, motivating us to test children as young as age 3 in Study 2.

5.5. Structural and semantic characteristics of early NMCCs

Some of the earliest argument, adjunct, and extended NMCCs are presented in (10)–(12), with the head nouns serving a variety of roles in the clausal events from AGENT and INSTRUMENT to BY-PRODUCT, EXTERNAL CAUSE, and ACCESSORY. In the argument NMCC in (10), the head noun is co-indexed with gapped subject position in the modifying clause (i.e. *huoche* "train"). In the adjunct and extended NMCCs in (11) and (12), the modifying clauses have syntactic gaps with unrealised, covert subject and objects, but the head nouns are not co-indexed with those gaps. The majority of early adjunct (95 out of 148) and extended (72 out of 99) NMCCs have covert clausal subjects, resulting in a simpler [V(-N)-*de*-head N] sequence than a full NMCC, as in (11) and (12a–b).

(10) Earliest argument NMCC.
[[___$_i$ bian-cheng zhe yangzi] de huoche$_i$].
change-become this appearance *de* train
"The train that changes into this appearance." (WU 1;9, *TCCM*)

(11) Earliest adjunct NMCC.
[[__ xie zi] de bi]. (INSTRUMENT)
write (Chinese) character *de* pen
"The pen that (one) writes (with)." (XUE 1;8)

(12) Earliest extended NMCC.
a. [[__ chui __] de shengyin]. (BY-PRODUCT)
blow *de* sound
"The sound that (results from one) blowing (candles)" (PAN 2;0, *TCCM*)

b. mama dian [[__ fangpi] de zhe-ge]. (EXTERNAL CAUSE)
mom click fart *de* this-CL
"Mom, click this button such that (we make the toy) fart." (TONG 2;4, *Tong*)
(Authors' note: the toy has a button which you can press to hear a farting sound.)

c. zhe shi [[tamen da-zhen] de ka]. (ACCESSORY)
this COP 3PL get-injection *de* card
Lit. "This is the card that (is related to) them getting vaccines."
"This is their vaccine record card." (3;0, *LiZhou*)

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

5.6. Interim summary

Study 1 investigated how Mandarin-learning children produce Mandarin NMCCs in naturalistic child speech. We analysed 728 tokens of NMCCs extracted from a total of 119,764 utterances produced by 369 children. The results show that NMCCs emerge quite early, typically at or before 2;0. Although the frequency is low at the early stages (less than 6 tokens per thousand utterances from 1;2 to 3;11), it increases with age (see Table 2), despite considerable individual differences (see Table 3). Different production patterns are found across the three types of NMCCs. Argument NMCCs make up a larger percentage in child speech than adjunct and extended NMCCs across age groups and individuals. In the cross-sectional data, argument and adjunct NMCC emerge at 2;0, earlier than extended NMCCs, which are produced after 3;0. Qualitative analysis showed that early NMCCs instantiate a variety of head–clause relations derived from the modifying event and the conversational context. An empty clausal argument could either be subject to a coreferential relationship with the head noun, or subject to pervasive argument dropping in Mandarin (Comrie, 1998).

Overall, our results indicate a primacy of argument NMCCs over the other types in language acquisition, consistent with the developmental hierarchy in (8). However, the naturalistic data present limitations: lower frequency and later emergence may be partially due to a lack of production opportunities; individual differences in NMCC production may be associated with uncontrolled child-internal factors as well as potential sampling gaps across the disparate corpora. An experimental study was therefore designed to remedy this weakness and test specific NMCC structures with Mandarin children controlling for NMCC-level and child-level factors.

6. Study 2: experimental study

6.1. Research questions and predictions

In the second study, we sought to test knowledge of Mandarin NMCCs in children aged 3–4 years, focusing on the role of conceptual prominence of the head noun's thematic role in the clausal event. To achieve this, in our experimental stimuli we controlled for word order, sentence length, and animacy of the head noun across NMCC types, narrowing down the targets to a subset of NMCC structures. Our experimental design was guided by the following research questions: How accurately do 3–4-year-old Mandarin children produce NMCCs in felicitous elicitation contexts? How do accuracy rates vary across NMCC types? How does the effect of NMCC type vary with age?

Since previous studies and our Study 1 show that children begin to spontaneously produce all three types of NMCCs between ages 2 and 3, we predict that some 3–4-year-olds will be able to produce all three types of NMCCs successfully in felicitous elicitation contexts. Nevertheless, age, WM, and general grammatical proficiency may play moderating roles. In cases where children fail to produce the target NMCC, they will likely substitute the complex, embedded construction with a simpler structure. Across NMCC types, we expect the highest accuracy rate in argument NMCCs, given the dominance of argument NMCCs in naturalistic input and the high prominence of AGENT and PATIENT in events. We also predict that children will perform better in adjunct NMCCs than in extended ones, considering the conceptual prominence scale in (7). However, the disparities in accuracy among NMCC types should diminish as children approach ceiling performance.

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

6.2. Methods

Participants. The participants were recruited from a public kindergarten in Shandong Province in northern China, where Mandarin is the dominant language. Priority was given to the youngest testable children in elicited production tasks, namely, 3- and 4-year-olds. 127 monolingual Mandarin-speaking children were invited to participate in the study, and 121 (3;1–4;11, 59 girls) completed all the tasks (*M*age in months = 49). No substantial exposure to other varieties of Chinese or other languages was reported by the kindergarten or the parents. This study was approved by the Survey and Behavioral Research Ethics Committee of The Chinese University of Hong Kong (No. SBRE-24-0072). Written consent was obtained from individual parents prior to data collection.

Tasks, materials, and procedures. A contextualised sentence repetition (SR) task was used to elicit NMCC. SR is a sensitive measure of children's grammatical knowledge (Lust et al., 1996) and has been used to investigate young children's knowledge of complex structures like RCs (Diessel & Tomasello, 2005; Kidd et al., 2007). Our task was presented as a "treasure hunt game" in which the children were required to repeat previously heard instructions (i.e. prompt sentences with NMCCs) to Peppa Pig (a popular cartoon character) so that Peppa could follow their instructions to find the treasures (e.g. diamond, pearl). The children were told that Peppa only took instructions from children, rather than adults, and that they should repeat the instructions that they heard from the experimenter to Peppa most clearly and accurately. An example of the prompt sentence and its preceding lead-in is presented in (14) and (13). A lead-in sentence spells out the target entity in this search (e.g. coin), followed by a test item (one of (14)) which is a prepositional phrase in the "*zai*-[NMCC]-locative" frame, identifying the location of the target entity. *Zai* is the only preposition used in the experiment, and the locatives include *shangmian* "top," *xiamian* "bottom," *limian* "inside," and *pangbian* "side." Previous studies have shown that 3-year-olds demonstrate productive use of *zai* and the four locative expressions used in our stimuli, and can interpret them in largely adult-like manners (Deng & Yip, 2016).

Note that the experimental design controlled for word order and head noun animacy across the three NMCC types to align with Study 2's focus on conceptual prominence, within the constraints of a limited item set. Only object RCs with inanimate PATIENT head nouns (i.e. [S-V-*de*-O]) were included to represent argument NMCCs. Subject RCs were avoided because their non-agent-first [V-O-*de*-S] word order and preference for animate heads (see (3) and (9)) would have rendered them considerably different from the adjunct and extended NMCCs and confounded the results (Hao et al., 2025). The adjunct NMCC items included INSTRUMENT and LOCATION heads, and the extended NMCC items ACCESSORY and BY-PRODUCT heads.

(13) Lead-in sentence.
Women lai kan zhe fu tu de jinbi zai nali.
1PL come see this CL picture *de* gold coin at where
"Let us see where the gold coin is in this picture."

(14) Test items (to be repeated by children)
a. Argument NMCC prompt in [S-V-*de*-O]:
Zai [xiongmao didi chui de lazhu]$_{\text{NMCC}}$ pangbian
at Panda brother blow *de* candle side
" Next to the candle that Brother Panda blows out."

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

b. Adjunct NMCC prompt in [S-V-O-*de*-N]:
Zai [xiongmao chui lazhu de yizi]$_{\text{NMCC}}$ xiamian
at Panda blow candle *de* chair bottom
"Under the chair where Panda blows out a candle."

c. Extended NMCC prompt in [S-V-O-*de*-N]:
Zai [xiongmao chui lazhu de zhaopian]$_{\text{NMCC}}$ shangmian
at Panda blow candle *de* photo top
Lit.: "Above the photo that (is related to) Panda blowing out a candle."
"Above the photo such that Panda blows out a candle."

To engage the young children and control for contextual factors which might affect the interpretation of the NMCCs, we created coloured pictures visualising the events denoted by the test items to serve as secondary prompts, as exemplified in Figure 2. Note that the visual prominence of the depicted objects (e.g. AGENT, PATIENT, INSTRUMENT, BY-PRODUCT), which may or may not be an integral part of conceptual prominence, was not deliberately controlled in the pictures – the objects were depicted in realistic proportion to one another. A total of 36 test items (3 NMCC types × 12 items per type) and accompanying pictures were created. We controlled for the number of syllables of the clausal verb (all monosyllabic, e.g. *chui* "to blow") and the NMCC prompts (11 syllables per sentence). To avoid priming effects among items depicting similar events (like those in (14)), we created three lists such that test items with the same clausal verb were distributed across the lists, rendering 12 test items per list. NMCC types were evenly distributed across lists. Additionally, each list included three "easy" filler items appearing in the [*zai*-[N1-*he*-N2]-locative] sequence, in which N1 and N2 were simple nouns coordinated by *he* "and." The filler items helped to ascertain whether the participants were engaged in the task. The SR task began with three additional NMCC warm-up items using verbs different from the critical items. The SR task lasted for 10–15 minutes in total. See Tables S3, S4 in Supplementary for the full experimental stimuli.

Since the acceptability of extended NMCCs is known to be conditioned by felicitous pragmatic contexts, a norming study was conducted to select the most appropriate test items. This norming study focused exclusively on extended NMCCs because, unlike argument and adjunct NMCCs, their well-formedness hinges on a loose semantic-pragmatic link between the head noun and the clausal event, which is not yet sufficiently



Figure 2. Sample pictures for three types of NMCC prompts in (14) (from left to right: argument, adjunct, extended).

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

understood in the literature and thus requires case-by-case norming. Prior to the SR task, 50 adult native Mandarin speakers (*M*age = 28.38, *SD* = 6.96) judged 24 sentences on the "plausibility" of the relevant event depicted by the sentences using a five-point Likert scale from 5 "completely acceptable" to 1 "completely unacceptable." Twelve extended NMCCs that were rated above 3 were selected for the SR task. Their mean ratings are included in Table S4 in Supplementary.

To assess the prominence of the objects depicted in the pictures during language encoding, we asked 45 adult native Mandarin speakers (*M*age = 30.60 yrs) to describe each of the 12 pictures within one single sentence using the intended AGENT as the sentential subject (e.g. panda in Figure 2). In the 452 valid responses, 95.58% encoded the PATIENT (e.g. candle in Figure 2); 24.78% encoded TIME, LOCATION, and INSTRUMENT (e.g. chair in Figure 2). Only 10.84% encoded entities that are ACCESSORY or BY-PRODUCT of the events (e.g. photo on the wall in Figure 2). This descending order of object encodability aligns with the conceptual prominence scale and validates the effectiveness of our visual materials.

In addition to age, children's memory span and general linguistic knowledge are central to their ability to reconstruct sentences for repetition (Frizelle et al., 2017). We measured children's WM using the Mandarin version of the Comprehensive Test of Phonological Processing, Memory for Digits (Wagner et al., 1999) and their grammatical knowledge using the Mandarin Receptive Grammar Test (Zhou et al., 2025). In the WM task, children repeated strings of numbers varying in length from 2–8 digits/syllables. The task terminated when the child provided three incorrect repetitions consecutively. In the receptive grammar (RG) task, children listened to a sentence and pointed to the matching picture among four given pictures in 51 items testing 19 prominent grammatical structures in Mandarin such as negation, classifier, and passive (see Supplementary Table S4 in Zhou et al., 2025 for the list of structures tested).

The children were individually tested in a quiet classroom at the kindergarten. The tasks were administered by the first author and a research assistant, both native Mandarin speakers with postgraduate training in linguistics. Typically, it took 20–30 minutes for a child to complete all the tasks. Children's responses were audio-recorded for transcription and analysis.

Scoring and data analysis. For the SR, a repetition was considered "correct" only when it was identical to the original prompt, or contained minor changes without altering the meaning and structure of test items (see coding scheme in Supplementary Table S5). Preliminary scoring was conducted by individual experimenters on site, and the first author double-checked the scoring against the recording and discussed the discrepancies until consensus was reached. The WM and RG tasks were scored following the respective manuals. Data analysis was conducted in R (R Core Team, 2024).

6.3. Quantitative results

In the SR task, 116 out of the 121 children performed 100% accurately in the three fillers; four of them provided correct responses in two out of three fillers; one of them only responded to one filler item correctly. We removed the last child, reducing the sample size to 120. Table 4 presents the raw scores in the tests. In SR, the children performed most accurately in argument NMCCs (M = 3.33, SD = 1.01), followed by adjunct NMCCs (M = 2.68, SD = 1.47), with extended NMCCs exhibiting the lowest accuracy (M = 2.2, SD = 1.55). Paired-sample tests revealed significant differences between argument and adjunct NMCC (t = 5.45, p < .001) and between adjunct and extended NMCC (t = 5.03,

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

Table 4. Descriptive statistics of performance scores in Study 2

	Mean	*SD*	Range
Sentence repetition (SR) task			
Argument NMCC (max = 4)	3.33	1.01	0–4
Adjunct NMCC (max = 4)	2.68	1.47	0–4
Extended NMCC (max = 4)	2.20	1.55	0–4
SR total (max = 12)	8.21	3.50	0–12
Working memory (WM) (*max* = 21)	14.82	3.31	8–21
Receptive grammar (RG) (*max* = 51)	35.72	6.67	14–48

$p < .001$). Overall, the children displayed an average accuracy rate of 68% (8.21 out of 12) in the SR task. In the WM task, only four children performed at ceiling (8 syllables) and scored 21. For RG, the children obtained an average score of 35.72 out of 51 items (70%) with no ceiling cases.

To examine the extent to which experimental conditions (NMCC type) and participant characteristics (age, WM, RG) predict the item-level binary responses (1 = correct; 0 = incorrect) in SR, we fitted generalised linear mixed-effects models (GLMMs) using *lme4* in R. An empty means (no predictor) model with participant ($N = 120$) and item (target NMCC, $N = 36$) as random intercepts had an intraclass correlation of .69, in which 52.4% of the variance in accuracy was due to participant mean differences and 16.6% due to item mean differences. Children's age (scaled), WM, and RG scores were entered as control variables, which significantly improved the empty means model, $-2\Delta LL(3) = 99.70$, $p < .001$. NMCC type (argument, adjunct, extended) and its interaction with age were entered as fixed effects, which further improved the control variable model, $-2\Delta LL(4) = 38.82$, $p < .001$. The variance inflation factor (VIF) showed low multicollinearity of predictors (VIFs ranging from 1.03–1.76) in the last model. Results from the GLMM indicated a significant main effect of NMCC type: adjunct NMCCs have significantly lower item accuracy than argument NMCCs ($\gamma = 1.49$, $p = .001$) and significantly higher accuracy than extended NMCCs ($\gamma = -.90$, $p = .03$). Table 5 summarises the GLMM analysis.

Figure 3 visualises the predicted probability of children producing the target NMCC correctly across age. Older children ($\gamma = .52$, $p = .017$) are more likely to produce accurate responses than younger children. Although the 3-year-olds showed a below-chance probability of producing the target extended NMCCs and a slightly above-chance probability of producing the target adjunct NMCCs, accuracy rates in both types increased sharply with age. The interaction between NMCC type and age is only significant between argument and adjunct NMCCs ($\gamma = -.66$, $p = .003$), suggesting that the performance difference between these two types narrows with age. Such interaction effect is not significant between adjunct and extended NMCCs ($\gamma = -.1$, $p = .631$), probably because both types grow considerably from 3;0 to 5;0 at similar rates and neither has reached ceiling by 5;0.

6.4. Analysis of errors

Of the 1440 responses, 451 responses were coded as incorrect. Within this subset, 54 were null or incomprehensible (NULL) and were excluded, leaving 397 analysable tokens.

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

Table 5. GLMM analysis of Sentence Repetition (fixed effects: NMCC type, Age, NMCC type~Age, working memory, receptive grammar; response variable: accuracy score in SR)

Coefficient	*γ*	*Std. Error*	*z* value	*p*
(Intercept)	−9.28	1.29	−7.17	<.001
NMCC type (adjunct versus argument)	1.49	.43	3.48	.001
NMCC type (adjunct versus extended)	−.90	.41	−2.17	.030
Age	.52	.22	2.39	.017
Working memory	.44	.06	7.99	<.001
Receptive grammar	.12	.03	3.96	<.001
NMCC type (adjunct versus argument):Age	−.66	.22	−3.01	.003
NMCC type (adjunct versus extended):Age	.10	.20	.48	.631

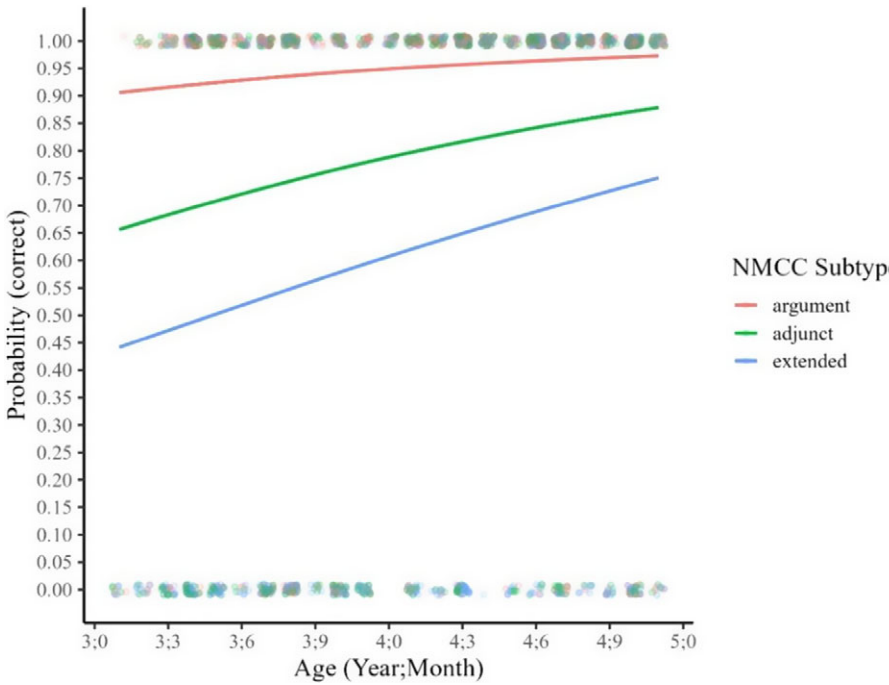


Figure 3. Probability of correct responses (= 1) in each NMCC type across ages (scores in working memory and receptive grammar tests centred by means).

From these, 48 instances were grammatical and plausible paraphrases of the target NMCCs (PRPH), including change of NMCC type from a lower type to a higher type (from right to left in the hierarchy), as in (15). Reverse change of NMCC type from a higher type to a lower type was not attested.

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

(15)

Target NMCC:	zai	[xiaomao	chi	mianbao]	de	guojiang	limian	
(extended)	at	Kitten	eat	bread	*de*	jam	inside	
	"(it's) in the jam such that Kitten eats bread"							

Response:	zai	[xiaomao	chi	guojiang]	de	beizi	limian
(adjunct NMCC)	at	Kitten	eat	jam	*de*	mug	inside
	"(it's) in the mug from which Kitten eats jam"						

From the 397 analysable tokens, 349 were incomplete, ungrammatical repetitions (IR) where the children dropped one or more lexical items in the target NMCC. Half of them (50.14%) appeared in a non-embedded [N-V-N] sequence; some (22.34%) were verbless sequences with a string of one to three nouns, as in [N(-N)(-N)]; others (9.17%) were structured as [N-*de*-N], forming a possessive relation with the head noun. For the adjunct and extended NMCC items, which were presented in the [$N1_S$-V-$N2_O$-*de*-N3] sequence in the prompts, the IR responses routinely dropped either the head noun (N3) or the clausal object (N2, 90.26%), while retaining the clausal subject (N1, 82.7%), which is the semantic AGENT across the prompts. An example is shown in (16). A chi-square test shows that the proportions of each type of incorrect responses (NULL, PRPH, IR) do not differ across NMCC types, X^2 (4, N = 451) = 5.38, p = .25.

(16)

Target NMCC:	zai	[$yazi_{N1}$	xi	$wazi_{N2}$]	de	$feizao_{N3}$	pangbian	
(adjunct)	at	Duck	wash	sock	*de*	soap	side	
	"(it's) beside the soap with which Duck washes socks"							

Response:	zai	$yazi_{N1}$	xi	$wazi_{N2}$	de	pangbian
	at	Duck	wash	sock	*de*	side
	"(it's) beside Duck washing socks"					

6.5. Interim summary

Study 2 tested 3–4-year-old children's knowledge of Mandarin NMCCs in a contextualised SR task, with age, WM, and general RG as control variables in GLMM models. The results showed a significant main effect of NMCC type and an advantage for argument over adjunct, and adjunct over extended NMCCs, after age, WM, and general RG were controlled for, conforming to the hierarchy in (8). The adjunct and extended NMCCs were produced slightly above or below chance levels by the youngest children, but the accuracy rates increased sharply with age.

Since the stimuli were controlled for syllable count, lexicon, visualisation, word order, and head noun animacy across NMCC types, the sole systematic variation was the conceptual prominence of the head noun's role in the event. Moreover, given that the 11-syllable elicitation prompts exceeded the children's phonological memory capacity (as evidenced by the digit span task where only 4 of 120 preschoolers reached the ceiling of 8 syllables), their repetitions should reflect their grammatical knowledge of NMCC, rather than rote parroting. One could argue that the argument NMCC prompts were cognitively less demanding than the adjunct and extended NMCC ones, because the former require holding only two event entities (agent, patient) in memory and the latter three (e.g. agent,

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

patient, instrument). Nevertheless, the difference is inherent to the issues under investigation – the event structure of adjunct and extended NMCCs is inherently more complex and nuanced. A methodological note is that we did not employ the traditional picture-identification task, in which the child produces an RC in order to identify the intended referents from their competitors. Picture identification would provide felicitous contexts for NMCC production. However, pilot testing revealed that the concurrent demands of producing full RCs and assessing visual competitors overwhelmed our youngest participants.

Our error analysis provides further insights into two dimensions. First, children only produced errors in one direction: they sometimes substituted a higher NMCC type for a lower one (e.g. argument for extended), but never the opposite. Second, when children encountered NMCCs beyond their grammatical capacity, they often dropped a postverbal noun (N2 or N3) in [$N1_S$-V-$N2_O$-*de*-N3], but consistently retained the clausal subject N1 and produced [N1-V] instead. This is notable because N1, referring to the only animate agent in the picture (e.g. the panda in Figure 2), had little referential value in identifying the agent in the SR task due to the lack of semantic competitors. Its retention is likely attributable to the canonical status of the [N-V] sequence in early sentence parsing (Diessel & Tomasello, 2005; Slobin & Bever, 1982) and the prominence of agents in incremental processing (Sauppe et al., 2023) and visual perception (Cohn & Paczynski, 2013). Both error patterns are consistent with the predictions of the proposed developmental hierarchy.

7. General discussion

7.1. Developmental hierarchy of Mandarin NMCC: a summary of Studies 1 and 2

We have reported the first comprehensive examination of the full range of Mandarin NMCCs in early childhood acquisition, analysing naturalistic production in existing corpora and elicited imitation in a controlled experiment. As summarised in Table 6, both studies showed that in Mandarin, argument, adjunct, and extended NMCCs develop at different rates. Specifically, argument NMCCs show an advantage over the other two: they emerged the earliest and were produced most frequently (Study 1) and accurately (Study 2). Adjunct NMCCs displayed significantly different patterns from the other two types in two measures (frequency in naturalistic production and accuracy in SR), lending support to their status as the "middle" type in the hierarchy. Extended NMCCs were identified as the most challenging type across analyses. The quantitative results largely conform to the developmental hierarchy proposed in (8), and the children's substitution and omission errors in the repetition task were consistent with the hierarchy.

Our hierarchy is predicated on the role of conceptual prominence in acquiring NMCCs, among others. Conceptually there exist more event roles beyond those listed in (7), and even those already listed in (7) can be further differentiated into finer-grained roles such as THEME, EXPERIENCER, and RECIPIENT. The scale in (7) is only preliminary. Future research in event cognition is essential to elaborate and refine the scale. Although we treat conceptual prominence as independent of grammatical function, the two are interrelated: agent and patient roles typically map onto core, obligatory verbal arguments (subjects/objects), while location and instrument are often encoded as non-core, optional adjuncts (Boland, 2005). Our developmental hierarchy for Mandarin NMCCs aligns most closely with Pan's (2022) Mandarin RC Recoverability Hierarchy (Argument > Adjunct > Gapless), which is based on the recoverability of a syntactic gap or a covert semantic

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

Table 6. Summary of findings in both studies

Study	Data	Measure	Findings				
1	Cross-sectional	Emergence	Argument	&	Adjunct	>	Extended
		Frequency	Argument	>	Adjunct	>	Extended
	Longitudinal	Frequency	Argument	>	Adjunct	>	Extended
2	Repetition	Accuracy	Argument	>	Adjunct	>	Extended
		Errors	[N-V] and AGENT advantage				
Proposed developmental hierarchy:			*Argument*	>	*Adjunct*	>	*Extended*

variable in the clause associated with the head noun. A key difference, however, is that our hierarchy is motivated by multiple factors critical to early language development (including but not limited to the status of the linking particle, input frequency, thematic prominence, and word order). Our model thus offers an analytical framework with a set of parameters that can be applied to other languages with differing configurations. Below we will return to this in our discussions of the high and low ends of the hierarchy (early emergence of argument NMCC and late acquisition of extended NMCC).

7.2. Early emergence of NMCC in Mandarin

Our analysis of naturalistic child Mandarin speech (1;2–6;9) reveals that NMCCs emerge by 2;0, notably earlier than RCs in languages like English. This early emergence can be attributed to the structural unity of noun modification in Mandarin. Unlike English, which employs prenominal adjectives and postnominal RCs, Mandarin places adjectives, possessives, and NMCCs consistently prenominally, all linked by the particle *de*. We posit that children's early acquisition of simpler prenominal modifiers (e.g. adjectives) facilitates their subsequent grasp of the more complex NMCC structure, consistent with the developmental trajectory from simple to complex constructions (Diessel & Tomasello, 2000). This structural isomorphism provides a scaffold, explaining the earlier onset of Mandarin NMCCs compared to RCs in other languages (Lai et al., 2023).

Although Cantonese and Mandarin NMCCs are structurally similar, our findings on Mandarin acquisition diverge from prior Cantonese studies. Lai et al. (2023) reported Cantonese NMCCs emerging around 2;2–2;5, approximately 6 months later than in our Mandarin data. This discrepancy may stem from methodological or linguistic factors. Methodologically, Lai et al.'s smaller Cantonese sample size, due to data scarcity, may have failed to capture the earliest productions. Linguistically, while Mandarin NMCCs appear in a single form with an overt and obligatory linking particle *de*, Cantonese NMCCs exhibit greater complexity and diversity, appearing in two different forms (the classifier form and the linking particle form). The classifier form is more common in colloquial Cantonese and arguably more frequent in child-directed speech, yet realised through a range of specific classifiers depending on specific head nouns (details in Yip & Matthews, 2007:159). This more complex form-to-function mapping likely elevates acquisition difficulty, explaining the later emergence and lower frequency of Cantonese NMCCs.

The presence of a general noun-modifying clause construction and the dependence of NMCC construal on an aggregate of grammatical, semantic, and pragmatic factors are not

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

unique to Mandarin. The facilitative effect of a uniform prenominal modification system may extend to languages like Japanese and Korean, where both lexical and clausal modifiers are prenominal. However, Japanese and Korean lack overt linking particles like *de* in Mandarin, which may weaken the facilitative effect. Furthermore, whether the same argument > adjunct > extended ordering applies to these languages will critically depend on multiple factors. For example, the canonical SOV word order in Japanese and Korean introduces a word order factor. In these languages, adjunct and extended NMCCs typically maintain the canonical [S-O-V-Head] sequence, whereas argument NMCCs deviate from the canonical SOV order and appear in either [S-V-O] (for object RC) or [O-V-S] order (for subject RC). This configuration creates a unique test case: while conceptual prominence predicts an advantage in acquiring argument NMCCs, there is a mismatch between NMCC and the canonical word order that may attenuate the advantage. Studying these languages can evaluate the interaction of these competing factors.

7.3. Late acquisition of extended NMCC

Extended NMCCs are licensed through semantic-pragmatic inferencing between the head noun and its modifying clause in Mandarin NMCCs (Cheng & Sybesma, 2006; LaPolla, 2017; Matthews & Yip, 2002). In our naturalistic speech data, we found that children above age 3 can successfully establish such relations in extended NMCCs based on encyclopedic, pragmatic knowledge about events, producing a non-trivial number and variety of extended NMCCs (Study 1). Nevertheless, the experimental results showed that the probability of reproducing an accurate extended NMCC is below .45 at 3;0 and approaches only .70 at 4;0, suggesting that 3–4-year-olds have not mastered this type. Another look at the corpus and experimental data revealed that two differences might have led to the discrepancy: (i) semantically, early extended NMCCs in spontaneous speech are restricted in function to express generic events associated with the head noun, whereas extended NMCC prompts in the SR had a specific clausal event illustrated by pictures; and (ii) syntactically, early extended NMCCs in naturalistic speech had a simpler [V(-N)-*de*-N] structure, whereas the extended NMCC prompts in the SR were fully fledged NMCCs with overt subjects and objects.

The excerpts in (17) and (18) from our corpus data illustrate the aforementioned semantic and syntactic differences. In (17) the child identifies a toy based on how it is used (one plays by pressing the little bear on the toy). Note that the clausal subject position is empty and receives a generic reading "one" or "people" in context. In (18), the child produced an extended NMCC in response to a *wh*-question "what happened in the forest?" Both NMCCs are innovative uses, which led to conversational repairs (mother's clarification question in (17) and experimenter's recast in (18)). Since the NMCC does not impose strict grammatical relations between the head noun and the modifying clause, it offers children a convenient structural frame to build up their prenominal modifiers, which sometimes leads to non-adult-like usages, showing children's developing competence in making semantic-pragmatic inferences in extended NMCCs.

(17) Meaning negotiation between child and mother regarding an extended NMCC.

CHI:	haiyou	na-ge	[[__	**an**	**xiao**	**xiong]**	**de na-ge]** .
	and	that-CL		**press**	**little**	**bear**	***de*** **that-CL**

"And the one (toy) such that (one can) press the little bear."

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

MOT: an xiao xiong de shenme dongxi a?
"What do you mean by pressing the little bear?"
CHI: jiushi na huang de che shang.
"It is on that yellow car."
MOT: o, you ge xiao xiong gei an xiaqu, shiba?
"Oh, there is a little bear to be pressed, right?"

(Tong, 2;4)

(18) Non-adult-like use of an extended NMCC by child.
EXP: senlin fasheng shenme shiqing?
"What happened in the forest?"
CHI: fasheng [[__ **zai dajia**] **de shi**] .
happen **ASP fight** ***de*** **thing**
"The thing such that (the animals) fighting is happening."
EXP: o, tamen zai da jia.
"Oh, they are fighting."

(5;7, *Chang1*)

8. Conclusions

This study presents the first developmental investigation of the full spectrum of NMCCs in Mandarin, testing the proposed acquisition hierarchy of argument > adjunct > extended NMCC. This hierarchy is theorised to be influenced by the conceptual prominence of the head noun's role in the modifying event, in addition to syntactic and frequency factors. Methodologically, we combined an exhaustive analysis of naturalistic CHILDES data with a controlled SR task. Our findings reveal that Mandarin NMCCs, as a group of syntactic constructions, emerge early. We argue that this early emergence is facilitated by the construction's uniform morphosyntactic form, which is consistently marked by a prenominal linking particle. Our quantitative and qualitative results confirm the predicted developmental hierarchy: Not only did argument NMCCs demonstrate primacy over both adjunct and extended types in terms of frequency and accuracy, but adjunct NMCCs also showed a significant advantage over extended NMCCs through the same measures, confirming our predictions that NMCCs with conceptually more prominent heads are acquired earlier than those with less prominent heads in Mandarin.

Mandarin-learning 3- and 4-year-olds demonstrate the ability to form fully grammatical argument NMCCs, showing mastery of syntactic dependencies and filler–gap relations. Meanwhile, their gradual acquisition of adjunct and extended NMCCs, which are built on nuanced event semantics and pragmatic knowledge, reveals that their acquisition of NMCC is not confined to syntactic filler–gap relations. This progression underscores a powerful human language learning mechanism in acquiring complex constructions. Our work establishes a foundation for cross-linguistic comparison, calling for parallel investigations in other languages with a general noun-modifying clause construction, such as Japanese and Korean.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S0305000926100750.

Data availability statement. The numeric data that support the findings of this study and the analysis codes are available from: https://osf.io/6bu3j/overview?view_only=fb23ef5fd8db42baabe02dba382f6bdb

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

Acknowledgements. We are grateful for the Child Language Data Exchange System (CHILDES) and the research teams who contributed Mandarin child speech samples to CHILDES, including our own team at Childhood Bilingualism Research Centre at the Chinese University of Hong Kong (CUHK). Without the rich corpus data, our first study would not have been possible. We thank all the children and teachers of Fushan Kindergarten in China for their participation in our experimental study. This article is an extension of Mengyao Shang's doctoral thesis project. Many colleagues and students shared their experience, suggestions, and constructive feedback over the years, especially Angel Chan, Lawrence Cheung, Haihua Pan, Zhenguang Cai, Jingyao Liu, as well as conference attendees at the 13th and 14th International Symposium on Bilingualism and the 16th International Congress for the Study of Child Language. This study is partially funded by a New Faculty Start-up Grant at CUHK awarded to Ziyin Mai. All remaining errors are our own.

Competing interests. The authors declare none.

Disclosure of use of AI tools. This paper did not use any AI tools to generate, analyse, or present ideas, words, data, or other materials.

Statement of ethical approval. This study was approved by the Survey and Behavioral Research Ethics Committee of The Chinese University of Hong Kong (No. SBRE-24-0072). Written consent was obtained from individual parents prior to data collection.

References

Boland, J. E. (2005). Visual arguments. *Cognition*, **95**(3), 237–274.

Brandt, S., **Kidd, E.**, **Lieven, E.**, & **Tomasello, M.** (2009). The discourse bases of relativization: An investigation of young German and English-speaking children's comprehension of relative clauses. *Cognitive Linguistics*, **20**(3), 539–570.

Casasola, M., **Bhagwat, J.**, **Doan, S. N.**, & **Love, H.** (2017). Getting some space: Infants' and caregivers' containment and support spatial constructions during play. *Journal of Experimental Child Psychology*, **159**, 110–128. https://doi.org/10.1016/j.jecp.2017.01.012

Casasola, M., **Cohen, L. B.**, & **Chiarello, E.** (2003). Six-month-old infants' categorization of containment spatial relations. *Child Development*, **74**(3), 679–693.

Chang, C. (1998). The development of autonomy in preschool mandarin Chinese-speaking children's play narratives. *Narrative Inquiry*, **8**(1), 77–111.

Chang, C. (2000). *Narrative performance across contexts and over time: Preschool Chinese children and mothers* [Unpublished doctoral dissertation]. Harvard Graduate School of Education.

Chen, J., & **Shirai, Y.** (2015). The acquisition of relative clauses in spontaneous child speech in mandarin Chinese. *Journal of Child Language*, **42**(2), 394–422. https://doi.org/10.1017/S0305000914000051

Cheng, L., & **Sybesma, R.** (2006). A Chinese relative. In H. Broekhuis, N. Corver, R. Huybregts, U. Kleinhenz, & J. Koster (Eds.), *Organizing grammar: Linguistic studies in honor of Henk van Riemsdijk* (pp. 69–76). De Gruyter Mouton.

Cheung, H., **Chang, C.**, **Ko, H.**, & **Tsay, J.** (2011). *Taiwan corpus of child Mandarin (TCCM).* NSC96-2420-H-002-030. http://doi.org/10.21415/T5Z616

Choi, S., & **Bowerman, M.** (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, **41**(1–3), 83–121.

Cohn, N., & **Paczynski, M.** (2013). Prediction, events, and the advantage of agents: The processing of semantic roles in visual narrative. *Cognitive Psychology*, **67**(3), 73–97. https://doi.org/10.1016/j.cogpsych.2013.07.002

Comrie, B. (1998). Attributive clauses in Asian languages: Towards an areal typology. In W. Boeder, C. Schroeder, K. H. Wagner, & W. Wildgen (Eds.), *Sprache in Raum und Zeit: In memoriam Johannes Bechert* (Vol. **2**, pp. 51–60). Gunter Narr.

Deng, X., & **Yip, V.** (2016). Cognition and perception in the linguistic encoding of space in child mandarin. *Journal of Chinese Linguistics*, **44**(2), 287–325.

Deng, X., & **Yip, V.** (2018). A multimedia corpus of child mandarin: The Tong corpus. *Journal of Chinese Linguistics*, **46**(1), 69–92.

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

Diessel, H., & **Tomasello, M.** (2000). The development of relative clauses in spontaneous child speech. *Cognitive Linguistics*, **11**, 131–151.

Diessel, H., & **Tomasello, M.** (2005). A new look at the acquisition of relative clauses. *Language*, **81**(4), 882–906.

Erbaugh, M. S. (1992). The acquisition of mandarin. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. **3**, pp. 373–455). Lawrence Erlbaum.

Ferretti, T. R., **McRae, K.**, & **Hatherell, A.** (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, **44**(4), 516–547.

Frizelle, P., **O'Neill, C.**, & **Bishop, D. V. M.** (2017). Assessing understanding of relative clauses: A comparison of multiple-choice comprehension versus sentence repetition. *Journal of Child Language*, **44**(6), 1435–1457. https://doi.org/10.1017/S0305000916000635

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, **68**(1), 1–76.

Hafri, A., **Trueswell, J. C.**, & **Strickland, B.** (2018). Encoding of event roles from visual scenes is rapid, spontaneous, and interacts with higher-level visual processing. *Cognition*, **175**, 36–52. https://doi.org/10.1016/j.cognition.2018.02.011

Hao, J., **Chondrogianni, V.**, & **Sturt, P.** (2025). Sources of children's difficulties with non-canonical sentence structures: Insights from mandarin. *Journal of Child Language*, **52**(6), 1295–1322. https://doi.org/10.1017/S0305000924000424

Hawkins, J. A. (1999). Processing complexity and filler-gap dependencies across grammars. *Language*, **75**(2), 244–285.

He, W., **Xu, N.**, & **Ji, R.** (2017). Effects of age and location in Chinese relative clauses processing. *Journal of Psycholinguistic Research*, **46**(5), 1067–1086. https://doi.org/10.1007/s10936-017-9480-4

Hsu, C. N., **Hermon, G.**, & **Zukowski, A.** (2009). Young children's production of head-final relative clauses: Elicited production data from Chinese children. *Journal of East Asian Linguistics*, **18**, 323–360.

Huang, C. (2016). The syntax and semantics of Prenominals: Construction or composition? *Language and Linguistics*, **17**(4), 431–475.

Huang, C., **Li, Y.**, & **Li, Y.** (2009). *The syntax of Chinese (Cambridge syntax guides)*. Cambridge University Press.

Johnston, J. R., & **Slobin, D. I.** (1979). The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language*, **6**(3), 529–545.

Keenan, E. L., & **Comrie, B.** (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, **8**(1), 63–99.

Kempson, R., & **Kurosawa, A.** (2009). At the syntax pragmatics Interface: Japanese relative clause construal. In H. Hoshi (Ed.), *The dynamics of the language faculty: Perspectives from linguistics and cognitive neuroscience* (pp. 47–84). Kurosio.

Kidd, E., **Brandt, S.**, **Lieven, E.**, & **Tomasello, M.** (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes*, **22**(6), 860–897.

Kurosawa, A. (2003). The interaction of syntax and pragmatics: The case of Japanese 'gapless' relatives. In K. M. Jaszczolt & K. Turner (Eds.), *Meaning through language contrast* (Vol. **1**, pp. 307–334). John Benjamins.

Lai, J. M., **Chan, W. S. A.**, & **Matthews, S.** (2023). Beyond relative clauses: The development of noun-modifying clause constructions in Cantonese. In W. Han & C. Brebner (Eds.), *Typical and atypical language development in cultural and linguistic diversity* (pp. 40–63). Routledge.

LaPolla, R. J. (2017). Noun-modifying clause constructions in Sino-Tibetan languages. In Y. Matsumoto, B. Comrie, & P. Sells (Eds.), *Noun-modifying clause constructions in language of Eurasia: Rethinking theoretical and geographical boundaries* (pp. 91–103). John Benjamins.

Lau, E., & **Tanaka, N.** (2021). The subject advantage in relative clauses: A review. *Glossa: a journal of general linguistics*, **6**(1), 34.

Lee, J. S., & **Lee, C.** (2012). Gap in" gapless" relative clauses in Korean and other Asian languages. In *Proceedings of the 26th Pacific Asia conference on language, information and computation* (pp. 640–645). Bali: Waseda University.

Li, L. H., & **Zhou, J.** (2008). *Metacommunication development of children aged from three to six during their collaborative pretend play* [Unpublished master thesis]. East China Normal University.

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

Li, X. Y., & **Zhou, J.** (2004). *The effects of pragmatic skills of mothers with different education on children's pragmatic development* [Unpublished master thesis]. Nanjing Normal University.

Liu, Z., & **Chan, A.** (2012). The role of qualia structure in mandarin children acquiring noun-modifying constructions. In *Proceedings of the 26th Pacific Asia conference on language, information and computation* (pp. 632–639). Faculty of Computer Science, Universitas Indonesia.

Lust, B., **Flynn, S.**, & **Foley, C.** (1996). What children know about what they say: Elicited imitation as a research method for assessing children's syntax. In D. McDaniel, H. S. Cairns, & C. McKee (Eds.), *Methods for assessing children's syntax* (pp. 55–76). MITPress.

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

Mai, Z., **Shang, M.**, **Liu, J.**, **Yan, S.**, **Matthews, S.**, & **Yip, V.** (2024) *Acquiring Chinese in US, Hong Kong and Beijing:Three new corpora and three verbal structures.* Paper presented at the XVIth International Congress for the Study of Child Language (IASCL-2024), Prague, Czech Republic. http://doi.org/10.21415/R1D0-BY71

Matsumoto, Y. (2017). General noun-modifying clause constructions in Japanese. In Y. Matsumoto, B. Comrie, & P. Sells (Eds.), *Noun-modifying clause constructions in languages of Eurasia: Rethinking theoretical and geographical boundaries* (pp. 23–43). John Benjamins.

Matsumoto, Y., **Comrie, B.**, & **Sells, P.** (2017). *Noun-modifying clause constructions in languages of Eurasia: Rethinking theoretical and geographical boundaries*. John Benjamins.

Matthews, S., & **Yip, V.** (2002). Relative clauses in early bilingual development: Transfer and universals. In A. G. Ramat (Ed.), *Typology and second language acquisition* (pp. 39–81). Mouton de Gruyter.

Matthews, S., & **Yip, V.** (2017). Noun-modifying clauses in Cantonese. In Y. Matsumoto, B. Comrie, & P. Sells (Eds.), *Noun-modifying clause constructions in languages of Eurasia: Rethinking theoretical and geographical boundaries* (pp. 105–120). John Benjamins.

Ming, T., & **Chen, L.** (2010). A discourse-pragmatic study of the word order variation in Chinese relative clauses. *Journal of Pragmatics*, **42**(1), 168–189.

Mintz, T. H., & **Gleitman, L. R.** (2002). Adjectives really do modify nouns: The incremental and restricted nature of early adjective acquisition. *Cognition*, **84**(3), 267–293.

O'Grady, W. (1997). *Syntactic development*. University of Chicago Press.

O'Grady, W. (2011). Relative clauses: Processing and acquisition. In E. Kidd (Ed.), *The acquisition of relative clauses: Processing, typology and function* (pp. 13–38). John Benjamins.

Ozeki, H., & **Shirai, Y.** (2007). The consequence of variation in the acquisition of relative clauses: An analysis of longitudinal production data from five Japanese children. In Y. Matsumoto, D. Oshima, O. Robinson, & P. Sells (Eds.), *Diversity in language: Perspectives and implications* (pp. 243–270). CSLI.

Packard, J. L. (1988). The first-language Acquisition of Prenominal Modification with DE in mandarin. *Journal of Chinese Linguistics*, **16**(1), 31–54.

Pan, H. (2022). Remarks on gapless relative clauses and complement clauses in mandarin Chinese. *Asian Languages and Linguistics*, **3**(1), 109–140.

R Core Team. (2024). *R: A language and environment for statistical computing* [Software]. R Foundation for Statistical Computing. https://www.R-project.org/

Radford, A. (2016). Analysing English sentences, *Cambridge textbooks in linguistics* (2nd ed.). Cambridge University Press.

Radvansky, G. A., **O'Rear, A. E.**, & **Fisher, J. S.** (2017). Event models and the fan effect. *Memory & Cognition*, **45**(6), 1028–1044. https://doi.org/10.3758/s13421-017-0713-4

Radvansky, G. A., **Zwaan, R. A.**, **Federico, T.**, & **Franklin, N.** (1998). Retrieval from temporally organized situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **24**(5), 1224–1237.

Rigney, J., & **Wang, S. H.** (2015). Delineating the boundaries of infants' spatial categories: The case of containment. *Journal of Cognition and Development*, **16**(3), 420–441.

Sauppe, S., **Naess, Å.**, **Roversi, G.**, **Meyer, M.**, **Bornkessel-Schlesewsky, I.**, & **Bickel, B.** (2023). An agent-first preference in a patient-first language during sentence comprehension. *Cognitive Science*, **47**(9), e13340. https://doi.org/10.1111/cogs.13340

Slobin, D. I., & **Bever, T. G.** (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition*, **12**(3), 229–265.

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press

Swallow, K. M., **Zacks, J. M.**, & **Abrams, R. A.** (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, **138**(2), 236–257. https://doi.org/10.1037/a0015631

Tardif, T., **Fletcher, P.**, **Liang, W.**, **Zhang, Z.**, **Kaciroti, N.**, & **Marchman, V. A.** (2008). Baby's first 10 words. *Developmental Psychology*, **44**(4), 929–938. https://doi.org/10.1037/0012-1649.44.4.929

Tsai, W.-T. D. (1997). On the absence of island effects. *Tsing Hua Journal of Chinese Studies*, **27**, 125–149.

Wagner, R. K., **Torgesen, J. K.**, **Rashotte, C. A.**, & **Pearson, N. A.** (1999). *Comprehensive test of phonological processing: CTOPP*. Pro-ed.

Wu, T. (2022). The general noun-modifying clause construction beyond Eurasia. *Linguistic Typology*, **26**(1), 89–127.

Yang, W., **Chan, A.**, **Chang, F.**, & **Kidd, E.** (2020). Four-year-old mandarin-speaking children's online comprehension of relative clauses. *Cognition*, **196**, 104103.

Yip, V., & **Matthews, S.** (2007). *The bilingual child: Early development and language contact*. Cambridge University Press.

Yoon, J.-H. (1993). Different semantics for different syntax: Relative clauses in Korean. In A. Kathol & C. Pollard (Eds.), *Working papers in linguistics no. 42: Papers in syntax* (pp. 199–226). Department of Linguistics.

Zacks, J. M. (2020). Event perception and memory. *Annual Review of Psychology*, **71**(1), 165–191.

Zhang, L., & **Zhou, J.** (2009). The development of mean length of utterance in mandarin-speaking children. In J. Zhou (Ed.), *The application and development of international corpus-based research methods* (pp. 40–58). Education Science Publishing House.

Zhou, J. (2001). *Pragmatic development of Mandarin speaking young children: from 14 months to 32 months* [Unpublished doctoral dissertation]. The University of Hong Kong.

Zhou, J., **Mai, Z.**, **Lau, E.**, **Lum, C.**, **Thian, A. L.**, & **Yip, V.** (2025). Cross- and within-language associations between phonological, lexical, and grammatical domains in mandarin-English bilingual preschoolers in Singapore. *Journal of Speech, Language and Hearing Research*, **68**(2), 636–653.

Cite this article: Shang, M., Mai, Z., Matthews, S., & Yip, V. (2026). Noun-Modifying Clause Construction in Mandarin-Speaking Children. *Journal of Child Language* 1–27, https://doi.org/10.1017/S0305000926100750

https://doi.org/10.1017/S0305000926100750 Published online by Cambridge University Press