

ARTICLE

The sensitivity to pragmatic-conceptual representations in garden path double relatives in L1 Mandarin

Bing Bai^{1,2,3} 

¹College of Foreign Languages and Cultures, Xiamen University, Xiamen, China; ²School of Foreign Languages, Soochow University, Suzhou, China and ³National Language Resources Monitoring and Research Center for Education and Teaching Media, Xiamen University, Xiamen, China
Email: bbxmu@xmu.edu.cn

(Received 19 October 2023; Revised 03 July 2024; Accepted 25 July 2024)

Abstract

The aim of this study is to investigate how pragmatic-conceptual representations can be integrated into theories of first language acquisition. Experiment 1, using a sentence–picture judgment task, examined how children (N = 53, aged 4–6 years) used prosody boundaries as cues for a recursive interpretation when the recursive relatives (i.e., SO and OO)¹ were garden path structures. The results showed that children below six-year had a stronger preference for recursive reading than adults under the conjunction-biased prosody condition and that children after six years of birth exhibited an adult-like preference for recursive readings under the recursion-biased prosody condition. Experiment 2 explored whether and how reversibility (e.g., “a dog eats a banana” vs “a dog kisses a cat”) in the action schema affected the production of OO and SO in Mandarin-speaking children (N = 137, age: 4–8 years). The results showed that adult-like production of OO in both reversible and irreversible conditions appeared at the age of six. The adult-like production ability of SO showed a one-year delay in the reversible condition (seven years under the reversible condition versus six years under the irreversible condition). The study suggests that some pragmatic-conceptual representations (such as the action schema) may be precursors of language and serve as a default analysis in language acquisition, while the mapping of the prosody domain onto syntax matures over time.

Keywords: child language; interface; Mandarin; recursion; relative clause

¹Here are abbreviations used in this paper: Rel-relative marker; V-verb; C-classifier; D-classifier; RC-relative clauses (in this paper, only object/subject-gapped relative clauses were discussed, short for object RCs and subject RCs); NP-nominal phrase; OO-a relativized object embedded inside another relativized object; SO-a relativized subject embedded inside a relativized object. Following Lin and Bever (2011), we labeled OO and SO double relatives in the current study, meaning that there are two relativized nominals in the structure.



1. Introduction

Although the acquisition of recursively embedded structures (e.g., the apple on the plate under the table) has been extensively studied in various languages, the doctrine of syntax autonomy in the generative enterprise does not incorporate pragmatic knowledge as a part of grammar. Research in psychology literature have shown that children have innate separable systems (such as animacy, number, perceptual properties, and the generic concept) that form the basis for higher-level human cognition (Kinzler & Spelke, 2007; O'Bryan, 2004; Spelke, 2003). This suggests a level of independence between the thought system and language (Sauerland & Alexiadou, 2020). It is still unclear how children identify isomorphic connections between linguistic and pragmatic-conceptual notions. Therefore, further discussion is necessary to determine whether pragmatic-conceptual representation, or at least a certain aspect of it, should be seen as a precursor to language and how it should be integrated into acquisition theories.

For decades, prosody – one of the most commonly used pragmatic cues in language comprehension – has been a long-standing concern revolving around three issues in language acquisition: 1) how the prosodic boundary serves as a cue in syntactic analysis, 2) to what extent syntactic domains and prosodic domains are independent and interact and 3) whether children differ from adults in the syntactic analysis of a given ambiguous structure when prosodic cues are explicitly biased toward an analysis. If a matching relation exists between prosodic domains and syntactic domains, a strong hypothesis emerges that the match can facilitate the parsing and comprehension of syntactic constituents (e.g., Selkirk, 2011). A plethora of studies have shown that prosody could assist in lexical learning, categorization, clausal typing and structural analysis during infancy (e.g., Esteve-Gibert & Prieto, 2018; Godde et al., 2020; Speer & Ito, 2009) and that prosodic cues might aid in the processing of center-embedded structures in both artificial languages and natural languages (Mueller et al., 2010; Roncaglia-Denissen et al., 2013). Singh and Fu (2016) argued that, unlike non-tone languages, tone language learners need to acquire both a lexical tone system and an intonational structure. In fact, very few acquisition studies to date have explored whether the prosodic boundary aligns with the syntactic boundary in recursive phrasing in tone languages and whether the perception of prosody in recursive phrasing varies across ages. If adults and children demonstrate different sensitivity to the mapping of prosody cues onto recursive phrasing, prosody and recursion may develop in parallel; otherwise, the mapping of prosody onto syntax needs to be learned. Thus, one of the aims in the current study is to investigate the prosodic effect on recursive phrasing in Mandarin-speaking children and adults.

Another pragmatic-conceptual representation closely related to language acquisition is the action schema in the mental structure. Chomsky (2014) argued that expressions must be linked with some elementary preexisting mental structure within the mind, such as the actor–action link, and that the actor–action schema is represented by the predicate–argument form in human languages. However, even if the merge of a transitive verb and its argument fits the action schema represented by the verb–argument link, some structures that violate pragmatic plausibility (e.g., the cake bites the dog) denote implausible situations in real-world occurrence. The psychology of the schema has been articulated by the acquisition delay of passives such as *John was hit by Bill* (where the action can be reversible between the actors) in contrast with *the milk was drunk by John* (where the action cannot be reversible

between the actors) (e.g., Ambridge et al., 2021; Bencini & Valian, 2008; Stromswold, 2006) and by the comprehension of Mandarin relatives (Macdonald et al., 2020; Wu et al., 2012). Given that the conceptual representation in the mental structure might operate at the syntax–thought interface, another aim of this study is to further investigate whether and how the reversibility in the action schema affects the acquisition of double relatives and whether the sensitivity to the reversibility varies across structures.

To examine the effect of prosody and reversibility in the action schema on the acquisition of recursion altogether, garden path double relatives are ideal data in two directions. For one thing, the syntax–prosody interface can earn credit due to the earlier emergence of recursive structures whose syntactic boundaries match the prosodic boundaries. For another, in ambiguous structures whose prosody is controlled for recursion interpretation, if reversibility in the action schema could affect the acquisition of double relatives, it is fair to argue for some default mental structures as precursors in human intelligence, in line with psycholinguistic literature (e.g., Kako, 2006; Mahajan & Woodward, 2009; Sauppe et al., 2023; Ünal et al., 2021). Thus, structures that meet the strict interface of prosody and action schema and syntax are assumed to emerge earlier than those do not.

2. The garden path double relatives in Mandarin

Before moving to the temporal ambiguity in double relatives in Mandarin, we first look at the basic properties of relativized subject (e.g., 1a) and relativized object (e.g., 1b) in Mandarin.

1. a. Relativized subject

V	NP ₁		de	NP ₂
<i>e_i</i>	<i>chi</i>	<i>pingguo</i>	<i>de</i>	<i>jiejie_i</i>
eat	apple	de	sister	

the sister who eats an apple

b. Relativized object

NP ₁	V		de	NP ₂
<i>mao</i>	<i>zhui</i>	<i>e_i</i>	<i>de</i>	<i>gou_i</i>
cat	chase	de	dog	

the dog that the cat chases.

In Mandarin, relative clauses are head-final, meaning that the modifier precedes the relativized nominal phrase. The marker for relativization can be various lexical items such as the particle *de*, bare demonstratives *zhe* (this) or *na* (that), bare classifiers, demonstrative–classifiers and the numeral *yi* (one) (Arcodia, 2017). Due to the similar word order, the parser cannot distinguish between a simple clause (subject–verb–object) and an object relative clause (subject–verb–Rel–object) until *de* appears. However, the demonstrative and classifier (D-C) can help disambiguate the sentence due to their requirement for nominal phrases. For instance, when a subject RC is preceded by D-C (e.g., 2a), the parser predicts a nominal phrase since D-C cannot select a verb but a noun phrase. This prediction effect is less consistent in object RCs since they start with a nominal phrase (e.g., 2b) that can be selected by D-C. Recent corpus-based studies have found that D-C immediately precedes the

verb in subject RCs but the head noun in object RCs, consistent with reaction time data (Sheng & Wu, 2013; Tang, 2007; Wu, 2011; Wu et al., 2009).

2. a. na-ge e_i chi pingguo de jiejie_i
 that-C eat apple de sister
that sister who eats an apple
- b. na-ge mao zhui e_i de gou_i
 that-C cat chase de dog
the dog that that cat chases.

In Mandarin, embedding a relativized subject and a relativized object gives rise to two types of garden path structures: that is., OO (e.g., 3a-b) and SO (e.g., 4a-b).

3. a. OO under the reversible action-actor condition
- | | | | | | | | | |
|-------|-------|-------|-------|------------------|-----|-------|----|------------------|
| N_1 | V_1 | | N_2 | V_2 | | N_3 | | |
| zhu | la | e_i | de | mao _i | da | e_j | de | hou _j |
| pig | drag | | de | cat | hit | | de | monkey |
- the monkey the cat the pig drags hits*
- b. OO under the irreversible action-actor condition
- | | | | | | | | | |
|-------|-------|-------|-------|------------------|------|-------|----|--------------------|
| N_1 | V_1 | | N_2 | V_2 | | N_3 | | |
| zhu | la | e_i | de | mao _i | ti | e_j | de | beizi _j |
| pig | drag | | de | cat | kick | | de | cup |
- the cup the cat the pig drags kicks*
4. a. SO under the reversible action-actor condition
- | | | | | | | | | |
|-------|-------|-------|----|------------------|-------|-------|-------|------------------|
| | V_1 | N_1 | | N_2 | V_2 | | N_3 | |
| e_i | la | mao | de | gou _i | da | e_j | de | hou _j |
| | drag | cat | de | dog | hit | | de | monkey |
- the monkey the dog that drags the cat hits.*
- b. SO under the irreversible action-actor condition
- | | | | | | | | | |
|-------|-------|------------|----|-----------------|-------|-------|-------|------------------|
| | V_1 | N_1 | | N_2 | V_2 | | N_3 | |
| e_i | chi | xigua | de | hu _i | la | e_j | de | hou _j |
| | eat | watermelon | de | tiger | drag | | de | monkey |
- the monkey the tiger that eats watermelon pulls*

The complexity of OO and SO can be broken down into three aspects. Firstly, it is important to differentiate between two types of analyses due to temporal ambiguity: recursive analysis, where a relativized nominal phrase (N_3 in 3a-b and 4a-b) has only one clausal modifier; and conjunctive analysis, where a relativized nominal phrase (N_3 in 3a-b and 4a-b) has two clausal modifiers. For instance, the relativized head noun *hou* (monkey) in 3(a) can be modified by either one clause (i.e., the monkey that the cat hits) or two clauses (i.e., the monkey that is hit by the cat and dragged by the pig). In the same way, the relativized head noun *beizi* (cup) in 3(b) can be modified by either one clause (i.e., the cup that the cat kicks) or by two clauses (i.e., the cup that is kicked by the cat and is dragged by the pig). Similarly, the relativized head noun *hou* (monkey) in 4(a) can be modified by either one clause (i.e., the monkey that the dog hits) or two clauses (i.e., the monkey that drags the cat and that is hit by the dog). The

relativized head noun *hou* (monkey) in 4(b) can also be modified by either one clause (i.e., the monkey that the tiger drags) or two clauses (i.e., the monkey that eats the watermelon and that is dragged by the tiger). Secondly, the (ir)reversibility in the actor–action link as represented by the agent–patient relationship introduces a varying degree of complexity in syntax–semantics. The reversible condition (i.e., 3a and 4a) could increase the likelihood of conjunctive analysis. Specifically, the agent–patient relationship between the cat and the cup in 3(b) is not reversible, but the relationship is reversible between the monkey and the cat in 3(a). Reversibility is also allowed between the cat and the monkey in 4(a) but is not allowed between the monkey and the watermelon in 4(b). Thirdly, OO and SO differ in terms of syntax embedding branching. OO is derived from left-branching embedding, in alignment with the canonical word order of Mandarin, whereas SO is not. As per Sheldon (1974), we further argue that the high cost of deriving SO lies in the perceptual change in thematic roles. Specifically, the language analyzer only needs to operate the existing object RC structural representation and semantic representation twice to generate OO. In contrast, deriving SO requires a perceptual change of thematic roles from agent–action–patient in object RCs to action–patient–agent in subject RCs when the mapping across modules first occurs in object RC (alternatively, a perceptual change from action–patient–agent in subject RCs to agent–action–patient in object RCs if the mapping across modules first occurs in subject RC). Therefore, these structural differences may lead us to observe different prosodic effects on the comprehension and production of OO and SO.

In comparison with the garden path model (Frazier & Fodor, 1978), the challenge in comprehension arises from whether the representation can be supported by contexts or schemata (the general framework used to organize details based on previous experience) (Ferreira et al., 2002). For instance, the correct interpretation of the passive sentence *the dog was bitten by the man* can be provided by the syntactic algorithm, but the awkward meaning should be supported by the communicative context or world knowledge. If the relationship between dog as patient and man as agent is not reinforced, the syntax–semantics–pragmatics interface will result in a higher cost in understanding *the dog was bitten by the man* compared with *the man was bitten by the dog*. Similarly, pragmatic and world knowledge play a crucial role in sentence comprehension as emphasized in the RI-Val model (Cook & O'Brien, 2014; O'Brien & Cook, 2016). According to this model, complete sentence comprehension requires the activation of contextual information or general world knowledge in the initial stage, as well as the judgment of how activated concepts are mapped onto syntax. These models suggest that semantic anomalies and pragmatic plausibility act as cues for comprehension, which is consistent with Walsh et al. (2018).

3. The modulating effect of prosody on syntactic analysis

Various languages use distinct prosodic cues for syntactic phrasing. In Mandarin, pause serves as a more prominent cue for intonational phrasing compared with pre-boundary lengthening in English (Shen, 1993; Yang, 1997). The significant role of pause in indicating clause boundaries has also been observed in Dutch, German and English (as discussed in Baek, 2022; Männel et al., 2013). Yang (1997) among the first identified varying prosodic parameters representing hierarchical boundaries in Mandarin Chinese during spontaneous speech. Specifically, duration primarily

signals word boundaries, while both duration and pause have limited differences in signaling phrasal boundaries. Pause stands out as the primary prosodic cue for identifying sentential and clausal boundaries. Subsequent studies have consistently demonstrated the effectiveness of pause in Mandarin for disambiguating garden path sentences (e.g., Wang et al., 2003, 2004). In Mandarin, pauses can be denoted by word spacing or commas in written text (e.g., Bai et al., 2008; Luo et al., 2013). For instance, the structure $V + NP_1 + de + NP_2$ in Mandarin can be interpreted as a verb phrase if a pause follows the verb or as a relative clause if a pause occurs after *de*, as indicated by word spacing (Yu, 2011; Yu & Yan, 2015).

Furthermore, the sensitivity to prosodic information develops and matures over the course of language development. Zhou et al. (2012) discovered that four- to five-year-olds were unable to resolve the ambiguity of the focus particle *zhīyou* (only) indicated by stress in sentences like *only Xiaoming's clock is yellow*. However, they found that children could resolve the ambiguity of *shenme* (what) in speech acts, where questions had a rising tone and statements had a level tone. The delayed acquisition of lexical prosody may be attributed to the late mastery of the complex tonal patterns at six years of age or later (Wong & Leung, 2018). Beach et al. (1996) among the earlier studies observed that five-year-olds could use prosody to understand the structure of coordinated sequences of adjectives. In contrast, Männel et al. (2013) found that like adults, German-speaking six-year-olds could perceive intonational phrase boundaries independently of a pause, while three-year-olds seemed to need all available prosodic cues for intonational phrase boundary perception. Wiedmann and Winkler (2015) also noted that English-speaking five-to-six-year-olds could differentiate between sentences like “Mary draws the boy’s hammer” and ‘Mary draws. The boys hammer’ based on prosodic boundaries. In contrast, Korean five-to-six-year-olds had different results in a similar task (Choi & Mazuka, 2003). These findings suggest that syntax may develop earlier than the pragmatic use of prosody and that children around six years old may establish a firm syntax–pragmatic–syntax mapping if a universal developmental path exists. On the other hand, the delayed sensitivity to prosodic boundaries could be specific to certain languages.

Although prosodic cues facilitate syntactic analysis, the facilitation effect might be structure dependent. Mueller et al. (2010) created three syllable sequences (AiAjBiBi, AiAjBjBj and AiAjBiBj) that violated dependency relations in A^nB^n . All of these sequences were examined under four conditions (unsegmented speech stream, descending sentential prosody, pause between quadruplets and pause between corresponding pairs). They found that prosody served as a cue for perceiving the irregularity and regularity of the center-embedded dependencies. In comparison, Roncaglia-Denissen et al. (2013) found that regular rhythm only facilitated the processing of embedded object RCs in German. Recently, researchers have begun exploring the correlation between prosody awareness and recursive embedding. Gomes et al. (2017) found that adult listeners used prosody as a cue for recursive embedding in the production and perception of relatives in Karaja, but Hirayama et al. (2021) did not find an association between recursive possessives and the prosodic contrast when Japanese children produced recursive nominals.

The current study examined how prosodic boundary cues the recursive phrasing in two types of garden path relative structures in Mandarin and to what extent children differ from adults in perceiving the connection between prosody and recursive embedding.

4. The reversibility triggered by animacy in action schema

It is reasonable to argue that language users have a bias toward a default mode in their conceptual representation of the predicate–argument link. In other words, certain pragmatic–conceptual representations as prototypes emerge before the full syntax is developed, and these components may act as ‘triggering experiences’ that align with a grammatical analysis similar to other biological triggers (e.g., Chomsky, 2014; Morgan et al., 2020). The shared notion of mental state and scenario may serve as triggers for linguistic representation and the growth of grammar. A strong piece of evidence supporting this is the preference for irreversible conditions, which is supported by decoding of event roles. The reversibility in the actor–action link is attributed to the animacy of the argument(s) of a predicate (e.g., the agent–patient relationship in *a dog bites a cake* is irreversible due to pragmatic plausibility). Ferreira and Stacey (2000) discovered that it took 25% longer to judge the plausibility of events described in passives like *the dog was bitten by the man* compared with events like *the man bit the dog*, *the man was bitten by the dog* and *the dog bit the man*. This effect is also observed in children with ASD (Ambridge et al., 2021), where the ASD group made significantly more errors in reversible passives (e.g., *Wendy was hit by Bob* when Wendy hit Bob) than the typically developing group, suggesting that children understand the situation and map the mental representation to syntax (Ünal & Papafragou, 2016, 2019).

The processing and comprehension of relative clauses also support the psychology of the action schema, whether it is corpus-based or laboratory-setting-based. Wu et al. (2012) introduced the Animacy Preference Constraints based on the interaction of relative clause type and animacy of the head noun. Specifically, the subject tends to be animate, head nouns in object-extracted relative clauses tend to be inanimate, and subject-extracted modifying subject relative clauses have animate relativized head nouns, which aligns with corpus findings (e.g., Pu, 2007; Wu et al., 2009). Recently, Macdonald et al. (2020) explored the influence of animacy on the incremental processing of subject relative clauses and object relative clauses in a picture–sentence matching task. Children aged 4–6 years old and adults listened to subject relative clauses or object relative clauses. The eye-tracking data revealed that the inanimate head noun did not facilitate the anticipation of object relative clauses in both children and adult groups. The eye movement data also indicated no interaction between animacy and relative clause types in adults. These results suggest that children are sensitive to semantic reversibility and rely on their primitive knowledge of agents to decode the semantic relation encoded in relative clauses, which is consistent with previous studies on relatives across languages (e.g., English, Traxler et al., 2002; Spanish, Betancort et al., 2009; Dutch, Mak et al., 2006; Mandarin, He & Chen, 2013).

5. Empirical studies on double relatives across languages

One consistent finding in acquisition literature is that young children prefer to use conjunction analysis in recursive relatives. This preference has been observed in various languages, including Wapichana, Romanian, English, Mandarin and others (e.g., Amaral & Leandro, 2018; Avram et al., 2020; Yang et al., 2022). However, it is unclear how pragmatics and semantics interacted with syntax in acquisition of recursion.

Yang et al. (2022) discovered that children's production of OO can be adult-like at the age of six when the action schema is reversible, while the adult-like production of SO emerges at the age of eight to nine. Yang et al. (2023) also observed that the adult-like production ability of OO under the reversible condition shows a one-year advantage compared with the irreversible condition (i.e., seven-year-olds versus six-year-olds) and that adult-like production of SO under the reversible condition emerges much later than that under the irreversible condition (i.e., nine-year-olds versus six-year-olds). They attributed this delay in acquisition to the syntax of recursion.

5. a. gege yang de yu tu de paopao
brother feeds de fish blow de bubbles
the bubbles the fish the brother feeds blows Yang et al. (2023)
- b. chi xiangjiao de jiejie na de qiqiu
eats banana de sister holds de balloon
the balloon the sister who eats a banana holds Yang et al. (2023)

In our opinion, Yang et al. (2022, 2023) could be further refined. One area for improvement in the experimental design is the pragmatic implausibility that may limit the conjunction analysis. For example, in their design, OO (e.g., 5a) can be interpreted as conjunction, including (a) [[gege yang de] [yu tu de] paopao] and (b) [[gege yang de] [yu tu de paopao]], both interpreted as *the bubbles the fish blows* and *the brother feeds*. However, the predicate–argument represented by *yang-paopao* (feed-bubbles) is pragmatically implausible (since the brother cannot feed the bubbles). Similarly, SO (e.g., 5b) can also be interpreted as coordination, including (a) [[chi xiangjiao de] [jiejie na de] qiqiu] and (b) [[chi xiangjiao de][jiejie na de qiqiu]], both interpreted as *the balloon that the sister holds and that eats bananas*. However, the predicate–argument represented by *xiangjiao-chi-paopao* (banana-eat-balloon) is anti-pragmatic (since the balloon cannot eat the banana). It is therefore unclear whether this implausibility enhances or hinders the acquisition of OO and SO. Another area for improvement is that Yang et al. (2022, 2023) only used one test sentence in OO and SO in each condition, hurting the statistical power.

The current research thus aimed to investigate whether and how pragmatic-conceptual representations (i.e., prosody and reversibility in the action schema) influence the acquisition of recursively embedded relativized nominals.

6. Methods

6.1. Experiment 1: a comprehension task

6.1.1. Predictions

Building on the work of Kaland and Van Heuven (2010) and Yu et al. (2022), this experiment investigated whether prosodic boundaries could act as a cue for recursive reading in garden path double-relative structures in Mandarin. As discussed in Section 3, it typically takes five to six years for individuals to develop sensitivity to prosodic cues in syntactic analysis. This suggests that the secure alignment of syntax with pragmatic representation develops independently and this asymmetry in development may also be evident in syntactic recursion. Of particular interest is the variation in complexity between object–object (OO) and subject–object

(SO) structures. Firstly, the word order of OO structures bears similarity to the canonical word order in Mandarin. Additionally, the perception shift in the syntax–semantics mapping is easier in OO structures compared with SO. OO structures may therefore represent a strict interface between different domains. Thus, we made the following predictions:

- Prediction 1: It takes years for children to develop a secure mapping of prosodic boundaries and recursive phrasing in an adult-like manner.
- Prediction 2: The mapping of prosody onto syntax in OO becomes mature earlier than that in SO.

6.1.2. Participants

Fifty-three monolingual Chinese-speaking children from four to six were recruited, and 25 monolingual Chinese-speaking college students served as controls. All participants did not have hearing or visual impairment, and all children were typically developing. Children were divided into five age-groups. Children were divided into three groups: 4 years (female = 11, male = 7; $M = 4;06$), 5 years (female = 10; male = 10; $M = 5;06$) and 6 years (female = 6; male = 9; $M = 6;05$). They were not told the objective of the experiment. We obtained approval from the caretakers before conducting the experiment. In order to retain the focus of four-year-old children, a caretaker accompanied the participant but did not give any hints for the judgment. Each participant was paid 30 Chinese *yuan* after finishing the experiment.

6.1.3. Materials

The experimental design consisted of three prosodic conditions and two types of structures: OO and SO. Each sentence included two conjunction-biased prosodic conditions (i.e., R-R condition with two relative markers *de* immediately before a pause; R-N condition with a pause immediately after the first relative marker *de* and after the second relativized head noun) and one recursion-biased prosodic condition (i.e., the N-N condition with a pause immediately after the first and a pause immediately after the second relativized head noun). Two scenarios with corresponding items were presented in two pictures featuring the same objects and animals: one for conjoined reading and the other for recursive reading. To minimize the animacy effect, the animacy of the nominal phrases was controlled, resulting in two action schema conditions in OO and SO. Specifically, in OO under the irreversible condition, NP_1 and NP_2 were animate (animals), while NP_3 was inanimate (objects); in OO under the reversible condition, all NPs were animate (animals); in SO under the irreversible condition, NP_1 was inanimate (object), while NP_2 and NP_3 were animate (animals); in SO under the reversible condition, all NPs were animate (animals). Twelve sentences were created as critical items, with 24 fillers that were ambiguous in the form of ‘V + NP_1 + *de* + NP_2 ’, selected from Yu et al. (2019, 2022) and then modified for child-friendly literacy. All items were arranged in a Latin square design.

The distribution of D-C in relatives was controlled according to findings in the corpus and controlled tasks (see the review in Section 2). Specifically, D-C preceded the whole structure of subject RCs (i.e., D-C – [subject RC]), and they preceded the head in object RCs (i.e., D-C+ head noun). To minimize the effect of tone, all nominal

phrases had the same tone contour (214 + 214), and all the verbs had the same tone contour (55). To create a child-friendly audio, we did not use category-specific classifiers in critical items (e.g., *zhi* for ‘dog’) but used the category-general classifier *ge* that appears very early in child Mandarin. Then, a male native speaker was recruited to record the sentences. Following studies on the prosodic boundary in phrasing (e.g., Shen, 1993; Yang, 1997; Yu, 2011; Yu & Yan, 2015), we manipulated the pause 200 ms with Praat. Table 1 below gives an example of structures under the irreversible condition cued in three prosodic conditions (the prosodic boundary was indicated by ‘/’ in the following examples).

Table 1. Examples of OO and SO under the irreversible condition in three prosodic conditions

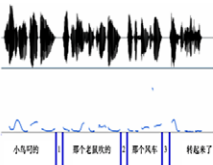

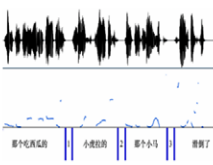

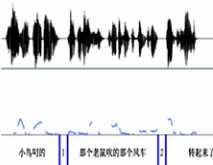

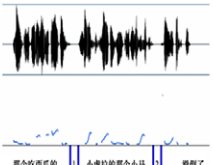

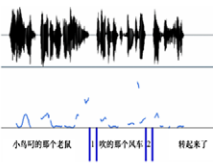

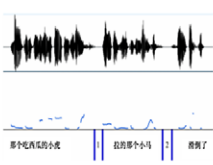

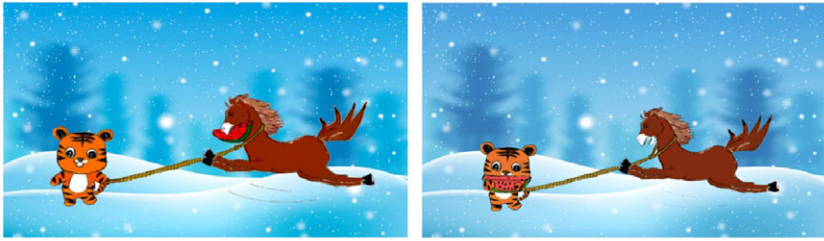
	OO	SO
R	<p>$[s_{21}xiao]naio \rightarrow diao \rightarrow de/na-ge[s_{21}laoshu] \rightarrow chui \rightarrow de/na-ge[s_{21}fengehe] \rightarrow zhuangqi$ bird \rightarrow bit \rightarrow de \rightarrow D-C \rightarrow mouse \rightarrow blew \rightarrow de \rightarrow D-C \rightarrow windmill \rightarrow spun-SF The windmill that the bird bit and that the mouse blew spun.^{2,3}</p>	<p>$na-ge \rightarrow chi \rightarrow [s_{21}xigua] \rightarrow de/[s_{21}xiao]hu \rightarrow la \rightarrow de/na-ge \rightarrow [s_{21}xiao]ma \rightarrow huadao$ D-C \rightarrow ate \rightarrow watermelon \rightarrow de \rightarrow tiger \rightarrow dragged \rightarrow de \rightarrow D-C \rightarrow horse \rightarrow slipped- \rightarrow The horse that ate the watermelon and that the tiger pulled slipped.^{2,3}</p>
R	 	 
R	<p>$s_{21}xiao]naio \rightarrow diao \rightarrow de/na-ge[s_{21}laoshu] \rightarrow chui \rightarrow de \rightarrow na-ge \rightarrow [s_{21}fengehe] / zhuangqi$ bird \rightarrow bit \rightarrow de \rightarrow D-C \rightarrow mouse \rightarrow blew \rightarrow de \rightarrow D-C \rightarrow windmill \rightarrow spun-S \rightarrow The windmill that the bird bit and that the mouse blew spun.^{2,3}</p>	<p>$na-ge \rightarrow chi \rightarrow [s_{21}xigua] \rightarrow de/[s_{21}xiao]hu \rightarrow la \rightarrow de \rightarrow na-ge \rightarrow [s_{21}xiao]ma / huadao$ D-C \rightarrow ate \rightarrow watermelon \rightarrow de \rightarrow tiger \rightarrow dragged \rightarrow de \rightarrow D-C \rightarrow horse \rightarrow slippe \rightarrow The horse that ate the watermelon and that the tiger pulled slipped.^{2,3}</p>
R	 	 
N	<p>$s_{21}xiao]naio \rightarrow diao \rightarrow de \rightarrow na-ge[s_{21}laoshu] / chui \rightarrow de \rightarrow na-ge \rightarrow [s_{21}fengehe] / zhuangqi$ bird \rightarrow bit \rightarrow de \rightarrow D-C \rightarrow mouse \rightarrow blew \rightarrow de \rightarrow D-C \rightarrow windmill \rightarrow spun-S \rightarrow The windmill that the mouse that the bird bit blew spun.^{2,3}</p>	<p>$na-ge \rightarrow chi \rightarrow [s_{21}xigua] \rightarrow de \rightarrow [s_{21}xiao]hu \rightarrow la \rightarrow de \rightarrow na-ge \rightarrow [s_{21}xiao]ma / huadao$ D-C \rightarrow ate \rightarrow watermelon \rightarrow de \rightarrow tiger \rightarrow dragged \rightarrow de \rightarrow D-C \rightarrow horse \rightarrow slipped \rightarrow \rightarrow The horse that the tiger that ate the watermelon pulled slipped.^{2,3}</p>
N	 	 

Table 2. The two contrastive visual representations of a sentence stimulus under the R-R condition**Audio recording:**[Ⓔ]

Zheliyou → liangge → xiaoma → huadaole.[Ⓔ]

Here → have → two → C → horse → slipped.[Ⓔ]

Qing → quanchulai → na-ge → chi → NP₂xigua → → de/[NP₂xiaohu] → la → → de/ → na-ge/[NP₃xiaoma][Ⓔ]

Please → you → circle → D-C → → ate → watermelon → de → tiger → → dragged → de → D-C → horse.[Ⓔ]

Please circle *the horse that ate the watermelon and that the tiger pulled.*[Ⓔ]

6.1.4. Procedures

The experiment included two tests: a pretest in the adult group to confirm the validity and reliability of using prosodic pause in conjunctive reading and recursive reading, and a formal test in the children group. All stimuli were presented on slides on an iPad. Both tests were conducted individually in Chinese following the same procedure in both groups. Participants had one trial for training to become familiar with the task. Before beginning the test in both groups, the examiner showed a video flash introducing the animals and objects in the stimuli sentences. Participants were then instructed to click the space bar when they were ready to start. When the space bar was clicked, two pictures were displayed for 10 seconds, followed by an audio recording played once. Participants were asked to circle the target referent from the two pictures simultaneously displayed on the screen (shown in Table 2). Throughout the formal test, the entire procedure and participants' judgments were recorded for data analysis.

6.1.5. Coding

Participants' responses were divided into recursive readings and non-recursive readings.

6.1.6. Results and analysis

First, a Friedman test was conducted based on the mean accuracy scores of recursive interpretations in OO and SO (Table 3) to investigate whether prosodic cues influenced recursion interpretation in adult group. Regarding the conjunction analysis in both OO and SO, the results indicated that both R-R condition and R-N condition triggered conjunction analysis and that these two conditions showed no significant difference. Regarding the recursive analysis in OO and SO, the results

Table 3. The mean of recursive reading each type of structures in three conditions in adults

		R-R condition		N-N condition		R-N condition	
		Mean	SD	Mean	SD	Mean	SD
OO	The irreversible condition	1.20	1.23	2.56	0.65	1.60	1.08
	The reversible condition	1.56	0.77	2.60	0.58	1.44	0.96
SO	The irreversible condition	1.40	0.96	2.92	0.28	1.60	1.00
	The reversible condition	1.36	1.04	2.88	0.33	1.52	0.96

Table 4. The mean of recursive interpretation in OO across ages in three prosody conditions across ages

	4 years (<i>N</i> = 18)		5 years (<i>N</i> = 20)		6 years (<i>N</i> = 18)		adults (<i>N</i> = 25)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
R-R condition	3.78	1.40	3.80	1.36	4.07	1.49	2.76	1.56
N-N condition	4.17	1.10	4.35	0.99	4.60	1.18	5.16	0.85
R-N condition	3.06	1.39	3.65	1.53	4.53	1.25	3.04	1.59

showed that the R-R condition yielded significantly lower scores than the N-N condition ($p < 0.001$), and the R-N condition also produced significantly lower scores than the N-N condition ($p < 0.001$), demonstrating that the N-N condition significantly prompted recursive analysis.

Secondly, the ANOVA test, with satisfied variance homogeneity, compared the mean accuracy scores of recursive interpretations of OO (shown in Table 4) among different age-groups. The results revealed significant differences in recursive reading among different age-groups in the R-R condition ($p = 0.022$), the R-N condition ($p = 0.012$) and the N-N condition ($p = 0.010$). Specifically, LSD-based multiple comparisons showed that in the R-R condition, the mean value of recursive reading in children group was significantly higher than that of adults (at four, $p = 0.027$; at five, $p = 0.020$; at six, $p = 0.008$). This indicated that children of all ages had a stronger preference for recursive reading than adults, when the R-R condition suggested a conjunctive reading. Furthermore, in the R-N condition, there was no significant difference in recursive reading between children at four and adults ($p = 0.973$) or between children at five and adults ($p = 0.171$). However, the preference for recursive reading of children at six was significantly higher than that of adults ($p = 0.003$), showing that children at six exhibited a stronger preference for recursive syntax than adults, when the R-N condition suggested a conjunctive reading. In the N-N condition, there was no significant difference in recursive reading between children at six and adults ($p = 0.095$), but the mean value of recursive reading in the adult group was significantly higher than that of children at four ($p = 0.002$) and at five ($p = 0.009$). These results suggested that children older than six years were able to successfully map prosodic boundaries onto recursive embedding.

Thirdly, with the satisfactory variance homogeneity, an ANOVA test was conducted to compare the mean accuracy score of recursive interpretation of SO (shown in Table 5) among different age-groups. The results indicated significant differences in recursive reading abilities among different age-groups in the R-R condition ($p = 0.006$) and the N-N condition ($p < 0.001$). There was no significant difference in recursive reading among different age-groups in the R-N condition ($p = 0.233$). Specifically, LSD-based multiple comparisons revealed that in the R-R condition,

Table 5. The mean of recursive interpretation in SO across ages in three prosody conditions across ages

	4 years (<i>N</i> = 18)		5 years (<i>N</i> = 20)		6 years (<i>N</i> = 18)		adults (<i>N</i> = 25)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
R-R condition	3.72	1.18	4.20	1.32	4.00	1.65	2.76	1.56
N-N condition	4.11	1.49	4.10	1.21	5.13	1.13	5.80	0.41
R-N condition	3.72	1.32	3.80	1.64	4.07	1.39	3.12	1.62

children across different age-groups had significantly higher preference for recursive reading compared with adults (at four, $p = 0.034$; at five, $p = 0.001$; at six, $p = 0.010$). This suggests that even when the R-R condition triggered a conjunction reading, children aged six still showed a stronger preference for recursive interpretation compared with adults. In the R-N condition, the preference for recursive reading in adults was not significantly different from that in children groups (at four, $p = 0.203$; at five, $p = 0.139$; at six, $p = 0.060$). This indicates that the recursive reading of children aged four to six were not significantly higher or lower than those of adults, even when the R-N condition cued a conjunction analysis. Additionally, in the N-N condition, the preference for recursive reading in adults was significantly higher than that in children at four ($p < 0.001$) and five ($p < 0.001$). However, there was no significant difference in preference for recursive reading between adults and children aged six ($p = 0.064$). These results suggested that children over the age of six exhibited adult-like sensitivity to prosodic boundaries and recursive embedding.

In summary, experiment 1 found that prosodic representation could trigger recursive syntax, but children did not display adult-like sensitivity to the mapping of prosodic boundary and recursive embedding in OO and SO until six years of age in the recursive-biased prosody condition (i.e., the N-N condition).

6.2. Experiment 2: a production task

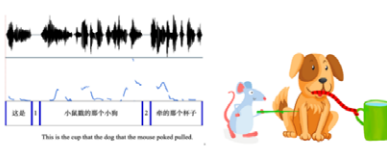
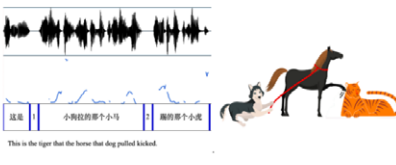
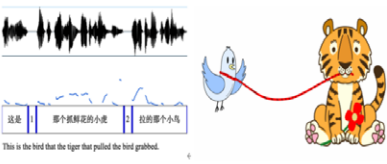
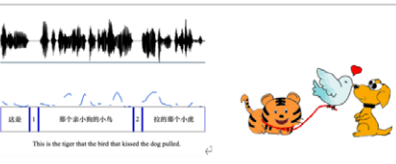
6.2.1. Predictions

Building on the findings of Yang et al. (2022), experiment 2 further explores the impact of reversibility in action schema on the production of OO and SO. While the reversible action-actor link may lead to increased complexity in OO and SO compared with the irreversible link (as discussed in Section 4), OO is comparatively simpler for children due to its word order, left-branching embedding and lower cost of perception change in thematic roles. Therefore, when explicit robust prosody focused on recursion (i.e., the N-N condition in experiment 1) was controlled, OO is projected to have a possible strict interface across various domains. As a result, we hypothesized that the adult-like production ability of OO may emerge earlier than that of SO.

6.2.2. Participants

One hundred and fifty-five children were recruited, but 18 children (including nine children aged four, five children aged five and four children aged six) were ruled out because they were not willing to continue the experiment in the mid-way. We obtained approval from the caretakers before conducting the experiment. Children below five years were accompanied by a caretaker who helped the examiner retain the participants' interest in the experiment. Children were aged from four to eight: four-

Table 6. Examples of the audio stimuli and corresponding picture display of OO and SO

The irreversible action schema condition	The reversible action schema condition
OO	
<p>zhe-shi-xiaoshu-chuo → de-na-ge → xiaogou-qian → de-na-ge → beizi¹² this-be→mouse¹ poked → de→D-C → dog → pulled → de→D-C → cup¹² This is the cup that the dog that the mouse poked pulled.¹²</p>	<p>zhe-shi-xiaogou-la → de-na-ge → xiaoma-ti → de-na-ge → xiaohu¹² this-be→dog → pulled → de→D-C → horse → kicked → de→D-C → tiger¹² This is the tiger that the horse that dog pulled kicked.¹²</p>
	
SO	
<p>zhe-shi-na-ge → zhu → xianhua-de-xiaohu-la → de-na-ge → xiaoniao¹² this-be→D-C → grabbed → flower → tiger → pulled → de→D-C → bird¹² This is the bird that the tiger that grabbed the flower pulled.¹²</p>	<p>zhe-shi-na-ge → qin → xiaogou-de-xiaoniao → la → de-na-ge → xiaohu¹² this-be → D-C → kissed → dog → de → bird → pulled → de → D-C → tiger¹² This is the tiger that the bird that the dog pulled.¹²</p>
	

year-olds (female = 10; male = 15; M = 4;06), five-year-olds (female = 17; male = 13; M = 5;06), six-year-olds (female = 14; male = 13, M = 5;06), seven-year-olds (female = 14; male = 15; M = 5;06) and eight-year-olds (female = 10; male = 16, M = 5;06). We also had 20 adult natives as controls, and all of them were undergraduates. All participants had no difficulty identifying animals and plants appearing in the pictures used in the current design. None of the participants were recruited in Experiment 1. Each participant was paid 30 Chinese *yuan*.

6.2.3. Materials

In Experiment 1, 12 critical items were identified as those containing explicit prosody cues for recursive reading (in other words, the N-N condition). Twelve fillers were also selected for Experiment 1. A male native speaker was recruited to record all the sentences. The critical sentence and its corresponding stimuli were minimal pairs, as only one nominal phrase differed between the two structures. One picture was created for the stimuli and another for the target item. In line with the approaches of Yang et al. (2022) underpinned in Ambridge and Rowland (2013), a flash video was developed by Python. This video began with a frame introducing the objects involved in the flash and then presented the critical items. The entire experiment was carried out in Mandarin Chinese. Table 6 below provides examples of the audio stimuli and their corresponding picture displays.

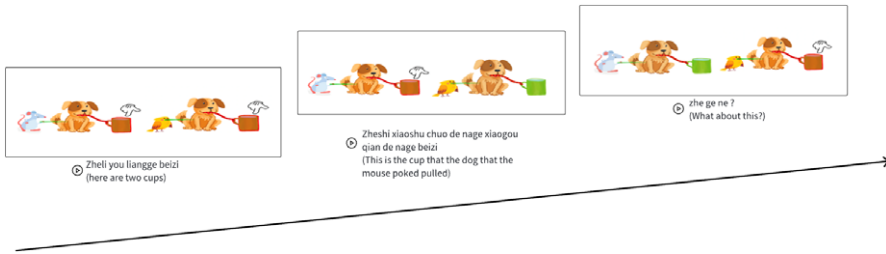


Figure 1. The time course of the presentation of one critical item (Note: Only pictures were displayed on the screen).

6.2.4. Procedures

Participants took part in our online experiment through Zoom meetings, Tencent meetings and ExplainEverything. Prior to the start of the experiment, two episodes were played multiple times to ensure that participants correctly identified objects. The examiner then clicked a button to display each that is one at a time. If a participant requested a replay, the examiner would comply. If a participant’s choices differed between two playbacks, a third playback would occur. The results from two playbacks were recorded for analysis. The entire session for each participant was recorded, and participants’ responses were transcribed for further analysis.

Here, we used one critical that is to illustrate the entire procedure of the experiment (Figure 1). Two pictures were presented on the screen, with two cups flickering for two seconds while simultaneously pointing a finger, immediately after an audio *zheli you liangge beizi* (here are two cups) was played. When an audio *zhesi xiaoshu chuo de nage xiaogou qian de nage beizi* (This is the cup that the dog that the mouse poked pulled) was heard, only the cup in the stimuli sentence flickered. When an audio *zhe ge ne* (what about this cup) played, the flickering of the cup in the stimuli sentence stopped, and the finger-pointing disappeared. Instead, the cup in the target sentence picture began flickering with a finger-pointing. The presentation for one trial shown below (Figure 1) demonstrates a complete process of one trial.

6.2.5. Coding

In accordance with Roeper and Oseki (2018) approach to recursion development, we examined the participants’ responses across different age-groups. If the focus of production was the targeted double relative structure, we assigned a code of 1; otherwise, a code of 0 was assigned for nontarget structures. In terms of nontarget production, we conducted an analysis of conjoined structures in various grammatical forms that were consistent with the occurrence in drawings. The categories of nontarget responses that we examined are listed below.

1. The situation-match conjoined analysis (the links between the nominal entities were correct, but the production was not recursive relatives).

①. Two conjoined simple sentences
 e.g., xiaoma la xiaohu, xiaohu qian xiaoiao
 little-horse pulls little-tiger, little-tiger leads little-bird
The little horse pulls the little tiger. The little tiger leads the little bird.

②. a simple sentence in juxtaposition with a free RC, or two RCs in juxtaposition.

e.g., xiaoma la xiaohu, xiaohu qian de na-ge xiaoniao
 little-horse pulls little-tiger, little-tiger lead de D-C little-bird
The little horse pulls the little tiger. The little bird the little tiger leads.

e.g., xiaoma la de na-ge xiaohu, xiaohu qian
 little-horse pulls de D-C little-tiger, little-tiger lead
 de na-ge xiaoniao
 de D-C little-bird
The little tiger the little horse pulls. The little bird the little tiger leads.

2. The conjoined-RC analysis (two grammatical RCs had the same head noun and matched the situation).

e.g., xiaoma la de na-ge xiaohu, xiaoniao qian de na-ge xiaohu
 little-horse pulls de D-C little-tiger, little-bird lead de D-C little-tiger
the little tiger that the little horse pulls and that the little bird leads.
 (Target response: *the little bird the little tiger the little horse pulls leads*)

3. The topic-prominent analysis (two clauses had the same referent and matched the situation).

e.g., zhe-shi xiaoniao qin-zhe na-ge laoshu la-zhe xiaohu
 this-be little-bird kissed-ASP D-C mouse pull-ASP little-tiger
This is the bird kissing the mouse and pulling the tiger.
 (Target response: *this is the little tiger that the little bird that is kissing the mouse pulling.*)

6.2.6. Results and analysis

First, a repeated-measures ANOVA was conducted to compare the mean accuracy score of the production of OO and SO (shown in Table 7) among different age-groups, treating reversible and irreversible action schema conditions as one group. Since Mauchly's test of sphericity was not met ($p < 0.001$). The Greenhouse-Geisser results revealed that the within-subjects effect was significant ($F = 37.910, p < 0.001$), indicating that the production of OO and SO were significantly different. The results also showed that age significantly affected the production ability of OO and SO ($F = 132.193, p < 0.001$). Specifically, the LSD-adjusted multiple comparisons indicated that children until six years old exhibited adult-like production ability of

Table 7. The mean of the recursive production in two types of structures

		4 years (N = 25)	5 years (N = 30)	6 years (N = 27)	7 years (N = 29)	8 years (N = 26)	adults (N = 20)
OO	Mean	1.76	4.47	5.44	5.52	5.69	5.85
	SD	1.09	1.28	0.64	0.69	0.62	0.37
SO	Mean	1.12	3.13	4.89	5.28	5.50	5.60
	SD	0.83	1.57	0.75	0.80	0.58	0.50

Table 8. The mean of the two types of structures across ages in two action schema conditions

Conditions			4years (N = 25)	5years (N = 30)	6years (N = 27)	7years (N = 29)	8years (N = 26)	Adults (N = 30)
OO	irreversible	Mean	0.96	2.33	2.78	2.83	2.85	2.95
		SD	0.46	0.61	0.42	0.38	0.37	0.22
	reversible	Mean	0.80	2.13	2.67	2.69	2.85	2.90
		SD	0.87	0.82	0.48	0.47	0.37	0.31
SO	irreversible	Mean	0.76	1.77	2.59	2.76	2.81	2.85
		SD	0.52	0.82	0.50	0.44	0.40	0.37
	reversible	Mean	0.36	1.40	2.30	2.59	2.69	2.75
		SD	0.49	0.86	0.47	0.50	0.47	0.44

OO ($SE = 0.254$, $p = 0.113$), while children until seven years old displayed adult-like production ability of SO ($SE = 0.274$, $p = 0.238$).

Secondly, we conducted a repeated-measures ANOVA test to compare the mean value of recursive production of OO (shown in Table 8) among different age-groups separately. Due to Mauchly's test of sphericity not being satisfied ($p = 0.002$), we used the adjusted results in GreenHouse–Geisser. The GreenHouse–Geisser adjusted result indicated that the production of OO and SO significantly differed between two action schemata conditions ($F = 21.656$, $p < 0.001$). The results also showed that age significantly affected the production of OO and SO ($F = 136.762$, $p < 0.001$).

The LSD-adjusted multiple comparison was used to compare the mean value of recursive production of OO in the reversible and irreversible conditions, indicating that the production of OO in the irreversible condition was not significantly different between adults and children after six years of birth ($SE = 0.129$, $p = 0.183$) and that the production of OO in the reversible condition was not significantly different adults and children after six years of birth ($SE = 0.178$, $p = 0.191$), suggesting that the reversibility of action schema did not affect the production of OO. Meanwhile, the LSD-adjusted multiple comparison compared the mean value of recursive production of SO in the reversible and irreversible conditions. The LSD-adjusted multiple comparison showed that the production of SO in the irreversible condition was not significantly different between adults and children after six years of birth ($SE = 0.160$, $p = 0.109$), while the production of SO in the reversible condition was not significantly different between adults and children after seven years of birth ($SE = 0.165$, $p = 0.324$), suggesting that the irreversibility of the action schema facilitated the production of SO in an adult-like manner.

Additionally, we described the nontarget production of children across different age-groups. At the age of four, we identified three types of conjoined analysis, with the situation-match analysis consistently being the most common in both OO and SO. The topic-prominent analysis was rarely used in OO, but it was the predominant strategy in SO regardless of the condition of action schema. OO tended to trigger more conjoined analysis compared with SO. By the age of five, children rarely utilized conjoined analysis in OO. From the ages of five to seven, the topic-prominent analysis emerged as the dominant nontarget strategy in SO.

In summary, experiment 2 demonstrated that, when the prosodic condition was controlled for recursive reading, achieving adult-like production in OO was not influenced by the irreversibility of action schema. However, adult-like production in SO appeared one year earlier in the irreversible condition compared with that in the reversible condition.

7. Discussion

The current study investigated the production and comprehension of OO and SO, two types of garden path double relatives. Experiment 1 suggested that children may need up to six years to develop a solid mapping of prosody boundaries onto recursive embedding. Experiment 2 showed that the ability to produce OO structures similar to adults emerged earlier than that of SO structures. Additionally, the current results showed that the facilitation effect of irreversible action schema is more pronounced in SO structures compared with OO structures. These findings suggest that language and cognitive development largely occur independently. Certain pragmatic and conceptual components, such as action schema, may act as precursors in language development, while others, like prosody mapping onto syntax, may take more time to fully develop.

While OO and SO structures differ in their composition, it may take children around six years to acquire an adult-like sensitivity to correctly mapping prosodic boundaries onto syntactic phrases. Similarly, many studies showed that children aged around five to six develop a secure mapping of prosody onto syntax (English, Beach et al., 1996; Wiedmann & Winkler, 2015; Mandarin, Wong & Leung, 2018; German, Männel et al., 2013; Korean, Choi & Mazuka, 2003). In response to the claim that Mandarin is assumed to be more difficult to acquire than a non-tone language because of lexical tone and intonational structure (Singh & Fu, 2016), the current results suggested that although Mandarin has a rich inventory tone system, Mandarin-speaking children are comparable to their counterparts in other tone languages regarding the emergence of mapping of prosody and syntactic phrasing in the adult-like manner. The current comprehension data showed that children after six years of birth showed the adult-like sensitivity to the mapping recursion-oriented prosody boundary (i.e., the N-N condition) onto recursive phrasing and that children below six showed stronger preference of recursive reading even if the prosody condition (i.e., the R-R condition; the R-N condition) cued conjunction analysis. Thus, it is suggested that children at six could have a secure mapping of prosodic domain and syntactic domain. Accounts arise in two directions: The recursive mechanism as the core property of human language might come earlier than the pragmatic use of prosody; alternatively, it might be true that both prosody and recursive function grow together, but their mapping grows later and then prosody could serve as a trigger for recursion, that is, the syntactic domains map onto the prosodic domains only when the pragmatic use of prosodic cues was acquired 2011.

Moreover, the effect of reversible action schema on the production ability of double relatives is more robust in SO than OO. The first evidence is that the facilitation effect of irreversible action schema was only found in the production of SO but not in OO. Another evidence is the productivity of OO used to replace SO to deliver the event structures correctly, as well as the situation-match analysis in OO and SO at four years. These suggest that children are born with a prototypical representation that fits visual experience and syntactic structures (Ambridge et al., 2021; Bencini & Valian, 2008). Specifically, the innate action schema might serve as trigger information to identify and establish the agent–patient relations relevant in the linguistic construction; that is, the contrast between the inanimate manipulable and the animate agent helps children reach the most economical representation in the preverbal infancy (Mahajan & Woodward, 2009; Ünal et al., 2021). Recently, Lau and Tanaka (2021) claimed that structures would be easier to process, when the animacy configuration matches the processor's expectations based on prototypical thematic/grammatical roles of arguments in a transitive clause (e.g., the rock that the hiker

rolled away), and that structures would be difficult to process when the animacy configuration leads to initial misanalysis (e.g., the hiker that the rock crushed).

Following Kennedy (2008) and Roeper (2014), we would argue that OO satisfies the strict interface in pragmatics, and semantics and syntax. In comparison with Yang et al. (2022)'s claim that the difficulty in acquiring double relatives should be ascribed to the immature autonomy of syntax, we would argue that the one-year delay emergence of adult-like production ability of SO in the reversible condition might be due to the unsatisfied interface across domains; that is, a successful acquisition should entail a successful recognizing and handling the interface relations. This could be elucidated in three directions. First, the interface of event structure and syntax expressed in object RCs is far more transparent than that in subject RCs. Specifically, the configuration of events in the form of 'entity–action–entity' in object RCs directly maps onto the syntax configuration 'subject–verb–object' of object RCs, and the syntactic roles of each nominal phrase directly onto the thematic roles (agent–predicate–patient). By comparison, the action schema of a subject RC cannot directly map onto its syntax–semantics configuration, due to the divergence between the thematic roles in event representation (agent–verb–patient) and syntactic derivation (verb–object–subject). In this case, a perception change in the thematic roles of agent to patient is needed to map the thematic role onto the representation of action schema and syntax, in line with Sheldon (1974). Following this vein, the fast and immediate mapping across syntax–semantics–pragmatics in Mandarin object RCs also extend into OO, because the derivation of OO only needs one more running of syntax–semantics–pragmatics interface. In other words, the bottom-up perceptual experience interplays with the event schema that store information about previously executed or observed actions (Elsner & Adam, 2021), and if the interplay fits the interface across modules, acquisition might be facilitated. Second, the iconicity of word order of object RCs and the canonical word order of Mandarin contribute to the fast mapping across modules. Abundant evidence shows that structures in a language's canonical word order should be easier to process and acquire, because a canonical word order is assumed to be used by the language processor (Diessel, 2009; Sekerina, 2003; Slobin & Bever, 1982; Tavakolian, 1981; see more in the review of Lau & Tanaka, 2021). Recently, Zhang and Chao (2019) argued that if a structure entailed the canonical word order of the language, it should be acquired earlier. For example, subject RCs should be acquired earlier than object RCs in English, because the word order of the subject RCs entails the SVO word order of English; object RCs should be acquired earlier than subject RCs in Mandarin, because the word order of object RCs 'S–V–*de*–O' in Mandarin (*de* is the relativization marker) is aligned with the canonical word order of Mandarin SVO. Third, the strict interface manifested in OO can also be explained by the branching direction in parsing. According to the continuum of the branching direction bias in different recursive nominals, recursive relative clause had a right-branching bias (Berg, 2012). Due to the parametrically determined headedness in Mandarin, OO is only expanded in one direction while SO is not. The knowledge of the psychological accommodation (e.g., expansion of a structure in the same direction is easier when the listener decodes the left-to-right speech stream) allows faster processing in OO than in SO.

Finally, the nontarget production showed that the structured meaningful representation was available to a child although the syntactic forms was restricted, lending support to the Meaning First Approach to language acquisition (Sauerland & Alexiadou, 2020) and the prior existence of core system in human intelligence (e.g., Kinzler & Spelke, 2007; O'Bryan, 2004; Spelke, 2003). In other words, to have adult-like

grammar, a child needs to deal with the linearization in articulation and the externalization of the logic form and the lexical items, all of which are ‘guided by various information-theoretic and communicative considerations, the context of use, the linguistic items, rules available in a given language, and considerations of learnability’ (Guasti et al., 2023). Thus, if the pragmatic-conceptual information was narrowed down to action schema, we would argue that the representation of action schema searches for all construction alternatives that are determined by the current developmental stage of grammar and match the target event, as suggested in Chomsky (2014).

8. Conclusion

In contrast to the illusion that pragmatic representation is not part of grammar, the experimental studies revealed that some pragmatic-conceptual representations (e.g., action schema) serve as triggers for language acquisition, while some pragmatic-conceptual representations (e.g., prosody) did not develop with syntax in parallel. Following Kennedy’s (2008) and Roeper’s (2014) elaboration on interface, the current author suggested that children seek a strict interface across syntax, semantics and pragmatics when they encounter complex grammar beyond reach, and that the default mode of mental representation triggers a quick syntactic analysis of a given structure. Explorations into language acquisition and language processing in the future lies in three directions: the neural basis of the triggering effect of pragmatic-conceptual components on recursion; the psychology of strict interface in other language domains and in different populations (e.g., children with language impairment, L2 learners and multilinguals); a comprehension model resolving around the type and strength of the pragmatic-conceptual cues that prompts a selectional attention to the syntactic analysis.

Data availability statement. The full database is publicly available, and we would also like to share the sum of target production in each type of structures under each condition. Please visit the link below https://osf.io/6fnyk/?view_only=c4612c45c62d4c93a70c3c7117838c38.

Acknowledgements. This work is a revised version of one chapter of the author’s dissertation. The author feels thankful to Prof. Yang Cai mei (the author’s supervisor in Soochow University), as well as participants in this work.

The author extends gratitude to Prof. Tom Roeper, who gave constructive comment on the first draft, as well as many colleagues (e.g., Prof. Jill de Villiers, Prof. Usha Lakshmanan, and Dr. Bleotu, Deborah Foucault Etheridge, to name just a few) in Recursion Meetings the University of Massachusetts Amherst, where the author presented this work and feel thankful to the audience.

Funding statement. This research is supported by the Central University Basic Research Fund of China (Grant No. ZK1125) and Fujian Provincial Federation of Social Sciences (Grant No. FJ2024C050).

References

- Amaral, L., & Leandro, W. (2018). Relative clauses in Wapichana and the interpretation of multiple-embedded “uraz” constructions. In L. Amaral, M. Maia, A. Nevins, & T. Roeper (Eds.), *Recursion across domains* (pp. 230–242). Cambridge University Press.
- Ambridge, B., Bidgood, A., & Thomas, K. (2021). Disentangling syntactic, semantic and pragmatic impairments in ASD: Elicited production of passives. *Journal of Child Language*, 48(1), 184–201.
- Ambridge, B., & Rowland, C. F. (2013). Experimental methods in studying child language acquisition. *Cognitive Science*, 4, 149–168.
- Arcodia, G. F. (2017). Towards a typology of relative clauses in Sinitic: Headedness and relativization strategies. *Cahiers de Linguistique Asie Orientale*, 46(1), 32–72.

- Avram, L., Sevenco, A., & Tomescu, V. (2020). The acquisition of recursively embedded noun modifiers in Romanian by Hungarian–Romanian bilinguals. *Bucharest Working Papers in Linguistics*, 22, 61–84.
- Baek, H. (2022). Prosodic disambiguation in first and second language production: English and Korean. *Language and Speech*, 65(3), 598–624.
- Bai, X., Yan, G., Liversedge, S. P., Zang, C., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: Evidence from eye movements. *Journal of Experimental Psychology. Human Perception and Performance* 34(5), 1277–1287.
- Beach, C. M., Katz, W. F., & Skowronski, A. (1996). Children's processing of prosodic cues for phrasal interpretation. *Journal of Acoustical Society of America*, 99, 1148–1160.
- Bencini, G. M. L., & Valian, V. V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language*, 59(1), 97–113.
- Berg, T. (2012). Branching direction in recursive structures. *English Language and Linguistics*, 16, 385–401.
- Betancort, M., Carreiras, M., & Sturt, P. (2009). Short article: The processing of subject and object relative clauses in Spanish: An eye-tracking study. *Quarterly Journal of Experimental Psychology*, 62, 1915–1929.
- Choi, Y., & Mazuka, R. (2003) Young children's use of prosody in sentence parsing. *Journal of Psycholinguistic Research*, 32, 197–217.
- Chomsky, N. (2014). Minimal recursion: exploring the prospects. In T. Roeper, & M. Speas (Eds.), *Recursion: Complexity in cognition. Studies in theoretical psycholinguistics* (pp. 1–15). Springer.
- Cook, A. E., & O'Brien, E. J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes*, 51(1–2), 26–49.
- Diessel, H. (2009). On the role of frequency and similarity in the acquisition of subject and non- subject relative clauses. In T. Givón, & M. Shibatani (Eds.), *Syntactic complexity: Diachrony, acquisition, neuro-cognition, evolution* (pp. 251–276). John Benjamins.
- Elsner, B. & Adam, M. (2021). Infants' goal prediction for simple action events: The role of experience and agency cues. *Topics in Cognitive Science*, 13, 45–62.
- Esteve-Gibert, N., & Prieto, P. (2018). Early development of the prosody-meaning interface. In N. Esteve-Gibert, P. Prieto (Eds.), *The development of prosody in first language acquisition* (pp. 227–246). John Benjamins.
- Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11(1), 11–15.
- Ferreira, F., & Stacey J. (2000). The misinterpretation of passive sentences.
- Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6(4), 291–325.
- Godde, E., Bosse, M. L., & Bailly, G. (2020). A review of reading prosody acquisition and development. *Reading and Writing*, 33, 399–426.
- Gomes, J. N., França, F. A., Maia, M., & Rilliard, A. (2017). Prosody as a recursive embedding tool in production and perception of Karaja: An acoustic and neuro-psycholinguistic investigation. *Journal of Speech Sciences*, 5, 101–123.
- Guasti, M. T., Alexiadou, A., & Sauerland, U. (2023). Under compression errors as evidence for conceptual primitives. *Frontiers in Psychology*, 14, 1104930.
- He, W., & Chen, B. (2013). The role of animacy in Chinese relative clause processing. *Acta Psychologica*, 144(1), 145–153.
- Hirayama, M., Colantoni, L., & Pérez-Leroux, A. T. (2021). Can prosody encode recursive embedding? Children's realizations of complex NPs in Japanese. *Journal of Child Language*, 48(3), 429–453.
- Kako, E. (2006). Thematic role properties of subjects and objects. *Cognition*, 101(1), 1–42.
- Kaland, C., & van Heuven, V. J. (2010). The structure-prosody interface of restrictive and appositive relative clauses in Dutch and German. *Speech Prosody*. Paper 064. [https:// doi.org/10.21437/SpeechProsody.2010-186](https://doi.org/10.21437/SpeechProsody.2010-186).
- Kennedy, C. (2008). Argument contained ellipsis. In K. Johnson (Ed.), *Topics in ellipsis* (pp. 95–131). Cambridge University Press.
- Kinzler, K. D., & Spelke, E. S. (2007). Core systems in human cognition. *Progress in Brain Research*, 164, 257–264.
- Lau, E., & Tanaka, N. (2021). The subject advantage in relative clauses: A review. *Glossa: A Journal of General Linguistics*, 6(1), 34.
- Lin, C.-J., & Bever, T. (2011). Garden path and the comprehension of head-final relative clauses. In H. Yamashita, et al. (Eds.), *Processing and producing head-final structures, studies in theoretical psycholinguistics* (pp. 277–297). Springer.

- Luo, Y. Y., Yan, M., & Zhou, X. L. (2013). Prosodic boundaries delay the processing of upcoming lexical information during silent sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 915–930.
- Macdonald, R., Brandt, S., Theakston, A., Lieven, E., & Serratrice, L. (2020). The role of animacy in children's interpretation of relative clauses in English: Evidence from sentence-picture matching and eye movements. *Cognitive Science*, 44(8), e12874.
- Mahajan, N., & Woodward, A. L. (2009). Seven-month-old infants selectively reproduce the goals of animate but not inanimate agents. *Infancy: The Official Journal of the International Society on Infant Studies*, 14(6), 667–679.
- Mak, W., Vonk, W., & Schriefers, H. (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*, 54(4), 466–490.
- Männel, C., Schipke, C. S., & Friederici, A. D. (2013). The role of pause as a prosodic boundary marker: language ERP studies in German 3- and 6-year-olds. *Developmental Cognitive Neuroscience*, 5, 86–94.
- Morgan, E. U., van der Meer, A., Vulchanova, M., Blasi, D. E., & Baggio, G. (2020). Meaning before grammar: A review of ERP experiments on the neurodevelopmental origins of semantic processing. *Psychonomic Bulletin & Review*, 27(3), 441–464.
- Mueller, J. L., Bahlmann, J., & Friederici, A. D. (2010). Learnability of embedded syntactic structures depends on prosodic cues. *Cognitive Science*, 34(2), 338–349.
- O'Brien, E. J., & Cook, A. E. (2016). Coherence threshold and the continuity of processing: The RI-Val model of comprehension. *Discourse Processes*, 53(5–6), 326–338.
- O'Bryan, E. L. (2004). Event structure in language comprehension. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 64, 3666.
- Pu, M.-M. (2007). The distribution of relative clauses in Chinese discourse. *Discourse Processes*, 43(1), 25–53.
- Roeper, T. (2014). Strict interface principles and the acquisition engine: From unlabeled to labeled and minimal modular contact. *Language Sciences*, 46, 115–132.
- Roeper, T. & Oseki, Y. (2018). Recursion in the acquisition path for hierarchical syntactic structure. In L. Amaral, M. Maia, A. Nevins, & T. Roeper (Eds.), *Recursion across domains* (pp. 267–278). Cambridge University Press.
- Roncaglia-Denissen, M. P., Schmidt-Kassow, M., & Kotz, S. A. (2013). Speech rhythm facilitates syntactic ambiguity resolution: ERP evidence. *PLoS One*, 8(2), e56000.
- Sauerland, U., & Alexiadou, A. (2020). Generative grammar: A meaning first approach. *Frontiers in Psychology*, 11, 571295.
- Sauppe, S., Næss, 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, e13340.
- Sekerina, I. A. (2003). Scrambling and processing: dependencies, complexity, and constraints. In S. Karimi (Ed.), *Word order and scrambling* (pp. 301–324). Blackwell.
- Selkirk, E. (2011). The syntax-phonology interface. In J. Goldsmith, J. Riggle, & A. C. L. Yu (Eds.), *The handbook of phonological theory* (2nd ed., pp. 435–484). Blackwell Publishing.
- Sheldon, A. (1974). The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior*, 13, 272–281.
- Shen, X. (1993). The use of prosody in disambiguation in Mandarin. *Phonetica*, 50(4), 261–271 (in Chinese).
- Sheng, Y., & Wu, F. (2013). Demonstrative-classifier positioning in Chinese relative clauses and its underlying reasons: A spoken corpus study. *Modern Foreign Languages*, 2, 150–157 (in Chinese).
- Singh, L., & Fu, C. S. (2016). A new view of language development: the acquisition of lexical tone. *Child Development*, 87(3), 834–854.
- 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.
- Speer, S. R., & Ito, K. (2009). Prosody in first language acquisition –acquiring intonation as a tool to organize information in conversation. *Language and Linguistics Compass*, 3, 90–110.
- Spelke, E. S. (2003). What makes us smart? Core knowledge and natural language. In D. Gentner, & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 277–311). Boston Review.
- Stromswold, K. (2006). *Why children understand and misunderstand sentences: An eye tracking study of passive sentences*. Rutgers University Center for Cognitive Science Technical Report TR-85.
- Tang, Z. (2007). Correlations between demonstrative constituents and the positions of relative clauses in spontaneous conversation: A typological perspective. *Contemporary Linguistics*, 9(2), 139–150 (in Chinese).

- Tavakolian, S. (1981). The conjoined clause analysis of relative clauses. In S. Tavakolian (Ed.), *Language acquisition and linguistic theory* (pp. 167–187). MIT Press.
- Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, 47(1), 69–90.
- Ünal, E., & Papafragou, A. (2016). Interactions between language and mental representations. *Language Learning*, 66, 554–580.
- Ünal, E., & Papafragou, A. (2019). How children identify events from visual experience. *Language, Learning and Development*, 15, 138–156.
- Ünal, E., Richards, C., Trueswell, J. C., & Papafragou, A. (2021). Representing agents, patients, goals and instruments in causative events: A cross-linguistic investigation of early language and cognition. *Developmental Science*, 24(6), e13116.
- Walsh, B., Usler, E., Bostian, A., Mohan, R., Gerwin, K. L., Brown, B., Weber, C., & Smith, A. (2018). What are predictors for persistence in childhood stuttering?. *Seminars in Speech and Language*, 39(4), 299–312.
- Wang, B., Yang Y., & Lv S. (2004). Acoustic analysis on prosodic hierarchical boundaries of Chinese. *Acta Acustica*, 29(1), 29–36 (in Chinese).
- Wang, D., Zheng, B., & Yang, Y. (2003). An experimental research on the role of prosodic features in disambiguation. *Journal of Psychological Science*, 4, 667–671 (in Chinese).
- Wiedmann, N. & Winkler, S. (2015). The influence of prosody on children’s processing of ambiguous sentences. In S. Winkler (Ed.), *Ambiguity: Language and communication* (pp. 185–198). De Gruyter.
- Wong, P., & Leung, C. T.-T. (2018). Suprasegmental features are not acquired early: perception and production of monosyllabic Cantonese lexical tones in 4- to 6-year-old preschool children. *Journal of Speech, Language, and Hearing Research*, 61, 1070–1085.
- Wu, F. (2011). Frequency issues of classifier configurations for processing Mandarin object extracted relative clauses: A corpus study. *Corpus Linguistics and Linguistic Theory*, 7(2), 203–227.
- Wu, F., Kaiser, E., & Andersen, E. (2009). The effects of classifiers in predicting Chinese relative clauses. In M. Grosvald & D. Soares (Eds.), *Proceedings of the 38th western conference on linguistics* (pp. 330–339). Department of Linguistics, University of California.
- Wu, F., Kaiser, E., & Andersen, E. (2012). Animacy effects in Chinese relative clause processing. *Language and Cognitive Processes*, 27(10), 1489–1524.
- Yang, C., Dong, X., & Zhao, J. (2023). A study of Mandarin Children’s acquisition of recursive relative clauses. *Journal of Foreign Languages*, 46(1), 49–62 (in Chinese).
- Yang, C., Hu, Y., Fan, J., Dong, X., & Jeschull, L. (2022). An experimental study of Mandarin-speaking children’s acquisition of recursion under a formal definition/classification system for recursion. *Lingua*, 271, 103228.
- Yang, Y. (1997). Prosodic cues to syntactic boundaries. *Acta Acustica*, 22(5), 414–421.
- Yu, M. (2011). An experimental research on the role of prosodic features in disambiguating “V+N1+de+N2” ambiguous phrases. *Journal of Nanjing University of Science and Technology (Social Science Edition)*, 24(5), 82–87 (in Chinese).
- Yu, M., Sommers, B., Yin, Y., & Yan, G. (2019). Effects of implicit prosody and semantic bias on the resolution of ambiguous Chinese phrases. *Frontiers in Psychology*, 10, 1308.
- Yu, M., & Yan, G. (2015). Effects of implicit prosody and contexts on ambiguity structure processing: evidence from eye movements. *Journal of Tianjin University (Science and Technology)*, 48(8), 686–691 (in Chinese).
- Yu, M., Zhou, S., & Long, J. (2022). The effect of implicit prosody and context on the disambiguation of Chinese balanced ambiguous structure. *Studies of Psychology and Behavior*, 20(6), 739–746 (in Chinese).
- Zhang, Y. & Chao, D. (2019). ertong jufa zenyang shengzhang: jiyu yuyan gongxing de xide gouxiang. *Journal of Capital Normal University (Social Sciences Edition)*, 6, 127–137 (in Chinese).
- Zhou, P., Crain, S., & Zhan, L. (2012). Sometimes children are as good as adults: The pragmatic use of prosody in children’s on-line sentence processing. *Journal of Memory and Language*, 67(1), 149–164.

Cite this article: Bai, B. (2025). The sensitivity to pragmatic-conceptual representations in garden path double relatives in L1 Mandarin. *Language and Cognition*, 17, e5, 1–23. <https://doi.org/10.1017/langcog.2024.40>