Applied Psycholinguistics (2021), **42**, 181–205
doi:10.1017/S0142716420000697

CAMBRIDGE
UNIVERSITY PRESS

ORIGINAL ARTICLE

Real-time comprehension of garden-path constructions by preschoolers: A Mandarin perspective

Peng Zhou[1,*], Jiawei Shi[1,*] and Likan Zhan[2]

[1]Tsinghua University, Beijing, China and [2]Beijing Language and Culture University, Beijing, China
*PZ and JS contributed equally to this work. Corresponding authors. E-mails: zhoupeng1892@mail.tsinghua.edu.cn; shijw20@mails.tsinghua.edu.cn

(Received 04 July 2019; revised 12 September 2020; accepted 11 October 2020; first published online 11 December 2020)

Abstract

The present study investigated whether 4- and 5-year-old Mandarin-speaking children are able to process garden-path constructions in real time when the working memory burden associated with revision and reanalysis is kept to minimum. In total, 25 4-year-olds, 5 25-year-olds, and 30 adults were tested using the visual-world paradigm of eye tracking. The obtained eye gaze patterns reflect that the 4- and 5-year-olds, like the adults, committed to an initial misinterpretation and later successfully revised their initial interpretation. The findings show that preschool children are able to revise and reanalyze their initial commitment and then arrive at the correct interpretation using the later-encountered linguistic information when processing the garden-path constructions in the current study. The findings also suggest that although the 4-year-olds successfully processed the garden-path constructions in real time, they were not as effective as the 5-year-olds and the adults in revising and reanalyzing their initial mistaken interpretation when later encountering the critical linguistic cue. Taken together, our findings call for a fine-grained model of child sentence processing.

Keywords: eye movements; garden-path constructions; preschool children; real-time processing; reanalysis

In the past few decades, numerous investigations on sentence processing suggest that adults exhibit incrementality in parsing (e.g., Ferreira, 2003; Ferreira & Clifton, 1986; Ferreira & Lowder, 2016; Frazier, 1979, 1987, 1989; Trueswell, Tanenhaus, & Garnsey, 1994). They do not postpone their parsing decisions in order to gain substantial information for accurate analysis, but rather they are actively engaged in incorporating the processed linguistic information to form a single dynamic representation, on the basis of which they then make predictions about the upcoming information.

(1) The horse raced past the barn fell.

However, such parsing strategy may lead to misinterpretations when temporary ambiguity is encountered in a sentence, and thus to correctly understand such a sentence requires later revision or reanalysis. These types of sentences are often referred to as garden-path constructions. One well-known example is given in (1). Before the parser encounters the last word, *fell*, there are two possible interpretations for the phrase *raced past the barn*: it can be either the predicate of the sentence or the postnominal modifier of the subject *the horse*. Upon the disambiguation point, that is, the verb *fell*, the predicate analysis collapses, thereby leaving the modifier analysis as the only plausible option, and the temporary ambiguity is then resolved. However, it has been reported that readers tend to adopt the predicate analysis as their initial interpretation even though both interpretations are available to them, and they revise the initial interpretation as the modifier analysis when later encountering the disambiguation word (Frazier, 1979; Frazier & Rayner, 1982).

To explain the parser's initial preference and subsequent revision, several accounts have been proposed. For instance, the garden-path theory (e.g., Ferreira & Clifton, 1986; Frazier, 1979, 1987, 1989; Frazier & Rayner, 1982) hypothesizes that the parser analyzes sentences according to their syntactic structure, and syntactic analysis can proceed without reference to other nonsyntactic sources of information. The garden-path theory is fundamentally syntax driven. However, it should be noted that the theory does not claim that other nonsyntactic sources of information are not important in processing. Rather, it argues for the separation of syntactic information from other information at some stage during sentence processing (e.g., Frazier & Rayner, 1982; Traxler, 2002, 2005; van Gompel & Pickering, 2007). In contrast, other accounts, such as the constraint-based theory (e.g., Boland, Tanenhaus, & Garnsey, 1990; MacDonald, 1994; Taraban & McClelland, 1988; Trueswell et al., 1994), the referential theory (Crain & Steedman, 1985), and the good-enough approach (e.g., Ferreira, 2003; Ferreira & Lowder, 2016), propose that the parser's initial commitment incorporates several sources of information, such as structural, verb subcategorization, and referential/contextual information. These various types of information interact to determine the analysis of a sentence. Empirical evidence in favor of the constraint-based theory and the referential theory was reported by Tanenhaus, Spivey-Knowlton, Eberhard, and Sedivy (1995), in which English-speaking adults were found to misinterpret the first prepositional phrase (PP) *on the towel* in (2) as the destination of the verb *put* significantly less often when contextual information supporting the correct modifier analysis was provided than when no relevant contextual information was provided.

(2) Put the apple on the towel in the box.

Overall, prior research on adult sentence processing has shown that when interpreting a sentence, the parser incrementally computes the structural representation and possible meanings of the sentence while drawing on different sources of linguistic and nonlinguistic information (e.g., Altmann & Kamide, 1999; 2007; Kamide, Altmann, & Haywood, 2003; Omaki, 2010; Pickering, Traxler, & Crocker, 2000; Staub & Clifton, 2006; Tanenhaus et al., 1995; van Berkum, Brown, Zwitserhood, Kooijman, & Hagoort, 2005; Zhan, 2018).

Recent studies on child sentence processing seem to suggest that when listening to a sentence, children also incrementally compute the structural representation and possible meanings of the sentence (Andreu, Sanz-Torrent, & Trueswell, 2013; Choi & Trueswell, 2010; Fernald, Zangl, Portillo, & Marchman, 2008; Lew-Williams & Fernald, 2007; Nation, Marshall, & Altmann, 2003; Omaki, 2010; Özge, Küntay, & Snedeker, 2019; Özge, Marinis, & Zeyrek, 2015; Sekerina & Trueswell, 2012; Trueswell, Sekerina, Hill, & Logrip, 1999; van Heugten & Shi, 2009; Zhou, Crain, & Zhan, 2014; Zhou, Ma, Zhan, & Ma, 2018). However, it has also been shown that although children process sentences incrementally, they fail to incorporate the referential information provided by the contexts into their initial interpretation (e.g., Choi & Trueswell, 2010; Kidd & Bavin, 2005, 2007; Kidd, Stewart, & Serratrice, 2011; Lassotta, Omaki, & Franck, 2016; Omaki, Davidson White, Goro, Lidz, & Phillips, 2014; Snedeker & Trueswell, 2004; Trueswell et al., 1999; Weighall, 2008; but cf., Meroni & Crain, 2003). In addition, compared with adults, children are more likely to fail to revise their initial interpretation when later encountering the disambiguating linguistic information, dubbed as the kindergarten-path effect by Trueswell et al. (1999). For example, Trueswell et al. (1999) adopted the test materials as in (3), similar to the ones used by Tanenhaus et al. (1995), to investigate whether 5-year-old English-speaking children could use the referential information provided in the acted-out scene to correctly interpret garden-path constructions. Both online eye movement data and offline action data were collected. The results showed that, in both the 1-referent scene (e.g., there was only one frog) and the 2-referent scene (e.g., there were two frogs, one on the napkin and one on a towel), children tended to misinterpret the first PP *on the napkin* as the destination of the verb *put* and failed to revise it when later hearing the correct destination *in the box*, as shown by both the frequent eye movements at the incorrect destination and the offline actions involving incorrect destination. The findings have been interpreted as evidence attesting to children's inability to use the referential information and to reanalyze their initial interpretation. In another study, Kidd et al. (2011) investigated whether 5-year-old English-speaking children could successfully revise their initial interpretation and reanalyze a sentence using the semantic information of the noun phrase (NP) at the end of the sentence. For instance, in sentence (4) the verb *cut* exhibits a strong bias to select the NP after the preposition *with* as its instrument, leading the readers to initially use the *with*-phrase to modify the verb. The semantic implausibility of the NP *the candle* as the instrument should then serve as the trigger to reanalyze the *with*-phrase as the modifier of the NP *the cake* rather than the modifier of the verb *cut*. Kidd et al. found that when presented with sentences as in (4), 5-year-old English-speaking children failed to recover from their initial misinterpretation caused by the verb selectional bias by using the disambiguating semantic plausibility information indicated by the NP.

(3) Put the frog on the napkin in the box.
(4) Cut the cake with the candle.

In addition to studies on English-speaking children, Choi and Trueswell (2010) explored whether 4- and 5-year-old Korean-speaking children could use the thematic role assignment information by the verb at the end of the sentence, as in (5), to recover from the initial misinterpretation. Korean is a subject–object–verb language, and the case marker –*ey* can be used either as a locative marker indicating the destination of the verb or as a genitive marker indicating that the marked noun is the modifier of the following NP. However, verbs like *cipu* "pick up" cannot assign a destination role to the initial PP, so when encountering such verbs, the parser abandons the destination analysis, leaving the modifier analysis as the only plausible interpretation. The findings were that although both children and adults initially preferred the locative interpretation of the word *naypkhin* "napkin," adults could take advantage of the thematic role assignment information by the verb at the end of the sentence to reanalyze *naypkhin* "napkin" as the modifier, whereas children could not use the information to revise their initial interpretation.

(5) Naypkhin-ey kaykwuli-lul cipu-sey-yo.
napkin-Loc/Gen frog-Acc pick up-Hon-SE
"Pick up the frog on the napkin."

One potential cognitive factor that has been proposed to account for the kindergarten-path effect exhibited by young children is their immature inhibitory control ability (e.g., Choi & Trueswell, 2010; Kidd et al., 2011; Mazuka, Jincho, & Onishi, 2009; Novick, Trueswell, & Thompson-Schrill, 2005; Omaki et al., 2014; Trueswell et al., 1999; Weighall, 2008; Woodard, Pozzan, & Trueswell, 2016; but cf., Huang & Hollister, 2019). For instance, Novick et al. (2005) argued that reanalysis is closely associated with the ability of inhibitory control. Due to young children's immature inhibitory control, when they incrementally process a sentence and establish provisional representations of the sentence, they have difficulties in inhibiting their initial provisional interpretation using the later encountered linguistic information. Woodard et al. (2016) directly investigated the relation between 4- and 5-year-old English-speaking children's ability of inhibitory control and their ability to reanalyze garden-path constructions as in (3), and found that children's reanalysis ability was positively correlated with their inhibitory control ability, thus providing some evidence attesting to the relation between the two.

Another cognitive factor that might contribute to children's difficulty with reanalysis is their limited working memory capacity (e.g., Choi & Trueswell, 2010; Kidd et al., 2011; Trueswell et al., 1999; Weighall, 2008; but cf. Woodard et al., 2016). There are theoretical models discussing how working memory capacity is related to reanalysis ability (e.g., Just & Carpenter, 1992; Lewis & Vasishth, 2005; Lewis, Vasishth, & van Dyke, 2006). For instance, Just and Carpenter (1992) proposed that two components in working memory are involved in reanalysis, storage, and processing, and the two components share the same resources pool. When encountering the ambiguous word, the parser initial stores its multiple interpretations in working memory, which costs additional working memory resources. As parsing continues, the less preferred alternative interpretation might have to be abandoned, if the remaining working memory resources are not sufficient for the processing of ensuing elements. If the alternative interpretation, which needs to be retrieved for reanalysis, is abandoned before the disambiguation point, then it would cause processing difficulties for reanalysis. Lewis and colleagues proposed a theory called the activation-based model (Lewis & Vasishth, 2005; Lewis et al., 2006). On this parallel processing model, when the parser initially encounters the ambiguous word, it activates one of its interpretations and the alternative interpretation starts to decay. Reanalysis is costly for working memory resources, because when encountering the disambiguating linguistic information, the parser needs to deactivate the initial preferred misinterpretation and simultaneously reactivate the correct alternative interpretation that has started to decay after the ambiguous word.

Overall, both theories assume that individuals with lower working memory capacity exhibit more difficulties with reanalysis. In addition, both theories posit that the linear distance between the ambiguous word and the disambiguation point positively correlates with the level of difficulties in reanalyzing garden-path constructions. When the ambiguous word is adjacent to the disambiguation point, that is, when the linear distance between the two is minimized, so are the difficulties associated with reanalysis. According to the theory by Just and Carpenter (1992), as the linear distance between the ambiguous word and the disambiguation point becomes greater, the processing load is also increasing. Because storage and processing share the same resources pool, the increase of processing load will automatically take up more resources available for the storage of less preferred alternative interpretation. Therefore, it is more likely for the parser to abandon the less preferred interpretation before the disambiguation point in working memory, and thus reanalysis becomes more difficult. By contrast, if the ambiguous word is adjacent to the disambiguation point, reanalysis occurs shortly after the parser stores the two interpretations in working memory, and therefore the processing load due to the storage of the less preferred interpretation is reduced to minimum. In this case, the less preferred interpretation might still be stored in working memory when reanalysis occurs, thereby alleviating the difficulties in reanalysis that requires the retrieval of the less preferred interpretation in working memory.

Similarly, in the activation-based model by Lewis et al. (2006), longer linear distance between the ambiguous word and the disambiguation point leads to longer decaying time of the correct alternative interpretation. Therefore, more memory resources are required to reactivate the decayed interpretation when later encountering the disambiguation point, resulting in reanalysis difficulties. In contrast, if the ambiguous word is adjacent to the disambiguation point, the decaying time of the correct alternative interpretation becomes much shorter, and thus to reactivate the decayed interpretation requires much fewer memory resources, thereby reducing the difficulties in reanalysis. The correlation between linear distance and reanalysis of garden-path constructions has been investigated and confirmed by experimental studies on adult sentence processing (e.g., Tabor & Hutchins, 2004; van Dyke & Lewis, 2003).

Although the two theories have not been directly tested using data from children's processing of garden-path constructions, the predictions should be fairly straightforward: young children are more likely to exhibit difficulties in reanalyzing garden-path constructions than adults, because they have more limited working memory capacity as compared to adults (e.g., Case, Kurland, & Goldberg, 1982; Gathercole, Pickering, Ambridge, & Wearing, 2004). In addition, the adjacency between the ambiguous word and the disambiguation point should presumably reduce the difficulties in reanalysis, because the working memory burden due to the linear distance between the two is reduced to minimum. On the basis of the two working memory models, the present study aims to investigate whether young children are able to revise their initial interpretation of garden-path constructions in which the ambiguous word is adjacent to the disambiguation point.

Prior research on children's reanalysis of garden-path constructions mainly focused on the nature of the lexical elements that are causing the ambiguity (Choi & Trueswell, 2010; Kidd et al., 2011; Trueswell et al., 1999), that is, the initial interpretation is either due to the verb argument structures, including the thematic role assignment by the verb (e.g., the verb *put* typically requires an object NP as its theme and a PP as its location) and the bias of the verb in the selection of its arguments (e.g., the verb *cut* exhibits a strong bias in selecting a *with*-phrase as its instrument), or due to the bias of the case marker (e.g., the case marker *–ey* in Korean is favored as a locative marker over a genitive marker). Much less is known about whether children could revise their initial interpretation when the ambiguous word is adjacent to the disambiguation point. In addition, most prior research focused on the kindergarten-path effect of English-speaking children, with only a few studies investigating this effect in other languages (see Choi & Trueswell, 2010, for Korean-speaking children, and Özge et al., 2015, for Turkish-speaking children). The cross-linguistic perspective has proven helpful and illuminating in revealing whether the kindergarten-path effect observed in English-speaking children holds for children across languages.

The present study offers a cross-linguistic perspective by looking at how preschool Mandarin-speaking children process garden-path constructions in real time. Mandarin is ideally suited for exploring this issue, which we discuss below.

Garden-path constructions in Mandarin are mostly associated with the grammatical morpheme *DE* (Lee, 2006).[1] Consider the Mandarin example in (6). It has the structure: "NP1 + Modal + Verb + NP2 + DE + NP3." The morpheme *DE* is a possessive marker, so NP2 + DE + NP3 indicates a possessive relation where NP2 is the animate possessor (e.g., *xiaogou* "dog") and NP3 is the inanimate possessee (e.g., *piqiu* "ball"). The verb *ti* "kick" is the verb that causes the initial misinterpretation, as it could take either an animate or inanimate entity as its plausible complement, so in (6) NP2 *xiaogou* "dog" could be a perfect complement for the verb. If the parser incrementally computes the structural representation and possible meanings of the sentence, it might initially analyze the structure "NP1 + Modal + Verb + NP2" as a complete sentence, as in (7), after hearing the verb *ti* "kick" and before encountering the marker *DE*. In other words, when processing (6), the parser might initially analyze NP2 *xiaogou* "dog" as the object NP of the verb *ti* "kick," rather than the modifier of the actual object NP *xiaogou DE piqiu* "dog's ball." The possessive marker *DE*, which is adjacent to the ambiguous NP2 *xiaogou* "dog," is the disambiguation point (trigger for reanalysis). Upon encountering the marker *DE*, the parser will need to revise its initial analysis of NP2 (*xiaogou* "dog") and reanalyze it as the modifier of the object NP (*xiaogou DE piqiu* "dog's ball").

(6) Xiaomao yaoqu ti xiaogou DE piqiu
cat will kick dog DE ball
"The cat is going to kick the dog's ball."

(7) Xiaomao yaoqu ti xiaogou
cat will kick dog
"The cat is going to kick the dog."

Note that English also has possessive markers like *'s*, as in *John's apple*. However, compared with its English counterpart, the Mandarin possessive marker *DE* is more suitable for experimentally investigating garden-path effects. First, although both *DE* and *'s* are subject to coarticulation of its preceding morpheme, *DE* can be pronounced fully and independently from its surrounding morphemes. In other words, Mandarin-speakers can opt out of co-articulating *DE* and the naturalness of the sentence will not be affected. By contrast, *'s* has to be coarticulated with its preceding morpheme (i.e., *'s* is pronounced as /s/ after voiceless nonsibilant consonants, and /z/ after voiced nonsibilant consonants). This feature of *DE* is particularly helpful when dividing time windows for analysis in a visual world eye-tracking study.

Second, it has been suggested that grammatical morphemes like *–ed* and *'s* are particularly vulnerable for young English-speaking children, because these morphemes are shorter in duration and phonologically weaker than adjacent morphemes, and thus the perception of these weak morphemes exhausts the processing resources available to young children (Leonard, 2014; Leonard, Caselli, Bortolini, McGegor, & Sabbadini, 1992; Leonard, Eyer, Bedore, & Grela, 1997). By contrast, the perception of the Mandarin morpheme *DE* should presumably be less challenging for young Mandarin-speaking children, due to its acoustic and phonological features discussed above. Previous research has shown that Mandarin-speaking children start to use *DE* as a possessive marker at 2 years of age (Kong, Zhou, & Li, 1990; Li, 2004) and by 4 years of age they have acquired the syntactic and semantic features of the possessive *DE* construction (Shi & Zhou, 2018).

The present study took advantage of the property of Mandarin *DE* to investigate children's processing of garden-path constructions. In particular, we were interested to investigate whether young children are able to reanalyze garden-path constructions when the ambiguous word is adjacent to the disambiguation point.

Note that in previous research the ambiguous word and the disambiguation point are nonadjacent elements, and thus the linear distance between the two is relatively long. For example, in the garden-path constructions used by Trueswell et al. (1999; see example [3]), there were two words between the ambiguous word *on* and the disambiguation point *in*. Similarly, two elements intervened between the ambiguous case marker *–ey* and the disambiguation verb in the Korean garden-path constructions used by Choi and Trueswell (2010; see example [5]). The relatively long linear distance might have posed difficulties for young children, due to their limited working memory capacity.

To minimize children's difficulties with reanalysis associated with working memory load, the present study used garden-path structures as in (6), where the disambiguation point, the marker *DE*, is adjacent to the ambiguous word *xiaogou* "dog," and therefore the linear distance between the two elements is kept to minimum. By reducing the linear distance, this maneuver might presumably reduce the computational burden posed on working memory, according to Just and Carpenter (1992) and Lewis et al. (2006).

More specifically, in the present study we were interested to find out whether 4- to 5-year-old Mandarin-speaking children are already able to process Mandarin garden-path constructions in an adultlike manner, by maximizing their chances to revise and reanalyze their initial interpretation due to the features of the Mandarin construction discussed above. To our knowledge, this is the first experimental study to investigate Mandarin-speaking children's real-time processing of garden-path constructions.

The present study

Participants

Twenty-five Mandarin-speaking 4-year-olds (age range 4;1-4;11; mean 4;6) and 25 Mandarin-speaking 5-year-olds (age range 5;1–5;10; mean 5;6) participated in the study. They were recruited from the Beijing Taolifangyuan Kindergarten, and had no reported history of speech, hearing, or language disorders. In addition, 30 Mandarin-speaking adults (age range 18–24, mean 20) were tested as controls. They were all undergraduate students at Tsinghua University in Beijing, and had no self-reported speech, hearing, or language disorders. Four of the 25 4-year-olds did not complete the actual test, because they became distressed during the task, and refused to continue. Four of the 25 5-year-olds and 3 of the 30 adults did not proceed to the actual test, because we were unable to calibrate them on the eye tracker. The other participants successfully completed the task and were included in the final analyses.

The study was approved by the ethics committee of the School of Medicine, Tsinghua University, 20170018. Written informed consent has been obtained from each child participant's parents and each adult participant.

Materials and design

A total of 8 target items were created, each containing a visual image and a spoken sentence.[2] All the target spoken sentences had the same structure: "NP1 + Modal + Verb + NP2 + DE + NP3" as in (6), repeated here as (8). The modal word remained the same across trials, that is, the Mandarin *yaoqu* "will," denoting a future event. The verb was always monosyllabic in Mandarin and could take either an animate or inanimate entity as its plausible complement (e.g., *ti* "kick"). The morpheme *DE* was a possessive marker, so "NP2 + DE + NP3" indicated a possessive relation in which NP2 (e.g., *xiaogou* "dog") was the animate possessor and NP3 (e.g., *piqiu* "ball") was the inanimate possessee. All the NPs, including NP1, NP2, and NP3, were disyllabic in Mandarin. Each visual image consisted of five entities, including one animal corresponding to NP1 of the spoken sentence (e.g., the cat in Figure 1), and two possessor-possessee pairs, one was the target possessor–possessee pair and one was the contrast possessor–possessee pair. The target possessor–possessee pair (e.g., the two entities in the left panel of Figure 1) corresponded to NP2 (e.g., *xiaogou* "dog") and NP2 + DE + NP3 (e.g., *xiaogou DE piqiu* "the dog's ball"), where NP2 functioned as the modifier of the object NP2 + DE + NP3 in the target sentence. We refer to the two areas as the target modifier area (e.g., the area containing the dog) and the target object area (e.g., the area containing

Figure 1. An example of a target visual image in the study.

the dog’s ball). The contrast possessor–possessee pair also indicated a modification relation between a possessor and a possessee (e.g., the rooster and the rooster’s ball in the right panel of Figure 1). We refer to the two areas as the contrast modifier area (e.g., the area containing the rooster) and the contrast object area (e.g., the area containing the rooster’s ball).

In the visual image, the possessive relation was established by drawing an icon of the possessor (e.g., the dog) on the possessee (e.g., the dog’s ball). The animal character corresponding to NP1 always occurred at the center of the visual image, whereas the positions of the other four entities corresponding to the target and contrast possessor–possessee pairs were counterbalanced across the visual images.

(8) Xiaomao yaoqu ti xiaogou DE piqiu
cat will kick dog DE ball
“The cat is going to kick the dog’s ball.”

In addition, 8 control and 8 filler trials were constructed, each containing a visual image and a spoken sentence. The visual images of the control and filler trials were similar to those on the target trials. The control sentences had the following structure: NP1 + Modal + Verb + NP2 + Adverb. In the control sentences, all the NPs were disyllabic in Mandarin like in the target sentences, the modal verb was always *yaoqu* “will,” and the adverb was always *yixia* “once,” as in (9). In this sentence, *xiaogou* “the dog” was the object of the verb *ti* “kick,” so no reanalysis was involved. Control sentences such as (9) were used as a baseline condition, because the structure of the control sentences followed the structure of the target sentences up until the point of disambiguation, but crucially did not involve a garden path, thus serving a good baseline for how likely the participants were to look away from a particular image by chance or visual preference. The filler sentences had the following structure: NP1 + Modal + Verb + NP2 + HE + NP3, where the Mandarin conjunction word *he* "and" was used between NP2 and NP3, as in (10). With a conjunctive phrase *xiaoyang he yizi* "the goat and the chair," the sentence means that the deer is going to kick the goat and the chair. The target, control, and filler trials were presented to the participants in random order. A full list of target, control, and filler sentences can be found in Appendix A.

(9) Xiaoji yaoqu ti xiaogou yixia.
chicken will kick dog once
"The chicken is going to kick the dog once."

(10) Xiaolu yaoqu ti xiaoyang he yizi.
deer will kick goat and chair
"The deer is going to kick the goat and the chair."

To ensure that NP3 (e.g., *piqiu* "ball") was not more plausible than NP2 (e.g., *xiaogou* "dog") as the object of the verb (e.g., *ti* "kick") in the target sentences, we did a survey on 20 Mandarin-speaking adults (9 males and 11 females; age range 19–27; mean 23) where they were asked to rate the plausibility of the two verb–object pairs, "verb + NP2" (e.g., *ti xiaogou* "kick the dog") and "verb +NP3" (e.g., *ti piqiu* "kick the ball") in all the target sentences using a 5-point Likert scale, with 5 representing *the most plausible* and 1 representing *the least plausible*. The mean plausibility score for the "Verb + NP2" pair was 4.08 ($SD = 1.08$), and the mean plausibility score for the "Verb + NP3" pair was 3.89 ($SD = 1.33$). No significant difference was found between the mean scores of the two pairs ($p = .09$). In addition, the median scores for the two pairs were both 4. To further examine the distribution of the plausibility scores, we divided the scores into two clusters, the low score cluster and high score cluster. The low score cluster includes three scores, 1, 2, and 3, while the high score group includes two scores, 4 and 5. The proportion of low scores for the "Verb + NP2" pair was 28% (45 out of 160 responses), and the proportion of low scores for the "Verb + NP3" pair was 31% (50 out of 160 responses). All these comparisons showed that the scores for "Verb + NP3" pair were statistically similar to the scores for the "Verb + NP2" pair, indicating that NP3 was not more plausible than NP2 as the object of the verb.

Production of the test stimuli

All the spoken sentences were produced by a female native speaker of Beijing Mandarin. She was asked to produce these sentences word by word in a child-directed manner. The recording took place in a sound-attenuated recording booth at Tsinghua University. To ensure the consistency of prosodic features (i.e., duration and prosody) of each element across the spoken sentences, the original recorded sentences were later edited in Praat: for each element, only one sample in the recording was selected and then used for all the relevant sentences that contained the

Table 1. Duration analysis of the target sentences used in the study (standard deviation in parentheses)

Duration of the time window	Duration of the element	Duration of the inserted pause
2500 ms	NP1: 866 ms (74)	1634 ms
1500 ms	Modal (*yaoqu*): 860 ms (0)	640 ms
1500 ms	Verb: 506 ms (30)	994 ms
1500 ms	NP2: 929 ms (66)	571 ms
1200 ms	*DE*: 346 ms (0)	854 ms
1800 ms	NP3: 807 ms (126)	993 ms

Note: Each time window consisted of the element of the sentence and an inserted pause. The same recorded sample of the modal verb *yaoqu* "will" and the same recorded sample of the morpheme *DE* were used across the target sentences.

element. For instance, all the spoken sentences (including the target, control, and filler sentences) used the same sample of the modal verb *yaoqu* "will," all the target sentences used the same sample of the morpheme *DE*, all the control sentences used the same sample of the adverb *yixia* "once," and so on. This maneuver was also used to control for the effect of prosodic cues on sentence comprehension. In addition, to create clear time windows for each element in the sentences, we added pauses between each element such that each element had the time window of the same length across sentences: NP1 (2500 ms), the modal (1500 ms), the verb (1500 ms), NP2 (1500 ms), *DE* (1200 ms), *HE* (1200 ms), NP3 (1800 ms), and the adverb (3000 ms). In other words, each time window consisted of the element of the sentence and an inserted pause. Table 1 provides a duration analysis of each time window in the target sentences. Note that in order to keep consistent the prosodic patterns across the target, control, and filler sentences, pauses were also inserted between each element in the control and filler sentences. The only difference was that the morpheme *DE* only occurred in the target sentences; the adverb *yixia* "once" only occurred in the control sentences, and conjunction word *HE* only occurred in the filler sentences. All the spoken sentences were 10 s long.

To ensure the naturalness of the edited target and control sentences, we did a survey on 20 Mandarin-speaking adults (8 males and 12 females; age range 19–27; mean 23). In the survey, they were asked to judge the naturalness of the target and control sentences using a 5-point Likert scale, with 5 representing *the most natural* and 1 representing *the least natural*. The mean naturalness score for the target sentences containing *DE* was 4.24 ($SD = 0.90$), and the mean naturalness score for the control sentences was 4.36 ($SD = 0.85$). No significant difference was observed in the naturalness ratings of the target and control sentences ($p = .10$), indicating that these edited target and control sentences sounded natural and intelligible.

Procedure

Both children and adults were tested using the visual-world paradigm (Cooper, 1974; Tanenhaus et al., 1995). They were presented with a spoken sentence while viewing a visual image. They were instructed that they were going to see some pictures and the puppet (the little kitten), was going to tell them what would happen in these pictures. The participants' eye movements were recorded by an EyeLink 1000 eye tracker (by SR Research Ltd., Mississauga, Ontario, Canada) interfaced with a PC computer. The eye tracker allows remote eye tracking without a head support. It provides information about the participant's point of gaze at a sampling rate of 500 Hz and has an accuracy of 0.5 degrees of visual angle. The visual images were displayed on the PC monitor, and the spoken sentences were presented to the participants through two external speakers connected to the computer. The distance between the participants' eye and the monitor was about 60 cm.

Before the actual experiment, we had an introduction session where the participants were familiarized with task and the objects shown in the visual images. The participants were also instructed that if there was an icon of an animal on the object, then the animal owned the object. After the introduction session, the experimental session began. Before each trial, a black dot was shown at the center of the PC monitor, which anchored the beginning of the trial, and also served to capture the participants' attention.

The spoken sentences started 500 ms after the appearance of the visual stimulus. The participants' eye movements were recorded for 10 s from the onset of the sentence until the sentence was completed.

Predictions

If the participants incrementally computed the structural representation and possible meanings of the spoken sentences, when presented with the target sentences as in (8) they might initially analyze "NP1 + Modal + Verb + NP2" as a complete sentence, meaning "The cat is going to kick the dog," after hearing the verb *ti* "kick" and before encountering the marker *DE*. In other words, when processing (8), the participants might initially analyze NP2 *xiaogou* "dog" as the object NP of the verb *ti* "kick," rather than the modifier of the actual object NP *xiaogou DE piqiu* "dog's ball." This interpretation process would lead the participants to initially look more at the dog in Figure 1 (the target modifier area), after hearing the verb *ti* "kick" and before hearing the possessive marker *DE*. The possessive marker *DE* is the disambiguation point (trigger for reanalysis). Upon encountering the possessive marker *DE*, the participants would need to revise their initial analysis of NP2 (*xiaogou* "dog") and reanalyze it as the modifier of the object NP (*xiaogou DE piqiu* "dog's ball"). This reanalysis process would lead the participants to switch their eye movements from the dog to the dog's ball in Figure 1 (the target object area), so a significant increase of fixations in the target object area and a significant decrease of fixations in the target modifier area should be expected after the onset of *DE*. As discussed, to provide a baseline measure of how likely participants were to look away from a particular image by chance or visual preference, sentences like (9) were used as a baseline control condition.[3] If the participants were able to revise and reanalyze their initial interpretation, and then successfully recovered from the garden path in the target sentences, then they should be expected to exhibit more looks to the target object area (e.g., the dog's ball in Figure 1) when hearing *DE* in the target sentences than when hearing the adverb *yixia* "once" in the control sentences;

by contrast, an opposite pattern should be observed in the target modifier area (e.g., the dog in Figure 1): hearing *DE* in the target sentences should trigger fewer looks to this area than hearing *yixia* “once” in the control sentences.

Results

To analyze the eye movement data, we first defined five equal-sized areas of interest in the visual image: the *Agent* area (corresponding to NP1 in the spoken sentence), the target modifier area (*Target_Mod*, corresponding to NP2 in the sentence), the target object area (*Target_Obj*, corresponding to NP2 + DE + NP3), the contrast modifier area (*Contrast_Mod*), and the contrast object area (*Contrast_Obj*). The contrast modifier and the contrast object areas corresponded to the other possessor–possessee pairs depicted in the visual image. As discussed in the Materials and Design section, on the example target trial the five areas of interest referred, respectively, to the cat (*Agent*), the dog (*Target_Mod*), the dog’s ball (*Target_Obj*), the rooster (*Contrast_Mod*), and the rooster’s ball (*Contrast_Obj*).

In preparing the eye movement data, we deleted the samples where the participants’ eye movements were not detected, for example, when they blinked their eyes. This process affected approximately 10% of the recorded data. To reduce the number of statistical tests carried out, we then down-sampled the data into a series of time bins, each with a duration of 50 ms. After that we computed the proportion of fixations for each area of interest under each temporal bin, for each participant and each trial. The proportion of fixations for a particular area of interest in a specific temporal bin was treated as the dependent variable. For example, if 5 fixation points in a temporal bin were recorded, with 2 fixation points located in that specific area of interest, then the proportion of fixations on that area was 2/5.

To visually present the data, we first averaged the coded data for all the trials and participants in each sample point under each condition and each age group. The results are summarized in Figure 2, where the target (i.e., Target [DE]) and control sentences (i.e., Control [Yixia]) are represented using solid and dotted lines, respectively, and the average fixation proportions in the two areas of interest, *Target_Mod* and *Target_Obj*, are presented in the left and right columns respectively. As indicated in the left column of Figure 2, hearing the possessive marker *DE* in the target sentences triggered fewer fixations on the *Target_Mod* area than hearing the adverb *yixia* in the control sentences for all the three age groups. In contrast, an opposite pattern was observed in the *Target_Obj* area, as shown in the right column of Figure 2. All the three age groups exhibited more looks to this area when hearing *DE* in the target sentences than when hearing *yixia* in the control sentences. As predicted, for all the three age groups, hearing the possessive marker *DE* switched the participants’ eye movements from the *Target_Mod* area to the *Target_Obj* area, indicating that the 4-year-olds and the 5-year-olds, like the adults, were able to revise and reanalyze their initial interpretation using the information encoded in the possessive marker *DE*, and thus successfully recovering from the garden path in the target sentences.

To statistically examine the observed effects, we first transformed the fixation proportions using the empirical logit formula (Barr, 2008): $probability = ln([y+0.5]/[n-y+0.5])$, where y is the number of samples in which the participants’

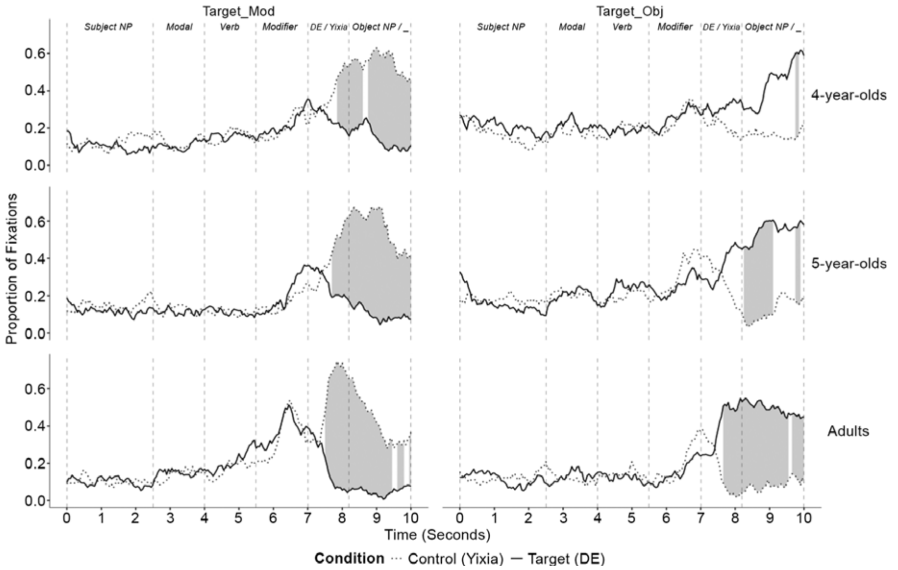

Figure 2. Average fixation proportions in the *Target_Mod* area (left column) and in the *Target_Obj* area (right column) by the 4-year-olds (upper panel), the 5-year-olds (middle panel), and the adults (lower panel). For illustration purposes, the *y*-axis gives the original mean proportions of fixations, instead of the transformed ones. The gray areas indicate significant differences between the target and control baseline conditions on the basis of the adjusted *p* values ($p < .05$).

fixation was located in a specific area of interest during a particular temporal bin; n is the total number of samples where the participants' eye fixations were recorded. To compare the target and control conditions, we then fitted a linear mixed-effects model (LMM) to the transformed data, under each temporal bin, each area of interest, and each age group. The LMM model contained only one fixed term, *condition*, and two random terms, *participant* and *trial*. The model is the maximum one as suggested by Barr, Levy, Scheepers, and Tily (2013): *Transformed-Proportion ~ 1 + condition + (1 + condition | participant) + (1 + condition | trial)*. The fitting process was conducted via the *MixedModels* package (Bates et al., 2019) in *Julia* language (Bezanson, Edelman, Karpinski, & Shah, 2017). The obtained *p* values were then Bonferroni adjusted, that is, the obtained *p* value times the number of comparisons in a specific area of interest and a specific age group. The model results are also summarized in Figure 2, where the gray areas indicate significant differences between the two conditions on the basis of the adjusted *p* values. The model results confirmed the observed eye gaze patterns. Note that for illustration purposes, the *y*-axis of Figure 2 displays the original mean proportions of fixations, instead of the transformed ones.

To statistically analyze the latencies of the obtained effects between different age groups, we first identified the latency for each participant and each area of interest by applying the LMM model, *Transformed-Proportion ~ 1 + condition + (1 | trial)*, to each temporal bin for each participant and each area of interest. We then compared the obtained latencies for each participant by applying the LMM model, *Latency ~ 1 + age_group + (1 | participant)*, to each area of interest. Using the 5-year-olds as the baseline, we found that the observed effects in the *Target_Obj* area occurred significantly earlier in the 5-year-olds than in the 4-year-olds (8.25 s vs. 9.65 s, $b = 14.78$, $z = 3.13$, $p < .01$), but compared with adults, the observed effects in the 5-year-olds occurred significantly later (7.65 s vs. 8.25 s, $b = -10.59$, $z = -2.92$, $p < .01$). The results indicated that it took longer time for the younger children than the older children and for the older children than the adults to revise and reanalyze their initial misinterpretation.

Discussion

The present study sought to investigate whether 4- and 5-year-old Mandarin-speaking children are able to process Mandarin garden-path constructions associated with the grammatical morpheme *DE*. The obtained eye gaze patterns show that the 4- and 5-year-olds, like the adults, committed to an initial misinterpretation and later successfully revised their initial interpretation when encountering the morpheme *DE*. This is the first experimental study that investigated Mandarin-speaking children's real-time processing of garden-path constructions and observed that preschool Mandarin-speaking children successfully recovered from the garden path in the relevant constructions. In addition, the findings are consistent with previous research observing 5- to 8-year-old children' successful recovery from their initial misinterpretation using later-encountered linguistic information, when processing filler-gap sentences involving local ambiguity (e.g., Özge et al., 2015).

One possible explanation for children's successful recovery is the minimized linear distance between the ambiguous word and the disambiguation point in the sentence. As discussed in the two working memory models, the linear distance between the ambiguous word and the disambiguation point positively correlates with the level of difficulties in reanalyzing garden-path constructions. When the ambiguous word is adjacent to the disambiguation point, that is, when the linear distance between the two is minimized, the difficulties in reanalysis are also reduced to minimum. In most previous research, the ambiguous word and the disambiguation point are nonadjacent elements, and thus the linear distance between the two is relatively long. For instance, there were two words between the ambiguous word and the disambiguation point in the garden-path constructions used by Trueswell et al. (1999; see example [3]), and two elements intervened between the ambiguous case marker and the disambiguation verb in the Korean garden-path constructions used by Choi and Trueswell (2010; see example [5]). The relatively long linear distance might have posed difficulties for young children, because young children have limited working memory capacity and thus are more likely to abandon the correct interpretation before the disambiguation point (on the account by Just & Carpenter, 1992) or they might exhibit more difficulties in reactivating the correct interpretation due to longer decaying time (on the account by Lewis et al., 2006).

Unlike previous research, the present study took advantage of the Mandarin garden-path construction, where the disambiguation point, the possessive marker *DE*, is adjacent to the ambiguous word, and thus the linear distance between the two elements is kept to minimum. By reducing the linear distance, this maneuver might presumably reduce the computational burden posed on working memory, because shorter linear distance might significantly increase the chance for young children to hold the correct interpretation in working memory (according to Just & Carpenter, 1992), or might reduce the decaying time of the correct interpretation so that it becomes easier for young children to reactivate (according to Lewis et al., 2006).

Alternatively, the reduction of children's working memory burden in reanalysis might also be linked to the syntactic structure of the sentences in the current study, in addition to the adjacency between the ambiguous word and the disambiguation point. According to the processing model by Pritchett (1988, 1992), difficulties in reanalysis are related to the number of syntactic nodes involved in the reanalysis process: more syntactic nodes involved in reanalysis pose more difficulties in reanalysis. More specifically, when encountering the ambiguous word, the parser assigns an initial position to it in the syntactic structure. As the parsing continues, the parser has to reanalyze the syntactic structure of the sentence when encountering the disambiguation point, whereby the ambiguous word has to be removed from its initial syntactic position and be relocated to the revised syntactic position. If the initial syntactic position and the revised syntactic position reside under the same syntactic node (e.g., both are under the same NP node), then reanalysis poses relatively lower burden on working memory, because it only involves the processing of one syntactic node. By contrast, if the initial syntactic position and the revised syntactic position do not reside in the same syntactic node (e.g., the initial syntactic position is under an NP node, and the revised position under a verb phrase node), then reanalysis poses relatively higher burden on working memory, because it engages the processing of different syntactic nodes.

Compared with the classic garden-path structures such as (1) and the structures used in prior research, the Mandarin garden-path constructions in the current study might be structurally simpler, and thus are relatively easier to be processed. In the English garden-path constructions in Trueswell et al. (1999), the ambiguous phrase *on the napkin* was first analyzed as the destination of the verb under the verb phrase node, and later reanalyzed as the modifier of the noun under the NP node when encountering the disambiguation point. As the initial syntactic position and the revised syntactic position do not reside in the same syntactic node, the reanalysis process would probably induce high working memory burden. By contrast, in the current study, the ambiguous word (e.g., *xiaogou* "dog") was initially analyzed as the object noun under the object NP node. Encountering the disambiguation word *DE* led the parser to reanalyze it as the modifier of the object noun under the same object NP node. As the initial and the revised syntactic positions are under the same object NP node, reanalysis involved only one syntactic node and thus induced relatively lower working memory burden.

Nonetheless, we wish to note that although the features of the Mandarin garden-path construction facilitated children's real-time comprehension of the construction, the younger children exhibited more difficulties than the older children and the adults in revising their initial misinterpretation using the later-encountered linguistic information, as evidenced by the finding that the effect of fixating more on the *Target_Obj* area (indicating the correct interpretation) occurred significantly later in the 4-year-olds than in the 5-year-olds and the adults. Our findings suggest that although the 4-year-olds successfully recovered from the garden-path in the target constructions, they were not as effective as the older children and the adults in revising and reanalyzing the initial misinterpretation even when the working memory burden was kept to minimum, indicating that in addition to working memory, other cognitive factors, such as children's immature cognitive control ability, also play a role in children's processing of garden-path constructions (Choi & Trueswell, 2010; Kidd et al., 2011; Mazuka et al., 2009; Novick et al., 2005; Omaki et al., 2014; Trueswell et al., 1999; Weighall, 2008; Woodard et al., 2016).

The findings of the current study open up new questions for building a fine-grained model of child sentence processing. To the best of our knowledge, most previous research attributed the kindergarten-path effect to children's immature cognitive abilities like limited working memory capacity or immature cognitive control ability, without specifying how exactly each cognitive component relates to children's difficulties with reanalysis. To better understand the kindergarten-path effect, a fine-grained child sentence processing model is required in which the respective roles of working memory and cognitive control are spelled out in detail. The present study is an attempt in this direction by investigating whether children could revise their initial interpretation when the working memory burden is reduced to minimum, regardless of linear distance or structural reasons. To confirm the role of linear distance and structural property in children's processing of garden-path constructions, future research is required to tease apart these two factors by investigating how each of these two factors contributes to the reduction of working memory burden in reanalysis.

We also wish to acknowledge a few limitations of the current study. The current study did not measure the participants' working memory capacity, because we assume that young children and adults differ in their working memory capacity on the basis of the general consensus in previous research that young children have more limited working memory capacity as compared to adults (e.g., Case et al., 1982; Gathercole, 2004). Future research is required to directly investigate the relation between distance effect and working memory capacity. In addition, our study only focused on the working memory burden, without considering in detail the nature of the ambiguous word/the disambiguation point (e.g., whether the frequent occurrence of the morpheme *DE* in the adult input helped children's reanalysis). Again, further research is needed to examine how the nature of the ambiguous word/the disambiguation point relates to children's processing of garden-path constructions. Finally, we wish to note that the observed effects in the data analysis may not have come out if a different *p* value adjustment method was used or a more conservative hypothesis was adopted for the Bonferroni correction (e.g., without adopting the assumption that data from different areas of interest are independent). Future studies using the visual-world paradigm might consider using a more appropriate adjustment method when dealing with the auto-correction problem between the time bins.

Overall, the present study advances our understanding of children's difficulty with reanalysis in processing garden-path constructions by showing that 4-year-old children can already successfully revise their initial misinterpretation and then arrive at the correct interpretation using the later-encountered linguistic information, when the working memory burden is kept to minimum as compared to previous research.

Acknowledgments. This work was supported by National Social Science Foundation of China Grant number 16BYY076 (to Peng Zhou). The authors would like to thank the children, the parents, and the teachers at the Taolifangyuan Kindergarten, Beijing, China, for their assistance and support in running the study.

Notes

1. An anonymous reviewer pointed out that in addition to the garden-path construction discussed in our study, there are other garden-path constructions in Mandarin that are associated with relative clauses containing the grammatical morpheme *DE*. We thank the reviewer for pointing this out. We are fully aware of garden-path constructions of this type, but we wish to note that there are two reasons for not testing children's comprehension of garden-path constructions containing relative clauses. First, relative clauses are already complex for young children. It has been reported that 5-year-old Mandarin-speaking children still have difficulties in producing and comprehending relative clauses (see, e.g., Hu, Gavarró, & Guasti, 2016; Hu, Gavarró, Vernice, & Guasti, 2016). It would pose even more difficulties for young children if garden-path constructions containing relative clauses were used. Second, garden-path constructions with relative clauses are not ideal for a child visual-world paradigm study, because it would be fairly hard to depict complex structures like this in a visual image.

2. All the test stimuli and the original eye movement data can be found in Open Science Framework via the link: http//www.osf.io/32mvf.

3. A second analysis using a different baseline condition is provided in Appendix B. In this second analysis, the comparison of looks to *the Target_Mod/Target_Obj* areas in the interest period before the verb was used as the baseline. We thank one anonymous reviewer for this suggestion to have more than one baseline to help place the current results. Overall, the current analysis in the main text and the analysis in Appendix B gave consistent results, though using different baseline conditions, confirming the reliability of our findings and the interpretation of the findings.

References

Altmann, G., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, **73**, 247–264.

Altmann, G., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, **57**, 502–518.

Andreu, L., Sanz-Torrent, M., & Trueswell, J. C. (2013). Anticipatory sentence processing in children with specific language impairment: Evidence from eye movements during listening. *Applied Psycholinguistics*, **34**, 5–44.

Barr, D. J. (2008). Analyzing "visual world" eye tracking data using multilevel logistic regression. *Journal of Memory and Language*, **59**, 457–474.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, **68**, 255–278.

Bates, D., Calderón, J. B. S., Noack, A., Kleinschmidt, D., Kelman, T., Bouchet-Valat, M., … Baldassari, A. (2019). dmbates/MixedModels.jl: v2.1.1 | Zenodo. doi: 10.5281/zenodo.3428819

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48.

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, **59**, 65–98.

Boland, J. E., Tanenhaus, M. K., & Garnsey, S. M. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language*, **29**, 413.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, **33**, 386–404.

Choi, Y., & Trueswell, J. C. (2010). Children's (in)ability to recover from garden paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of Experimental Child Psychology*, **106**, 41–61.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, **6**, 84–107.

Crain, S., & Steedman, M. K. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives*. Cambridge, MA: Cambridge University Press.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97–135). Amsterdam: Benjamins.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, **47**, 164–203.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, **25**, 348–368.

Ferreira, F., & Lowder, M. W. (2016). Prediction, information structure, and good enough language processing. *Psychology of Learning and Motivation*, **65**, 217–247.

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies*. Unpublished doctoral dissertation, University of Connecticut.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading*. Hillsdale, NJ: Erlbaum.

Frazier, L. (1989). Against lexical generation of syntax. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, **14**, 178–210.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, **40**, 177.

Hu, S., Gavarró, A., & Guasti, M. T. (2016). Children's production of head-final relative clauses: The case of Mandarin. *Applied Psycholinguistics*, **37**, 323–346.

Hu, S., Gavarró, A., Vernice, M., & Guasti, M. T. (2016). The acquisition of Chinese relative clauses: Contrasting two theoretical approaches. *Journal of Child Language*, **43**, 1–21.

Huang, Y. T., & Hollister, E. (2019). Developmental parsing and linguistic knowledge: Reexamining the role of cognitive control in the kindergarten path effect. *Journal of Experimental Child Psychology*, **184**, 210–219.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, **99**, 122–149.

Kamide, Y., Altmann, G., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, **49**, 133–156.

Kidd, E., & Bavin, E. L. (2005). Lexical and referential cues to interpretation: An investigation of children's interpretations of ambiguous sentences. *Journal of Child Language*, **32**, 855–876.

Kidd, E., & Bavin, E. L. (2007). Lexical and referential influences on on-line spoken language comprehension: A comparison of adults and primary school-age children. *First Language*, **27**, 29–52.

Kidd, E., Stewart, A. J., & Serratrice, L. (2011). Children do not overcome lexical biases where adults do: The role of the referential scene in garden-path recovery. *Journal of Child Language*, **38**, 222–234.

Kong, L., Zhou, G., & Li, X. (1990). Investigation on the use of DE by 1 to 5 years old children. *Psychological Science*, **6**, 14–20.

Lassotta, R., Omaki, A., & Franck, J. (2016). Developmental changes in misinterpretation of garden-path *wh*-questions in French. *Quarterly Journal of Experimental Psychology*, **69**, 829–854.

Lee, T. H.-T. (2006). A note on garden path sentences in Chinese. In D.-A. Ho, S. Cheung, W. Pan, & F. Wu (Eds.), *Linguistic studies in Chinese and neighboring languages: Festschrift in honor of Professor Pang-Hsin Ting on his seventieth birthday* (pp. 491–518). Taipei: Institute of Linguistics, Academia Sinica.

Leonard, L. B. (2014). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Leonard, L. B., Caselli, M. C., Bortolini, U., McGregor, K. K., & Sabbadini, L. (1992). Morphological deficits in children with specific language impairment: The status of features in the underlying grammar. *Language Acquisition*, **2**, 151–179.

Leonard, L. B., Eyer, J. A., Bedore, L. M., & Grela, B. G. (1997). Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, **40**, 741–753.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, **18**, 193–198.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, **29**, 375–419.

Lewis, R. L., Vasishth, S., & van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, **10**, 447–454.

Li, Y. (2004). *The development of child language*. Wuhan: Huazhong Normal University Press (in Chinese).

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, **9**, 157–201.

Mazuka, R., Jincho, N., & Onishi, H. (2009). Development of executive control and language processing. *Language and Linguistics Compass*, **3**, 59–89.

Meroni, L., & Crain, S. (2003). On not being led down the kindergarten path. In *Proceedings of the 27th Boston University Conference on Language Development* (pp. 531–544). Somerville, MA: Cascadilla Press.

Nation, K., Marshall, C. M., & Altmann, G. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology*, **86**, 314–329.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, **5**, 263–281.

Omaki, A. (2010). *Commitment and flexibility in the developing parser*. Unpublished doctoral dissertation, University of Maryland.

Omaki, A., Davidson White, I., Goro, T., Lidz, J., & Phillips, C. (2014). No fear of commitment: Children's incremental interpretation in English and Japanese *wh*-questions. *Language Learning and Development*, **10**, 206–233.

Özge, D., Küntay, A., & Snedeker, J. (2019). Why wait for the verb? Turkish speaking children use case markers for incremental language comprehension. *Cognition*, **183**, 152–180.

Özge, D., Marinis, T., & Zeyrek, D. (2015). Incremental processing in head-final child language: Online comprehension of relative clauses in Turkish-speaking children and adults. *Language, Cognition and Neuroscience*, **30**, 1230–1243.

Pickering, M. J., Traxler, M. J., & Crocker, M. W. (2000). Ambiguity resolution in sentence processing: Evidence against frequency-based accounts. *Journal of Memory and Language*, **43**, 447–475.

Pritchett, B. L. (1988). Garden path phenomena and the grammatical basis of language processing. *Language*, **64**, 539–576.

Pritchett, B. L. (1992). *Grammatical competence and parsing performance*. Chicago, IL: University of Chicago Press.

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Sekerina, I. A., & Trueswell, J. C. (2012). Interactive processing of contrastive expressions by Russian children. *First Language*, **32**, 63–87.

Shi, J., & Zhou, P. (2018). How possessive relations are mapped onto child language: A view from Mandarin Chinese. *Journal of Psycholinguistic Research*, **47**, 1321–1341.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, **49**, 238–299.

Staub, A., & Clifton, C., Jr. (2006). Syntactic prediction in language comprehension: Evidence from *either … or*. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **32**, 425–436.

Tabor, W., & Hutchins, S. (2004). Evidence for self-organized sentence processing: Digging-in effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **30**, 431.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, **268**, 1632–1634.

Taraban, R., & McClelland, J. L. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *Journal of Memory and Language*, **27**, 597–632.

Traxler, M. (2002). Plausibility and subcategorization preference in children's processing of temporarily ambiguous sentences: Evidence from self-paced reading. *Quarterly Journal of Experimental Psychology*, **55**, 75–96.

Traxler, M. J. (2005). Plausibility and verb subcategorization in temporarily ambiguous sentences: Evidence from self-paced reading. *Journal of Psycholinguistic Research*, **34**, 1–30.

Trueswell, J. C., Sekerina, I., Hill, N. M. & Logrip, L. (1999). The kindergarten-path effect: Studying online sentence processing in young children. *Cognition*, **73**, 89–134.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, **33**, 285.

van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **31**, 443–467.

van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, **49**, 285–316.

van Gompel, R. P. G., & Pickering, M. J. (2007). Syntactic parsing. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 289–307). Oxford: Oxford University Press.

van Heugten, M., & Shi, R. (2009). French-learning toddlers use gender information on determiners during word recognition. *Developmental Science*, **12**, 419–425.

Weighall, A. R. (2008). The kindergarten path effect revisited: Children's use of context in processing structural ambiguities. *Journal of Experimental Child Psychology*, **99**, 75–95.

Woodard, K., Pozzan, L., & Trueswell, J. C. (2016). Taking your own path: Individual differences in executive function and language processing skills in child learners. *Journal of Experimental Child Psychology*, **141**, 187–209.

Zhan, L. (2018). Scalar and ignorance inferences are both computed immediately upon encountering the sentential connective: The online processing of sentences with disjunction using the visual world paradigm. *Frontiers in Psychology*, **9**, 61.

Zhou, P., Crain, S., & Zhan, L. (2014). Grammatical aspect and event recognition in children's online sentence comprehension. *Cognition*, **133**, 262–276.

Zhou, P., Ma, W., Zhan, L., & Ma, H. (2018). Using the visual world paradigm to study sentence comprehension in Mandarin-Speaking children with autism. *Journal of Visualized Experiments*, **140**, e58452.

Appendix A

Target, control and filler sentences in the study (8 target, 8 control, and 8 filler sentences)

Target sentences

(1) Xiaomao yaoqu ti xiaogou DE piqiu
cat will kick dog DE ball
"The cat is going to kick the dog's ball."

(2) Xiaolu yaoqu mo xiaoyang DE maozi
deer will touch goat DE hat
"The deer is going to touch the goat's hat."

(3) Tuzi yaoqu pai xiaozhu DE zhuozi
rabbit will pat pig DE table
"The rabbit is going to pat the pig's table."

(4) Houzi yaoqu tui tuzi DE yizi
monkey will push rabbit DE chair
"The monkey is going to push the rabbit's chair."

(5) Xiaozhu yaoqu zhuang xiaoniao DE beizi
pig will bump bird DE cup
“The pig is going to bump the dog’s cup.”
(6) Xiaoma yaoqu bei xiaoji DE benzi
horse will carry chicken DE book
“The horse is going to carry the chicken’s book.”
(7) Hema yaoqu peng xiaolang DE qiche
hippo will hit wolf DE car
“The hippo is going to hit the wolf’s car.”
(8) Xiaoyang yaoqu la xiaolu DE qiqiu
lamb will pull deer DE balloon
“The lamb is going to pull the deer’s balloon.”

Control sentences

(9) Xiaoji yaoqu ti xiaogou yixia
chicken will kick dog once
“The chicken is going to kick the dog once.”
(10) Houzi yaoqu mo xiaoyang yixia
monkey will touch lamb once
“The monkey is going to touch the lamb once.”
(11) Hema yaoqu pai xiaozhu yixia
hippo will pat pig once
“The hippo is going to pat the pig once.”
(12) Xiaomao yaoqu tui tuzi yixia
cat will push rabbit once
“The cat is going to push the rabbit once.”
(13) Xiaogou yaoqu zhuang xiaoniao yixia
dog will bump bird once
“The dog is going to bump the bird once.”
(14) Xiaoyang yaoqu bei xiaoji yixia
lamb will carry chicken once
“The lamb is going to carry the chicken once.”
(15) Xiaoniao yaoqu peng xiaolang yixia
bird will hit wolf once
“The bird is going to hit the wolf once.”
(16) Tuzi yaoqu la xiaolu yixia
rabbit will pull deer once
“The rabbit is going to pull the deer once.”

Filler sentences

(17) Xiaolu yaoqu ti xiaoyang HE yizi
deer will kick goat and chair
“The deer is going to kick the goat and the chair(s).”
(18) Xiaomao yaoqu mo xiaoniao HE piqiu
cat will touch bird and ball
“The deer is going to touch the bird and the ball(s).”
(19) Xiaoniao yaoqu pai xiaomao HE maozi
bird will pat cat and hat
“The bird is going to pat the cat and the hat(s).”

(20) Xiaolang yaoqu tui xiaoma HE qiche
wolf will push horse and car
"The wolf is going to push the horse and the car(s)."

(21) Hema yaoqu zhuang xiaolu HE zhuozi
hippo will bump deer and table
"The hippo is going to bump the deer and the table(s)"

(22) Houzi yaoqu bei hema HE beizi
monkey will carry hippo and cup
"The monkey is going to carry the hippo and the cup(s)."

(23) Xiaoji yaoqu peng xiaozhu HE qiqiu
chicken will hit pig and balloon
"The hippo is going to hit the pig and the balloon(s)."

(24) Xiaoyang yaoqu la xiaogou HE benzi
lamb will pull dog and book
"The lamb is going to pull the dog and the book(s)."

Appendix B

This appendix contains the original analysis that used the comparison of looks to the *Target_Mod/Target_Obj* areas in the interest period before the verb as the baseline. We have decided to keep the original analysis by following an anonymous reviewer's suggestion to have more than one baseline to help place the current results. Overall, the current analysis in the main text and this original analysis gave consistent results, though using different baseline conditions, confirming the reliability of our findings and the interpretation of the findings.

Analysis

To examine the effect of the verb and the possessive marker *DE*, we *baseline-centered* the transformed proportions, that is, for each trial and each participant we subtracted the mean of transformed proportions prior to the onset of the verb from each obtained value after the onset of the verb. We then fitted a LMM to the baseline-centered data, for each temporal bin, each area of interest, and each age group. The LMM model contained only one fixed term, *intercept*, and two random terms, *participant* and *trial.* Since the empirical logit function is monotonic increasing, an intercept that is significantly bigger than zero means that the current proportion is significantly bigger than that of the baseline. The model fitting process was conducted via the *lmer* function from the *lme4* package (Bates, Mächler, BolKer, &Walker, 2015) under the R environment (R Core Team, 2019). The model formula used in R is $Transform\text{-}and\text{-}Centered\text{-}Proportion \sim 1 + (1|Participant) + (1|Trial)$. The p values were then obtained using Wald z tests, that is, the statistics is hypothesized to have a normal distribution with the parameter as its mean, and the standard error as its standard deviation. The obtained p values were then Bonferroni adjusted, that is, multiplying the obtained p values by 200 (the number of comparisons in each area of interest). The adjusted results are represented using colored horizontal lines in the Appendix Figure A.1, where the red line represents the 4-year-olds, the green line the 5-year-olds, and the blue line the adults. A temporal period that has a colored horizontal line indicates that for the relevant age group, there was a significantly higher fixation proportion than the baseline

Table A.1. The onset and offset of the significant effects represented by the colored horizontal lines in Figure A.1 for each group in each area of interest (seconds from the onset of the verb)

Area of interest	Age group	Onset (s)	Offset (s)
Target_Mod			
	4-year-olds	0.40	4.90
	5-year-olds	2.70	3.60
	Adults	0.95	3.55
Target_Obj			
	4-year-olds	5.80	6.00
	5-year-olds	2.65	6.00
	Adults	3.15	6.00

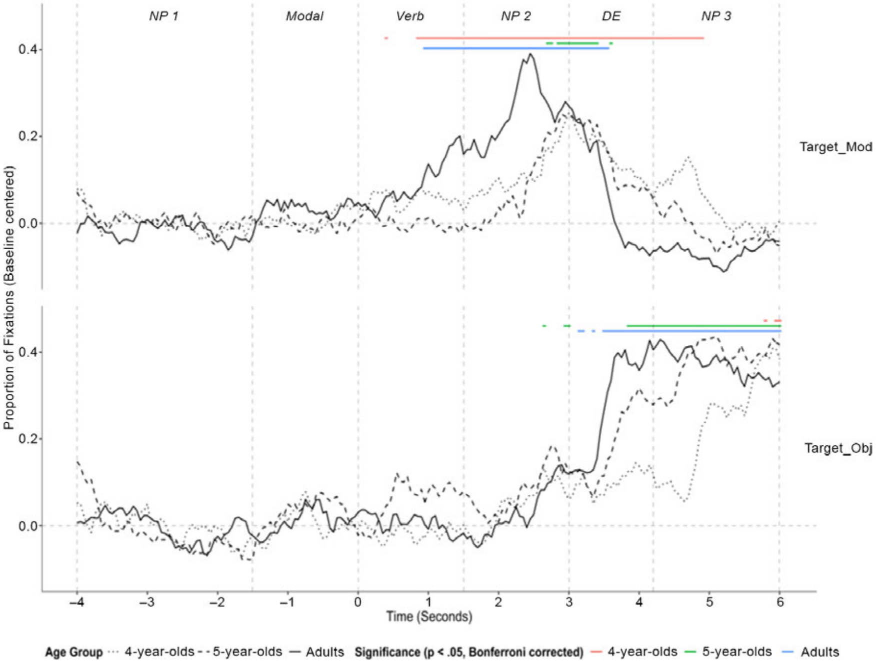

Figure A.1. (color online) Average fixation proportions in the *Target_Mod* area (upper panel) and in the *Target_Obj* area (lower panel) by the 4-year-olds (dotted line), the 5-year-olds (dashed line), and the adults (solid line). The illustrated proportions are baseline centered, that is, the mean fixation proportion in that area of interest prior to the onset of the verb is subtracted from the original proportions. The colored horizontal lines of each panel indicate that for the relevant age group, there was a significantly higher fixation proportion than the baseline in this area of interest during this temporal bin; the red line represents the 4-year-olds, the green line the 5-year-olds, and the blue line the adults.

(i.e., the fixation proportion prior to the onset of the verb) in this area of interest during this temporal bin. The onset and offset of the observed effects are summarized in the Appendix Table A.1.

The model results (see both Figure A.1 and Table A.1) show that the trend to look more at the *Target_Mod* area appeared relatively earlier in the 4-year-olds (0.40 s) and the adults (0.95 s) than in the 5-year-olds (2.70 s), and it disappeared relatively later in the 4-year-olds (4.90 s) than in the adults (3.55 s) and the 5-year-olds (3.60 s), indicating that it took longer time for the younger children to revise and reanalyze their initial misinterpretation. The effects in the *Target_Obj* area further confirmed this processing difficulty exhibited by the younger children. The 4-year-olds started to fixate more on the *Target_Obj* area (5.80 s) relatively later than the adults (3.15 s) and the 5-year-olds (2.65 s).

Cite this article: Zhou, P., Shi, J., and Zhan, L. (2021). Real-time comprehension of garden-path constructions by preschoolers: A Mandarin perspective. *Applied Psycholinguistics* **42**, 181–205. https://doi.org/10.1017/S0142716420000697