ORIGINAL ARTICLE

Comprehension of relative clauses in Chinese children with and without developmental dyslexia

Shenai Hu[1,2] ⓘ, Maria Teresa Guasti[2] ⓘ and Jing Zhao[3] ⓘ

[1]College of Foreign Languages and Cultures, Xiamen University, Xiamen, China, [2]Department of Psychology, University of Milano-Bicocca, Milan, Italy and [3]School of Psychology, Capital Normal University, Beijing, China
Corresponding authors: Shenai Hu; Email: shenai.hu@xmu.edu.cn, Jing Zhao; Email: conanzj@126.com

Abstract

This study investigated the comprehension of relative clauses (RCs) in Chinese children with and without developmental dyslexia (DD). Twenty-two children with DD, 22 chronological age-matched (CA) children, and 22 younger reading-level-matched (RL) children completed an RC comprehension task (measuring both accuracy and response latency), a receptive vocabulary task, and a working memory task. Results show that all three groups comprehended subject RCs more accurately than object RCs, supporting featural Relativized Minimality's prediction that structural intervention (i.e., syntactic configurations where an intervening element blocks dependency formation) is a crucial factor in children's RC comprehension. The DD group performed less accurately and more slowly on both structures compared to the CA group, but performed similarly to the RL group. Dyslexic children's receptive vocabulary knowledge was associated with higher accuracy and shorter response latencies in RC comprehension, and their phonological short-term memory was specifically linked to faster RC processing. These findings confirm the existence of syntactic difficulties in dyslexia and suggest that these difficulties may stem from limited vocabulary knowledge and phonological short-term memory deficits.

Keywords: Child language disorders; memory and language; spoken language comprehension

Introduction

Relative clauses (RCs) have been recognized as an important probe into language acquisition because of their syntactic complexity, and have received attention in dyslexic literature (e.g., Arosio et al., 2017; Byrne, 1981; Robertson & Joanisse, 2010; Shankweiler et al., 1984, 1995; Smith et al.,1989; Stella & Engelhardt, 2021). Most

dyslexic studies examined languages with head-initial RCs (e.g., English and Italian), with few studies focusing on head-final RCs (Chan, 2014, 2015). Cross-linguistically, head-final RCs are extremely rare in verb–object (VO) languages, with relevant exceptions such as Mandarin Chinese (henceforth Chinese), Cantonese and Wu (Dryer, 2013; Hu et al., 2018a). In the current study, we investigated the comprehension of RCs in Chinese children with DD, comparing with typically developing (TD) children matched for age or reading level, and explored their potential associations with vocabulary knowledge and verbal working memory. In this way, we gain a more in-depth understanding of the relationship between reading impairment and difficulties in complex syntactic processing, and help to disentangle the underlying causes of the syntactic deficits in dyslexia.

The article is organized as follows. We first briefly introduce the comprehension of subject relative clauses (SRCs) and object relative clauses (ORCs) in TD children and theories accounting for the asymmetry between two structures. Then, we reviewed previous studies on dyslexic children's comprehension of RCs and factors that may influence their processing of complex syntax. Next, we present the current study, and finally, offer a general discussion.

The comprehension of relative clauses and theoretical accounts

The comprehension of RCs has been extensively examined in TD children across a variety of languages. Specifically, studies often compare SRCs and ORCs with animate noun phrases (NPs). In languages with head-initial RCs, as illustrated in (1), a consistent SRC advantage has been observed (Tanaka et al., 2024). However, in languages with head-final RCs such as Chinese, as shown in (2), most studies report a similar SRC advantage (e.g., Hu et al., 2016), though some studies have found no such advantage (e.g., He et al., 2017). These conflicting findings have been attributed to various factors, such as word order, experimental tasks, and individual differences (e.g., Hu et al., 2020).

(1) a. the child$_i$ that t_i draws the teacher (English SRC)
    b. the child$_i$ that the teacher draws t_i (English ORC)

(2) a. t_i hua laoshi de haizi$_i$ (Chinese SRC)
           draw teacher REL child
      "the child that draws the teacher"
    b. laoshi hua t_i de haizi$_i$ (Chinese ORC)
      teacher draw REL child
      "the child that the teacher draws"

Various theories have been proposed for the subject-object asymmetry in the processing of RCs. In the present paper, we focus on two influential versions of the theories, namely, the Dependency Locality Theory (DLT; Gibson, 1998, 2000) and the featural Relativized Minimality (fRM; Rizzi, 1990, 2004, 2018). One reason is that both theories deal with filler–gap dependencies, but provide different predictions with respect to the processing of head-initial and head-final RCs.

Another reason is that both theories hypothesized that the processing of filler-gap dependencies relies heavily on computational resources such as working memory, while working memory impairment is pervasive in dyslexia (e.g., Chiappe et al., 2000; Smith-Spark & Fisk, 2007).

According to the DLT, sentence comprehension requires two computational resources: storage of the structure built thus far and integration of the current word into the existing structure. One key aspect of this account is that sentence complexity is related to the locality of integration between dependent syntactic elements (e.g., a dependent with a head). The locality is measured by the distance between these relevant elements, i.e., the number of new discourse referents (nouns and verbs) intervening between them. Within this framework, the longer an element has to be kept in working memory, the greater the computation resources required. The DLT predicts that, in languages like English, ORCs should be more difficult to process than SRCs because the relation between the relative head (i.e., the noun modified by the clause) and its trace in ORCs is resolved at a later stage. In particular, in (1a), the integration between the relative head *the child* and its trace is local. By contrast, in (1b), the integration between the relative head and its trace has to cross the embedded subject *the teacher* and the embedded verb *draws*, and is thus hypothesized to consume more computation resources. However, in languages like Chinese, SRCs should be harder to process than ORCs because the distance between the relative head and its trace in SRCs (2a) is longer than that in ORCs (2b), and by hypothesis requires more computation resources.

Under the fRM framework, when two elements that enter a local relation (i.e., the moved NP and the position where it is first merged) are hierarchically separated by an intervening element (i.e., another NP) matching the featural specification of the elements it separates, an intervention effect arises and the sentence is more complex to comprehend. In the case of RCs, the difficulty is modulated by the nature of the relative head and of the embedded noun. Under a raising analysis, the relative head is attracted by a complex attractor [+R, +NP], where R and NP represent the relative feature and the lexical restriction feature, respectively. In an English SRC (1a), the relative head, which bears [+R, +NP] features, does not cross over the lexically restricted element in the clause, i.e., the embedded object, which bears a [+NP] feature, a subset of the features of the relative head. In an English ORC (1b), the lexically restricted relative head has to cross over another lexically restricted element, i.e., the embedded subject, which also bears a [+NP] feature. To consider the relative head and the embedded noun distinct as required by RM, one has to compute the superset−subset relation, but by hypothesis, limited computational resources sometimes prevent younger children from making this computation. Accordingly, an RM violation arises and children have greater trouble in acquiring ORCs compared with SRCs (Belletti et al., 2012; Friedmann et al., 2009, 2021). The fRM approach also makes precise predictions for languages like Chinese. As shown in (3a), in a Chinese SRC, the relative head (*haizi* "child") structurally dominates (i.e., c-commands) its trace (marked as *t*, representing the head's original position in the clause). Critically, there is no intervening element between the relative head and its trace, namely, the embedded object (*laoshi* "teacher") does not disrupt this relationship. In contrast, in a Chinese ORC, as illustrated in (3b), the embedded subject (*laoshi* "teacher") structurally intervenes between the relative head (*haizi*

"child") and its trace. Here, the relative head c-commands the embedded subject, which in turn, c-commands the trace, creating a structural conflict. To reiterate, under the raising analysis, the embedded noun has a [+NP] feature, which shares a subset of the features of the relative head. According to the fRM approach sketched above, children encounter difficulty computing this subset relation and, thus, Chinese ORCs are also expected to be more difficult to acquire than SRCs, similar to English.

(3) a. Hierarchical structure of Chinese SRC
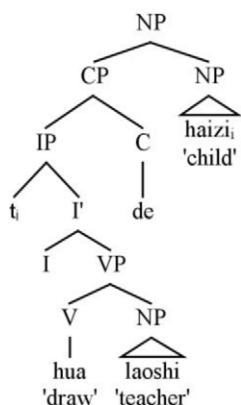
b. Hierarchical structure of Chinese ORC
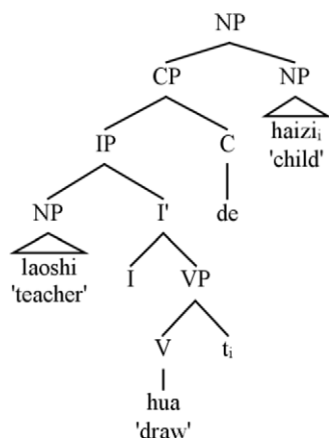
Critically, intervening elements can disrupt dependency formation both hierarchically and linearly. Contrary to structural intervention, defined by c-command, linear intervention is based on precedence. In head-initial RCs, such as those in English, structural and linear factors are intertwined and cannot be separated, whereas in Chinese RCs, the two factors can be disentangled (Hu et al., 2016). In Chinese SRCs, the embedded object (*laoshi* "teacher") linearly intervenes between the relative head (*haizi* "child") and its trace, but does not structurally

intervene (see 2a and 3a). By contrast, in Chinese ORCs, the embedded subject (*laoshi* "teacher") does not linearly intervene between the relative head (*haizi* "child") and its trace, but does structurally intervene (see 2b and 3b). Through comprehension experiments with TD children, Hu et al. (2016) demonstrated that both hierarchical and linear interventions create interference, but to varying degrees: structural intervention has a stronger effect than linear intervention.

To summarize, the SRC advantage is uniformly reported in the comprehension of head-initial RCs, as those in English, which can be explained by the DLT and the fRM. Conversely, this SRC advantage is not as clear-cut in the comprehension of Chinese RCs, and the two theories diverge in predicting the asymmetry between two structures: an ORC advantage predicted by the DLT, and an SRC advantage predicted by the fRM. Therefore, the first aim of our study is to use a clinical population to distinguish between these two approaches and thus to characterize the syntactic difficulties in the dyslexic population. In the next section, we turn to previous studies on the comprehension of RCs in dyslexia.

Relative clause comprehension in children with developmental dyslexia

Developmental dyslexia (DD) is a neurobiological condition affecting around 3−10% of the school population across languages (e.g., Stevenson et al., 1982; Zhang et al., 1996). Children with DD often show problems in reading fluently and comprehending written materials accurately, despite their normal intelligence and adequate learning opportunities. Besides, they also face difficulties in language domains such as syntax (e.g., Hu et al., 2018b, 2024; Shankweiler et al., 1984, 1995). Several studies on dyslexic children have reported deficits in the comprehension of RCs, with some differences depending on the manipulation of animate and inanimate NPs in the structure, the varying ages of the children, and the different experimental tasks used (Arosio et al., 2017; Bar-Shalom et al., 1993; Byrne, 1981; Casalis et al., 2013; Leikin & Assayag-Bouskila, 2004; Mann et al., 1984; Robertson & Joanisse, 2010; Stein et al., 1984).

With respect to head-initial RCs, a number of studies showed that the comprehension of SRCs is not impaired for children with DD, while the comprehension of ORCs is problematic for them. Using an act-out task, Stein et al. (1984) examined 7- to 10-year-old English-speaking reading-disabled children ($N = 20$) and typical readers ($N = 20$) with sentences containing two animate NPs and one inanimate NP. Results showed that reading-disabled children exhibited difficulties in comprehending ORCs (e.g., *The bear bites the lion that the ball hits*) compared to SRCs (e.g. *The lion hugs the bear that rolls the ball*). Using a picture-selection task, Arosio et al. (2017) examined the comprehension of SRCs and ORCs with animate NPs in Italian children with DD ($N = 13$, age range: 8; 7–13; 3) and TD controls matched for age or vocabulary. Results showed that SRCs were unproblematic for all the children, while ORCs were challenging for children with DD and vocabulary-matched TD controls ($N = 13$, age range: 7; 8–11; 0). Accordingly, the authors claimed that this finding supports intervention effects within the fRM framework.

Turning to head-final RCs, children with DD also have difficulties comprehending RCs, with only a couple of studies examining Chinese RCs (Chan 2014, 2015).

Using a listening and reading comprehension test, Chan (2014) found Chinese dyslexic children ($N = 43$, age range: 8; 4–12; 5) in Hong Kong comprehended SRCs and ORCs less accurately than TD children ($N = 43$, age range: 8; 4–12; 5). In addition, dyslexic children comprehended SRCs less accurately than ORCs (56% vs 75%), while TD children comprehended both structures similarly (83% vs 89%). The similar results were also reported in Chan (2015), which examined the reading comprehension of RCs using a sentence reading comprehension test. However, the ORC advantage observed in the dyslexic group must be cautious. In those studies, as illustrated in (4), the head noun (*na zhi xiao fei zhu* "that small fat pig") is a complex noun phrase, including a demonstrative (*na* "that"), a general classifier (*zhi*), two adjectives (*xiao* "small" and *fei* "fat"), and a noun (*zhu* "pig"), while the embedded noun is a noun with an adjective (*xiao tu* "small rabbit"). To fully address the theoretical issue related to the subject-object asymmetry of RCs, one should carefully match the head noun and the embedded noun. In addition, these studies did not explore how children do when they fail to understand RCs. It is essential to analyze errors, as it would provide an important window into their underlying deficit. Given these limitations, the current study would further examine Chinese children with DD by carefully manipulating the syntactic features of SRCs and ORCs and analyzing both correct and incorrect responses.

(4) Zhuizhe xiao   tu   de   na   zhi   xiao   fei   zhu   shi   huisede.
     chase   small   rabbit REL   DEM   CL   small fat   pig   is   grey
     "That small fat pig that is chasing the small rabbit is grey."

Dyslexic children's impairments in the comprehension of complex syntax may be associated with several sources. One could be a consequence of poor vocabulary knowledge. Vocabulary has been suggested to be a pivotal measure in evaluating individual differences in linguistic performance (e.g., Joshi, 2005; Tunmer & Chapman, 2012). Studies with children and adults showed that increased vocabulary knowledge affects both processes and representations shared with spoken language (Borovsky et al., 2012; Mani & Huettig, 2012; Nation et al., 2003). Crucially, vocabulary knowledge robustly accounted for the unique variance of prediction of spoken language beyond production fluency and nonverbal IQ (Hintz et al., 2017; Rommers et al., 2015). Van Dyke et al. (2014) administered a comprehensive skill battery and found that receptive vocabulary knowledge was the only significant predictor of comprehension performance when the variance shared with IQ was removed. They interpreted these results in the light of a model that emphasizes retrieval interference and the quality of lexical representations as key determinants of successful comprehension.

Working memory has been suggested to be a critical cognitive factor contributing to comprehension deficits in dyslexia (Mann et al., 1984; Shankweiler & Crain, 1986). Studies have shown that working memory deficits are pervasive in individuals with dyslexia (Chiappe et al., 2000; Smith-Spark & Fisk, 2007). In a study that directly examined the relationship between working memory and spoken sentence comprehension in dyslexic children, Robertson and Joanisse (2010) manipulated working memory loads by varying sentence length and the delay between the offset of the sentence and the presentation of picture stimuli across

different sentence types. Results showed that, compared to canonical sentences (i.e., active sentence and SRCs), English-speaking children with DD ($N = 14$, age range: 9; 1–12; 1) performed more poorly on noncanonical sentences (i.e., passive sentences such as *The boy in the dark blue pants is tapped by the girl with the nice blond hair* and passive RCs such as *This is the man in the light brown shirt that is waved at by the boy in the dark blue pants*) under a high working memory load. Research on dyslexic adults' sentence comprehension has also reported similar findings (e.g., Wiseheart et al., 2009): College students with dyslexia were significantly outperformed by their TD peers in comprehending complex sentences containing center-embedded RCs with high working memory loads.

To sum up, several critical factors have been considered to account for comprehension deficits in dyslexia, but none of the Chinese dyslexic studies have examined whether those factors could be a potential predictor for the RC comprehension. The current study would further investigate Chinese dyslexic children's comprehension of RCs and measure their vocabulary knowledge and verbal working memory, with the aim of establishing the source of processing difficulty in RCs.

The current study

The present study investigated the comprehension of subject and object RCs in Chinese children with and without DD. The first goal was to contribute to theoretical debates concerning the source of processing difficulty associated with subject and object RCs. The second goal was to determine whether Chinese children with DD have difficulty comprehending RCs, and whether vocabulary knowledge and working memory contribute to their RC comprehension.

To achieve these goals, we used an RC comprehension task, measuring both accuracy and response latency, to test Chinese children with DD and their chronological age-matched (CA) peers and reading-level-matched (RL) peers. The purpose of including two control groups was to help determine whether any observed syntactic processing difficulties were based on a delay that could potentially be explained by limited reading and vocabulary experience. In addition, we administered additional tasks to determine how children's vocabulary knowledge and verbal working memory relate to their RC comprehension. To characterize working memory capacity, we used forward and backward digit span tasks. The forward digit span is a measure of phonological short-term memory, as the task requires participants to maintain the correct order of an increasing sequence of digits and to repeat it. The backward digit span is a measure of central executive functioning, as the task requires participants to manipulate the retained information by calculating the reversed order of the digits, which implies retaining, manipulating, and recalling a given number sequence. By assessing vocabulary knowledge and working memory, we were able to identify the underlying factors responsible for RC processing difficulties.

For the first research question, we focused on analyzing whether there was an asymmetry between subject and object RCs across all the groups, and which errors children made in comprehending RCs. Regarding the theoretical psycholinguistic debate, the DLT predicts an ORC advantage, and the fRM predicts an SRC

advantage. We therefore anticipated significant differences between SRCs and ORCs either in accuracy or response latency. Additionally, due to linear intervention effects, we expected children not only to have difficulties in ORC comprehension but also to exhibit some difficulties in SRC comprehension. For the second research question, we focused on analyzing whether children with DD differed significantly from TD controls in RC comprehension, and whether vocabulary knowledge and verbal working memory influenced RC comprehension. We hypothesized that children with DD would show poorer comprehension accuracy and longer response latencies compared to the CA controls, and perform similarly to the RL controls. We also expected vocabulary knowledge and verbal working memory scores to correlate significantly with RC comprehension, particularly in children with DD.

Methods

Participants

Sixty-six Chinese children were recruited for three groups, namely, children with DD, the CA controls, and the younger RL controls. All the participants were right-handed (determined by writing habits) and had normal or corrected-to-normal vision, with no reported ophthalmologic or neurological abnormalities. All the information was verified through school records and teacher reports.

Dyslexic children were screened following a well-established procedure in the literature (e.g., Wang et al., 2014; Zhao et al., 2019). A total of 1,105 children from Grades 2 to 6 at a public primary school in Beijing were administered two screening tests: the Standardized Chinese Character Recognition Test (Wang & Tao, 1996) and the Chinese version of the Raven's Standard Progressive Matrices test (Zhang & Wang, 1985). Children were identified as being at risk of dyslexia if their scores on the character recognition test were at least −1.5 standard deviations below the grade-level mean, and if their nonverbal intelligence fell within the normal range (i.e., above the 5th percentile of the norm). In addition, the children's Chinese teachers were invited to evaluate whether the screened individuals demonstrated difficulties in their daily Chinese learning and to complete the revised Chinese version of the Swanson, Nolan, and Pelham Scale-IV (SNAP-IV; Zhou et al., 2013), which was used to assess attention-deficit/hyperactivity disorder based on the criteria from the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). To be included in the DD group, children were required to meet fewer than three items on both the inattention and hyperactivity-impulsivity subscales of the SNAP-IV. Based on these criteria, 60 children (5.43%) were identified as at risk for dyslexia among the 1,105 students screened. From this group, 22 children with DD were recruited to participate in the current study. CA and RL controls were selected from their peers.

To ensure the validity of the dyslexia diagnosis screening, all participants were administered multiple linguistic skill assessments. These included phonological awareness, assessed using an onset, rime, and lexical tone detection task; morphological awareness, assessed using homophonetic and homographic awareness tests; and rapid automatized naming, assessed using a rapid digit naming task. All tests have been used in prior research (Huang et al., 2021; Zhao et al., 2017). Table 1 presents the means and standard deviations of the measured

Table 1. Participant characteristics

	DD	CA	RL	*F* value
	M (SD)	*M* (SD)	*M* (SD)	F (2, 63)
No.	22	22	22	
Gender	17 males	14 males	13 males	
Age range	10;4–13;3	10;7–12;9	8;10–10;7	
Age	11;6 (0.81)	11;6 (0.62)	9;6 (0.59)	61.05*** DD = CA > RL
Character recognition[a]	2108.09 (251.40)	3187.99 (115.35)	1935.52 (401.05)	128.10*** DD = RL < CA
Raven	42.95 (4.93)	44.41 (3.98)	39.68 (3.73)	7.15** DD = CA > RL
Phonological awareness	15.54 (5.37)	20.55 (5.31)	17.82 (5.88)	4.52* DD < CA, DD = RL
Homophonetic awareness[b]	13.32 (2.44)	17.00 (1.88)	15.14 (2.05)	16.35*** DD < RL< CA
Homographic awareness[b]	12.45 (1.99)	14.86 (1.42)	13.50 (2.22)	8.81*** DD = RL < CA
Rapid digit naming	12.65 (3.00)	10.62 (2.60)	11.03 (2.59)	3.38* DD < CA, DD = RL

[a]Standardized Chinese Character Recognition Test.
[b]Morphological awareness was assessed with homophonetic and homographic awareness tests; *p < .05; **p < .01; ***p < .001

variables for the three groups, along with the ANOVA *F*-values for group differences on these measures and pairwise comparisons derived from Tukey's post hoc tests. The results revealed that children with DD, matched for age and Raven scores with their CA peers, had significantly lower scores in character recognition, phonological awareness, morphological awareness, and rapid digit naming compared to their CA peers. In addition, compared to their RL peers, they had older age, higher Raven scores, and lower homophonic awareness scores, but were matched for character recognition, phonological awareness and homographic awareness, and rapid digit naming.

Materials

The study consisted of an RC comprehension, vocabulary knowledge, and a working memory test.

First, children's RC comprehension was tested with a character-sentence matching task, developed by Hu et al. (2020). Chinese SRCs and ORCs are exemplified in (5a) and (5b), respectively.

(5) a. Zhichu   zhui   shizi   de   ma.
           point to chase lion   REL   horse
           "Point to the horse that is chasing the lions."

Figure 1. An example of experimental pictures in the RC comprehension task.

b. Zhichu   shizi   zhui   de   ma.
    point to   lion   chase   REL   horse
    "Point to the horse that the lions are chasing."

The task consisted of 16 black and white pictures with the same structure (i.e., one animal X on the left, a pair of animals Y in the middle, and another X on the right). Figure 1 is a sample of experimental pictures. The pictures depicted 8 actions, including *bite, chase, follow, hit, push, smell, spurt,* and *wipe*. They were presented with equal times, i.e., each action appears in the two types of RCs exemplified in (5). To avoid priming effects, two lists of the task (i.e., list A and list B) were created, and each picture was used only once in each list. Each list included 8 SRCs and 8 ORCs. In addition, there were 8 filler sentences involving intransitive verbs (e.g. *sleep*) or actional irreversible verbs (e.g. *drink*), and 3 practical items. In total, there were 27 items in each version of the RC test. All the sentences were recorded by a native female Chinese speaker. The pictures and the sentences were presented at the same time. The test was programed, using E-Prime 2.0 software, recording accuracy and response latency (Schneider et al., 2012).

Moreover, to assess working memory abilities, children were tested with a forward digit span (FDS) and a backward digit span (BDS) task, taken from the Wechsler Intelligence Scale for Children Revised in China (C-WISE; Gong & Cai, 1994). The FDS includes a series of digits of increasing length, from 3 to 13 digits, and BDS includes a series of digits of increasing length, from 2 to 12 digits. The lists were digitally recorded in audio files by a Chinese native speaker. Each list was read at a rate of one digit per second. Children were asked to listen carefully to the series of digits and immediately repeat them aloud, either in the same order (yielding the FDS) or in the reversed order (yielding the BDS). The task was stopped when children failed in repeating 2 out of 2 trials within one level.

Furthermore, to assess receptive vocabulary knowledge, children were tested with the Peabody Picture Vocabulary Test-Revised Chinese Version (PPVT-R; Sang & Miao, 1990). The test consists of 175 items. During the test, children were presented each word with four pictures, and were asked to select the picture that best corresponded to the word. The total score was calculated by subtracting the number of errors from the maximum achievable score.

Procedure

Participants were tested individually in a quiet room at the primary school. Written informed consent was obtained for the children from their parents and teachers

before assessment took place. The study was approved by the Research Ethics Committee of the School of Psychology at Capital Normal University, in accordance with the standards of the Helsinki Declaration (1964).

Participants took approximately 3 hours to complete all the tests, which were administered over two or three sessions within one week. The tests included those reported in Participants section, as well as the RC comprehension, the PPVT-R, the FDS, and the BDS test. Each task was explained to the children in detail. The experimental materials of the RC tests were presented on a laptop. Half of the children completed list A, and the other half completed list B. Children were asked to listen to a series of sentences while looking at corresponding pictures. They were instructed to choose the character referred to in the sentence by pressing specific keys on the keyboard: "V" for the left character, "B" for the middle character, and "N" for the right character. These keys corresponded to the positions of the characters from left to right in the picture. They were encouraged to respond as quickly and accurately as possible. Before the main experiment began, three practice trials were conducted to familiarize the participants with the task. No feedback was provided during the experimental trials. Test administrators were instructed to give children breaks if any signs of fatigue were observed.

Scoring and error coding

In the RC comprehension test, the dependent variable was the proportion of accurate responses, namely, the accuracy in identifying the correct character. When participants did not choose the correct character(s), we coded the response as Error. Errors were labeled as Reversal Error and Embedded Error, following Hu et al. (2020).

Consider "*the horse that the lions are chasing*" and Figure 1. A Reversal Error was coded if the horse on the right was chosen, i.e., the horse that is chasing the lions. In this case, the theta-roles are reversed, i.e., the relative head *the horse* is a Patient, but it was interpreted as an Agent. An Embedded Error was coded when children selected the middle characters, i.e., *the lions*, which correspond to the embedded noun within the RC. The Reversal Error may reflect a misunderstanding of the thematic assignment, whereas the Embedded Error may arise from children's confusion about the syntactic role of the relative head, indicating an insensitivity to the fact that the RC adds information to the relative head (Arnon, 2005; Hu et al., 2016).

Statistical analysis

Response accuracy and response latency were analyzed by employing mixed effects models (Baayen et al., 2008), based on the lme4 and lmerTest packages in the R environment (R Development Core Team, 2023). Models were constructed with a maximal random effects structure and were successively simplified when they failed to converge (Barr et al., 2013). The BOBYQA optimizer was used when models with the default optimizer did not converge. Filler sentences were performed at ceiling, with 98% correct responses, and thus, were excluded from the analysis.

The categorical accuracy data were analyzed via generalized linear mixed effects models, and the response latency data were analyzed with linear mixed effects models. The analysis of response latency was performed only on correct responses, with response latency being logarithmically transformed. Response latencies shorter than 200 ms or exceeding 2 SDs above the group mean were classified as outliers, resulting in 0.88% of the data being excluded from the analyses. To understand whether the asymmetry existed between subject and object RCs across all groups and whether children with DD differed significantly from TD controls in RC comprehension, Sentence Type (SRC vs ORC) and Group (DD vs CA vs RL) were included as potentially significant fixed factors in the analyses, with subjects and items as random factors. We used the ORC as the reference category for the Sentence Type factor, and the DD group for the Group factor. To investigate the effects of vocabulary knowledge and working memory, additional analyses were conducted using the PPVT-R, FDS, and BDS as continuous factors. Raw scores for the PPVT-R were selected for the analyses because the norms for the PPVT-R were established in 1990 (Sang & Miao, 1990) and are now outdated. Raw scores from the working memory tasks were also used to align with methodological precedents from prior studies (e.g., Bentea et al., 2016). Effects were evaluated one by one on the basis of likelihood ratio tests; both first-level effects and the interactions between the fixed factors were tested.

In addition, we explored dyslexic children's individual performance in RC comprehension against the CA group's mean score in terms of accuracy and response latency. Furthermore, we analyzed children's errors using Poisson regression models.

Results

We first report children's comprehension of RCs and then present their relationship with vocabulary knowledge and working memory.

RC comprehension

We report the results of RC comprehension, with the order of the analyses of correct responses, individual performance, and error analyses.

Correct responses

Table 2 reports means and standard deviations of the accuracy and response latency of correct responses in children with DD, their CA and RL controls. The comprehension of RCs in children with DD differed from that of the CA controls and resembled that of the RL controls. Across all groups, SRCs were comprehended more accurately than ORCs, whereas no significant difference in response latency was observed between the two structures.

Accuracy. Sentence Type and Group were initially entered into a factorial model, and significantly contributed to the model fit [$\chi^2(3) = 16.95, p < .001$]. The interaction between Sentence Type and Group did not significantly contribute to the goodness of fit of the model, as shown by the likelihood ratio test [$\chi^2(2) = 0.40, p = .82$]; therefore, it was removed. The best-fitting model included Sentence Type and Group as fixed factors. Table 3 shows the output of the analysis. Overall, the results indicate that SRCs

Table 2. Mean (and standard deviation) of accuracy and response latency (ms) for each group in RC comprehension

	DD	CA	RL
Accuracy			
SRCs	0.87 (0.18)	0.96 (0.06)	0.91 (0.14)
ORCs	0.74 (0.29)	0.88 (0.19)	0.82 (0.27)
Response latency			
SRCs	5257.32 (844.36)	4691.20 (605.89)	5497.49 (692.34)
ORCs	5260.80 (906.55)	4786.55 (586.63)	5616.06 (799.47)

Table 3. Fixed effects in the mixed-effects model for accuracy in the RC comprehension

	β	SE	Wald Z	p
(Intercept)	1.42	0.37	3.85	= .000***
Sentence Type	1.13	0.28	4.09	= .000***
Group (DD vs. CA)	1.11	0.50	2.22	= .03*
Group (DD vs. RL)	0.59	0.48	1.21	= .23

Reference level for Sentence Type = ORCs; reference level for Group = DD; *p < .05; ***p < .001.

were comprehended more accurately than ORCs; children with DD comprehended RCs significantly less accurately than their CA peers, and displayed a similar pattern with their RL peers.

Response latency. Sentence Type and Group were initially entered into a factorial model, and significantly contributed to the model fit [$\chi^2(3) = 15.35$, $p < .01$]. Again, the interaction between Sentence Type and Group did not significantly contribute to the goodness of fit of the model, as shown by the likelihood ratio test [$\chi^2(2) = 1.12$, $p = .57$]; therefore, it was removed. The best-fitting model included Sentence Type and Group as fixed factors. Table 4 shows the output of the analysis. Overall, the main effect of Sentence Type was not significant, revealing a similarity between subject and object RCs; children with DD comprehended RCs significantly more slowly than their CA peers, and performed similarly to their RL peers.

To sum up, there are two main findings. First, a clear SRC advantage over ORC in comprehending Chinese RCs was observed in accuracy across all the groups, but no significant difference between two structures was found in response latency. Second, children with DD demonstrated lower accuracy and longer response latencies than their CA peers, while they performed similarly to their RL peers.

Individual performance
We ran individual analyses to better understand the individual severity of the deficit. Following Arosio et al. (2017), we examined individual mean scores in SRC and ORC comprehension and their deviations from the mean scores of the CA group.

Table 4. Fixed effects in the mixed-effects model for response latency in the RC comprehension

	β	SE	t	p
(Intercept)	3.71	0.01	259.61	= .000***
Sentence Type	−0.01	0.01	−0.77	= .46
Group (DD vs. CA)	−0.04	0.02	−2.25	= .03*
Group (DD vs. RL)	0.03	0.02	1.70	= .09

Reference level for Sentence Type = ORCs; reference level for Group = DD; *$p < .05$; ***$p < .001$.

Accuracy. In comprehending SRCs, 12 out of 22 children with DD (54.55%) scored below the mean score of the CA group. Among these, 2 children (9.09%) scored more than 1 SD below the CA mean, and 2 children (9.09%) scored more than 2 SD below the CA mean. In comprehending ORCs, 13 out of 22 children with DD (59.09%) scored below the mean score of the CA group. Among these, 3 children (13.64%) scored more than 1 SD below the CA mean, and 2 children (9.09%) scored more than 2 SD below the CA mean. Notably, 3 of these children (13.64%) who scored more than 1 or 2 SD below the CA mean in ORC comprehension also scored more than 1 or 2 SD below the CA mean in SRC comprehension.

Response latency. For SRCs, 17 out of 22 children with DD (77.27%) responded more slowly than the mean response latency of the CA group. Among these, 3 children (13.64%) exceeded the CA group's mean by 1 SD, and of these, 4 children (18.18%) exceeded it by 2 SD. For ORCs, 18 out of 22 children with DD (81.82%) showed slower response latencies than the CA group's mean. Among these, 2 children (9.09%) exceeded the mean by 1 SD, and 2 children (9.09%) exceeded it by 2 SD. Notably, the same 4 children (18.18%) who exceeded 1 or 2 SD in ORC comprehension also exhibited the slowest response latencies in SRC comprehension.

To summarize, individual analyses demonstrate that a significant proportion of children with DD exhibit deficits in comprehending SRCs, ORCs, or both structures.

Error analyses

We examined the distribution of errors to better understand what children did when they failed to understand RCs. Figure 2 shows the means and standard deviation of Reversal and Embedded Error for each group in the comprehension of SRCs and ORCs.

We counted the number of each type of error made by each child and treated it as a count variable to run a Poisson regression model. In SRC comprehension, there was no significant difference between Reversal and Embedded Errors, a pattern observed across all groups (all $ps > .53$). In ORC comprehension, Embedded Errors occurred significantly more frequently than Reversal Errors ($\beta = -1.39$, SE = 0.25, Wald $Z = -5.55$, $p < .001$). This difference was consistent across the DD group ($\beta = -1.56$, SE = 0.39, Wald $Z = -4.01$, $p < .001$), the CA group ($\beta = -0.98$, SE = 0.48, Wald $Z = -2.05$, $p < .05$) and the RL group ($\beta = -1.47$, SE = 0.45, Wald $Z = -3.24$, $p < .01$).
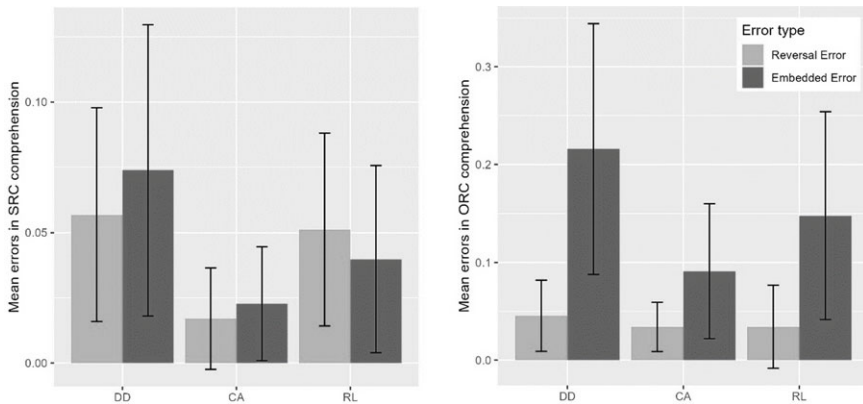
Figure 2. Means (and standard deviation) of errors in the SRC and ORC comprehension.

In summary, when the children were not able to comprehend SRCs, they made Reversal or Embedded Error. In other words, they chose another character randomly. When they did not comprehend ORCs, they were more likely to make Embedded Error than Reversal Error.

The relationship between RC comprehension and individual measures

We report children's performance on vocabulary knowledge and working memory tasks, followed by the analyses of their relationship with RC comprehension.

First, children with DD ($M = 141.77$, SD $= 13.14$) scored lower on the PPVT-R than their CA peers ($M = 149.55$, SD $= 10.15$), but scored quantitatively higher than their RL peers ($M = 137.55$, SD $= 18.35$). One-way ANOVA revealed significant differences between the DD and the CA group, $F(1, 42) = 4.82$, $p < .05$, $\eta^2 = 0.10$, as well as between the CA and the RL group, $F(1, 42) = 7.20$, $p < .05$, $\eta^2 = 0.15$, while there was no significant difference between the DD and the RL group, $F(1, 42) = 0.77$, $p = .39$, $\eta^2 = 0.02$.

Second, children with DD performed better on the FDS task than on the BDS task ($M = 6.82$, SD $= 1.33$; $M = 4.00$, SD $= 1.69$, respectively), and so did for the CA group ($M = 7.77$, SD $= 1.63$; $M = 4.77$, SD $= 1.45$, respectively) and the RL group ($M = 7.23$, SD $= 0.92$; $M = 4.05$, SD $= 1.43$, respectively). This pattern was as expected, consistent with previous studies on children and adults (Bentea et al., 2016; Wiseheart et al., 2009). Regarding the FDS, one-way ANOVA revealed a significant difference between the DD and the CA group, $F(1, 42) = 4.52$, $p < .05$, $\eta^2 = 0.10$, but no significant differences between the DD and the RL group, $F(1, 42) = 1.40$, $p = .24$, $\eta^2 = 0.03$, or between the CA and the RL group, $F(1, 42) = 1.86$, $p = .18$, $\eta^2 = 0.04$. Regarding the BDS, one-way ANOVA revealed no significant differences between the DD and the CA group, $F(1, 42) = 2.66$, $p = .11$, $\eta^2 = 0.06$, between the DD and the RL group, $F(1, 42) = 0.01$, $p = .92$, $\eta^2 = 0.01$, or between the CA and the RL group, $F(1, 42) = 2.82$, $p = .10$, $\eta^2 = 0.06$.

To investigate whether vocabulary knowledge and working memory contribute to the processing of complex syntactic structures such as RCs, we examined whether

scores on the PPVT-R, FDS, and BDS predicted accuracy and response latencies in the RC comprehension task.

Accuracy: We first carried out an analysis of collinearity between the PPVT-R, FDS, and BDS measures and RC accuracy to decide whether to include these measures in the mixed effects model. Interestingly, we found a significant correlation between PPVT-R and RC accuracy in the DD group ($r = 0.72$, $p < .001$), but not in the CA ($r = -0.34$, $p = .12$) and the RL group ($r = 0.30$, $p = .17$). Thus, we decided to run analyses without PPVT-R. We then tested whether the inclusion of the scores of FDS and BDS predicted accuracy in the comprehension task. We observed the FDS did not contribute to the fit of the model [$\chi^2(1) = 1.84$, $p = .18$], while the BDS contributed to the fit of the model [$\chi^2(1) = 6.09$, $p < .05$; $\beta = .35$, SE $= .14$, Wald $Z = 2.48$, $p < .05$], indicating that accurate responses were more likely in association with higher backward span.

We conducted an additional series of analyses to identify the differential roles exerted by working memory separately for each group in each condition. The FDS and the BDS did not significantly predict accuracy in each condition in the DD or the CA group (all $ps > .05$). In the RL group, the FDS did not contribute to the fit of the models (both $ps > .05$), while the BDS significantly predicted accuracy in both SRC and ORC conditions ($\beta = .87$, SE $= .42$, Wald $Z = 2.06$, $p < .05$; $\beta = 1.88$, SE $= .66$, Wald $Z = 2.86$, $p < .01$).

Response latency. As we did for accuracy measure, we further tested whether scores of vocabulary knowledge and working memory predicted response times to the comprehension task, considering all age groups collapsed. The PPVT-R predicted response latency to the RC comprehension task [$\chi^2(1) = 8.14$, $p < .01$; $\beta = -.01$, SE $= .01$, $t = -2.90$, $p < .01$], indicating that high scores in a receptive vocabulary task were associated with a faster comprehension of RCs. Neither the FDS [$\chi^2(1) = 1.17$, $p = .28$] nor the BDS [$\chi^2(1) = 0.07$, $p = .80$] contributed to the fit of the model.

We also conducted an additional series of analyses to identify the differential roles exerted by vocabulary knowledge and working memory separately for each group in each condition. In the DD group, the PPVT-R ($\beta = -.01$, SE $= .01$, $t = -4.47$, $p < .001$) and the FDS ($\beta = .02$, SE $= .01$, $t = 2.34$, $p < .05$) predicted the comprehension task in ORC condition, indicating that response latency was more likely in association with better vocabulary skills and higher forward span. In the CA and the RL group, none of the significant effects were observed (all $ps > .05$).

To sum up, there are two main findings. First, high scores in a receptive vocabulary task were associated with a more accurate and faster comprehension of RCs, and this was particularly evident in the DD group. Second, FDS was associated with a faster comprehension of ORCs, and this was only evident in DD group; BDS was associated with a more accurate comprehension of RCs, and this was particularly evident in the RL group.

Discussion

In the next sections, we discuss experimental findings to address our research questions and highlight some clinical implications and open questions in future research.

Subject advantage in the comprehension of relative clauses in dyslexia

The first goal of the present study was to contribute to theoretical debates concerning the source of processing difficulty associated with subject and object RCs.

According to the DLT, an ORC advantage should have been found in the comprehension of Chinese RCs. To recall, the comprehension difficulty is related to the locality of assembling two dependent syntactic heads: the earlier the dependency is resolved, the fewer computational resources are required. In the case of Chinese, the dependency in ORCs is resolved earlier than that in SRCs. Thus, ORCs require less computational resources than SRCs, and should be less difficult to process. On the contrary, according to the fRM, in Chinese SRCs, the embedded object does not structurally intervene between the relative head and its trace, whereas in Chinese ORCs, the embedded subject structurally intervenes between the relative head and its trace. Due to this structural intervention, children with DD and their CA and RL peers are expected to perform more poorly on ORCs compared to SRCs. Our data showed that all three groups comprehended ORCs less accurately than SRCs, revealing the SRC advantage in comprehension in line with the predictions based on the fRM, contrary to the DLT. This SRC advantage is consistent with a body of literature showing that children have difficulties comprehending ORCs (e.g., Belletti et al., 2012; Friedmann et al., 2009; Hu et al., 2016), confirming intervention effects in the acquisition of RCs.

It is worth pointing out that our results did not replicate the finding reported by Chan (2014, 2015). Recall that their Chinese children with DD comprehended SRCs significantly worse than ORCs, and TD children comprehended SRCs and ORCs quantitatively similarly. As pointed out earlier, their results may be due to the limitation of the experimental design, namely, the head noun has many more features than the embedded noun, which may facilitate the comprehension of ORCs. Our study manipulated the features of the head noun and the embedded noun carefully, as both are a noun that bears a [+NP] feature.

Our data also revealed that the pattern of errors was similar in children with DD and their TD peers. That is, when children were not able to comprehend SRCs, they made Reversal or Embedded Error; when they did not comprehend ORCs, they were more likely to make Embedded Error than Reversal Error. These results replicate what has been reported in previous studies on Mandarin-speaking TD children (Hu et al., 2016). As stated earlier, the Reversal Error seems to reflect a misunderstanding of the thematic assignment, whereas the Embedded Error may reflect children's confusion about the syntactic role of the relative head, indicating that children are not sensitive to the fact that the RC adds information to the relative head and they do not integrate the two sets of information, from the relative head and from the RC. As pointed out by previous studies (Hu et al., 2016), the finding on SRCs reveals that linear intervention is taxing for Chinese children, although to a lesser extent than structural intervention (operating in ORCs). In a SRC (*hua laoshi de haizi* "the child that draws the teacher"), the embedded object (*laoshi* 'teacher') linearly intervenes between the relative head (*haizi* 'child') and its gap. With respect to the findings on Chinese ORCs, an Embedded Error amounts to choosing the first NP heard as the Agent of the action, confirming that children were also influenced by the linear order when processing RCs.

To sum up, Chinese children with and without DD showed the SRC advantage in comprehension as predicted by the fRM, contrary to the DLT. In addition, the results of correct responses and errors confirmed that both hierarchical and linear interventions create interference, but to different degrees, namely, structural intervention has stronger effects than linear intervention.

Factors contributing to impaired comprehension of relative clauses in dyslexia

The second goal was to determine whether Chinese children with DD have difficulty comprehending subject and object RCs, compared to their TD peers, and whether vocabulary knowledge and working memory contribute to RC comprehension in dyslexia.

First, our data revealed that a subset of children with DD had deficits in the comprehension of SRCs. Significant difference was observed between the DD group and CA controls, but not between the DD group and RL controls. Specifically, 54.55% of children with DD scored below the CA group's mean accuracy for SRC, and 77.27% exhibited slower response latencies than the CA group's mean. Critically, 9.09% of children with DD scored more than 2 SD below the CA mean in accuracy, and 18.18% exceeded the CA group's mean by 2 SD in response latency. This contrasts with findings from languages with head-initial RCs (e.g., Italian), where children with DD achieve ceiling performance in SRC comprehension (Arosio et al., 2017). As theorized, these difficulties in Chinese SRCs can be attributed to linear intervention, where the embedded object linearly intervenes between the relative head and its trace.

Deficits were also observed in ORC comprehension. Similar to the SRC comprehension, children with DD performed significantly worse than the CA controls, and did not differ from the RL controls. In total, 59.09% of children with DD scored below the CA group's mean accuracy for ORC, and 81.82% exhibited slower response latencies than the CA group's mean. Critically, 9.09% scored more than 2 SD below the CA group's mean accuracy, and 9.09% exceeded the CA group's mean by 2 SD in response latency. These results are in line with cross-linguistic studies of head-initial RCs (e.g., Arosio et al., 2017). This suggests that a subset of children with DD have problems moving a relative head over an embedded subject endowed with a feature which is a subset of the relative head.

A significant proportion of children with DD showed comorbid deficits in both SRC and ORC comprehension. In terms of accuracy, 13.64% of children with DD who scored more than 1 or 2 SD below the CA group's mean accuracy in ORC comprehension also scored more than 1 or 2 SD below the CA mean in SRC comprehension. For response latency, 18.18% of children who exceeded 1 or 2 SD in ORC comprehension also exhibited slower response latencies in SRC comprehension. These results underscore that both linear and structural interventions contribute to impaired RC comprehension in Chinese children with DD.

Second, results showed that high scores in a receptive vocabulary task were associated with a more accurate and faster comprehension of RCs, and this was particularly evident in the DD group. This result confirmed that vocabulary was the best predictor of sentence comprehension, in line with previous studies (e.g., Joshi, 2005; Tunmer & Chapman, 2012). Importantly, the dyslexic group did not show a

significant difference in vocabulary knowledge compared to the RL group, and performed similarly to the RL group in both accuracy and response latency. Taken together, these results raised the possibility that the performance of sentence comprehension might be influenced by vocabulary knowledge. Research showed that literacy enhances vocabulary and syntactic knowledge, and increased vocabulary knowledge enables increased prediction of spoken language (Huettig & Pickering, 2019). The vital role of vocabulary seems to signify that it is a fundamental element in an architectural account of comprehension difficulty, according to which the memory retrieval mechanism may play a primary role (Van Dyke et al., 2014).

Note that our results also showed that FDS was associated with a faster comprehension of ORCs, and this was only evident in the DD group; BDS was associated with a more accurate comprehension of RCs, and this was particularly evident in the RL group. As introduced earlier, the FDS is a measure of phonological short-term memory, and the BDS is a measure of central executive functioning. Our results revealed a close relationship between phonological short-term memory and sentence processing in DD, while a close relationship between central executive functioning and sentence comprehension in RL controls. The contrast between the DD and the RL children seems to implicate different computational resources involved in comprehending complex structure such as RCs. For this reason, it is likely that the subtle sentence processing deficits found in dyslexia in the current study can also be explained by the processing deficit theory (Mann et al., 1984; Shankweiler et al., 1984, 1995; Smith et al., 1989). According to this hypothesis, spoken sentence comprehension deficits are grounded in a phonological deficit. In this view, it is not reading that is directly related to sentence processing, but rather phonological processing. Accordingly, a core deficit in phonology makes spoken language processing difficult, especially when working memory loads are high in tasks that are completely auditory.

Together, our results have further implications. First, our data revealed that a subset of children with DD had a consistent language deficit, as they scored drastically lower than 2 SD below the CA mean. This is particularly important, as norm-referenced language tests for children at these ages are not available in China, and it is not common to evaluate these children's grammatical abilities. Since there is a high comorbidity between reading disorder and language impairment (Snowling & Hulme, 2012), and children with DD could only be diagnosed during the school years, our results can help establish whether these children are also affected by language impairment. At a clinical level, the current study suggests that an evaluation of SRC and ORC comprehension should be included in testing materials for identifying language impairment in dyslexia. Second, our results revealed that vocabulary knowledge and phonological short-term memory play a vital role in dyslexic children's sentence comprehension. This indicates that clinical efforts focused on building children's vocabulary and working memory skills may improve their spoken language abilities. However, one must be cautious to consider these results' interpretation, as the study relied solely on a receptive vocabulary measure and a digit span task. Future research could be improved by adding measures of expressive vocabulary (Wiseheart & Altmann, 2018) and episodic buffer—a component of working memory responsible for integrating information from a

range of sources into a multidimensional code (Baddeley, 2000). Furthermore, given the influence of reading experience on vocabulary development (Cain & Oakhill, 2011), subsequent studies could investigate the specific effect of reading experience on the comprehension of complex syntactic structures.

Conclusion

This study investigated the comprehension of SRCs and ORCs in Chinese children with DD, comparing with TD children matched for age or reading level. All the groups showed the SRC advantage in accuracy, a finding that can be explained in terms of structural intervention within the fRM framework, contrary to the DLT. The DD group comprehended SRCs and ORCs less accurately and more slowly than the CA group, and performed similarly to the RL group. Also, a significant number of children with DD exhibit deficits in comprehending SRCs, ORCs, or both structures. These findings confirmed the existence of syntactic difficulties in dyslexia. Dyslexic children's receptive vocabulary knowledge was associated with higher accuracy and shorter response latencies in RC comprehension, and their phonological short-term memory was specifically linked to faster RC processing. These findings are most compatible with the claim that syntactic difficulties may stem from limited vocabulary knowledge and phonological short-term memory deficits.

References

Arnon, I. (2005). Relative clause acquisition in Hebrew: towards a processing-oriented account. In A. Brugos, M. R. Clark-Cotton, & S. Ha (Eds.), *BUCLD 29: Proceedings of the 29th Annual Boston University Conference on Language Development* (pp. 37–48). Somerville, MA: Cascadilla Press.

Arosio, F., Panzeri, F., Molteni, B., Magazù, S., & Guasti, M. T. (2017). The comprehension of Italian relative clauses in poor readers and in children with specific language impairment. *Glossa: A Journal of General Linguistics*, **2**(1), 9.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, **59**(4), 390–412.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, **4**(11), 417–423.

Bar-Shalom, E. G., Crain, S., & Shankweiler, D. (1993). A comparison of comprehension and production abilities of good and poor readers. *Applied Psycholinguistics*, **14**(2), 197–227.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language*, **68**(3), 255–278.

Belletti, A., Friedmann, N., Brunato, D., & Rizzi, L. (2012). Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew and Italian. *Lingua*, **122**(10), 1053–1069.

Bentea, A., Durrleman, S., & Rizzi, L. (2016). Refining intervention: the acquisition of featural relations in object A-bar dependencies. *Lingua*, **169**, 21–41.

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, **112**(4), 417–436.

Byrne, B. (1981). Deficient syntactic control in poor readers: is a weak phonetic memory code responsible? *Applied Psycholinguistics*, **2**(3), 201–212.

Cain, K., & Oakhill, J. (2011). Matthew effects in young readers: reading comprehension and reading experience aid vocabulary development. *Journal of Learning Disabilities*, **44**(5), 431–443.

Casalis, S., Leuwers, C., & Hilton, H. (2013). Syntactic comprehension in reading and listening: a study with French children with dyslexia. *Journal of Learning Disabilities*, **46**(3), 210–219.

Chan, C. Y. H. (2014). The role of working memory in the comprehension of relative clauses by Chinese dyslexic children. *Asian Journal of Education and e-Learning*, **2**(3), 165–172.

Chan, C. Y. H. (2015). Syntactic deficit and sentence comprehension of Chinese dyslexic children. *International Journal of Linguistics and Communication*, **3**(1), 32–37.

Chiappe, P., Hasher, L., & Siegel, L. S. (2000). Working memory, inhibitory control, and reading disability. *Memory & Cognition*, **28**(1), 8–17.

Dryer, M. S. (2013). Relationship between the order of object and verb and the order of relative clause and noun. In M. S. Dryer & M. Haspelmath (Eds.), *WALS Online (v2020.3)*. Zenodo.

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: types of intervention in the acquisition of A-bar dependencies. *Lingua*, **119**(1), 67–88.

Friedmann, N., Belletti, A., & Rizzi, L. (2021). Growing trees: the acquisition of the left periphery. *Glossa: A Journal of General Linguistics*, **6**(1), 131.

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

Gibson, E. (2000). The dependency locality theory: a distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain: Papers from the First Mind Articulation Project Symposium* (pp. 95–126). Cambridge, MA: MIT Press.

Gong, Y., & Cai, T. (1994). The Wechsler intelligence scale for children revised in China (C-WISC). *Chinese Journal of Clinical Psychology*, **2**(1), 1–6.

He, W., Xu, N., & Ji, R. (2017). Effects of age and location in Chinese relative clauses processing. *Journal of Psycholinguistic Research*, **46**(5), 1067–1086.

Hintz, F., Meyer, A. S., & Huettig, F. (2017). Predictors of verb-mediated anticipatory eye movements in the visual world. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **43**(9), 1352–1374.

Hu, S., Cecchetto, C., & Guasti, M. T. (2018a). A new case for structural intervention: evidence from Wenzhounese relative clauses. *Journal of East Asian Linguistics*, **27**(3), 247–273.

Hu, S., Costa, F., & Guasti, M. T. (2020). Mandarin-Italian dual-language children's comprehension of head-final and head-initial relative clauses. *Frontiers in Psychology*, **11**, 1379.

Hu, S., Delfitto, D., Yang, Y., & Vender, M. (2024). The comprehension of double negation in Chinese children with reading difficulties. *Clinical Linguistics & Phonetics*, **38**(12), 1192–1211.

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), 1–21.

Hu, S., Vender, M., Fiorin, G., & Delfitto, D. (2018b). Difficulties in comprehending affirmative and negative sentences: evidence from Chinese children with reading difficulties. *Journal of Learning Disabilities*, **51**(2), 181–193.

Huang, C., Liu, N., & Zhao, J. (2021). Different predictive roles of phonological awareness and visual attention span for early character reading fluency in Chinese. *The Journal of General Psychology*, **148**(1), 45–66.

Huettig, F., & Pickering, M. J. (2019). Literacy advantages beyond reading: prediction of spoken language. *Trends in Cognitive Sciences*, **23**(6), 464–475.

Joshi, R. M. (2005). Vocabulary: a critical component of comprehension. *Reading & Writing Quarterly*, **21**(3), 209–219.

Leikin, M., & Assayag-Bouskila, O. (2004). Expression of syntactic complexity in sentence comprehension: a comparison between dyslexic and regular readers. *Reading and Writing*, **17**(7), 801–822.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance*, **38**(4), 843–847.

Mann, V. A., Shankweiler, D., & Smith, S. T. (1984). The association between comprehension of spoken sentences and early reading ability: the role of phonetic representation. *Journal of Child Language*, **11**(3), 627–643.

Nation, K., Marshall, C. M., & Altmann, G. T. M. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology*, **86**(4), 314–329.

R Development Core Team. (2023). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.* https://www.R-project.org/.

Rizzi, L. (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.

Rizzi, L. (2004). Locality and the left periphery. In A. Belletti (Ed.), *Structures and beyond: The cartography of syntactic structures* (Vol. **3**, pp. 223–251). Oxford: Oxford University Press.

Rizzi, L. (2018). Intervention effects in grammar and language acquisition. *Probus*, **30**(2), 339–367.

Robertson, E. K., & Joanisse, M. F. (2010). Spoken sentence comprehension in children with dyslexia and language impairment: the roles of syntax and working memory. *Applied Psycholinguistics*, **31**(1), 141–165.

Rommers, J., Meyer, A. S., & Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attention, Perception, & Psychophysics*, **77**(3), 720–730.

Sang, B., & Miao, X. (1990). The revision of trail norm of Peabody Picture Vocabulary Test revised (PPVT-R) in Shanghai proper. *Journal of Psychological Science*, **5**, 20–25.

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools.

Shankweiler, D., & Crain, S. (1986). Language mechanisms and reading disorder: a modular approach. *Cognition*, **24**(1), 139–168.

Shankweiler, D., Crain, S., Katz, L., Fowler, A. E., Liberman, A. M., Brady, S. A., . . . Shaywitz, B. A. (1995). Cognitive profiles of reading-disabled children: comparison of language skills in phonology, morphology, and syntax. *Psychological Science*, **6**(3), 149–156.

Shankweiler, D., Smith, S. T., & Mann, V. A. (1984). Repetition and comprehension of spoken sentences by reading-disabled children. *Brain and Language*, **23**(2), 241–257.

Smith, S. T., Macaruso, P., Shankweiler, D., & Crain, S. (1989). Syntactic comprehension in young poor readers. *Applied Psycholinguistics*, **10**, 429–454.

Smith-Spark, J. H., & Fisk, J. E. (2007). Working memory functioning in developmental dyslexia. *Memory*, **15**(1), 34–56.

Snowling, M. J., & Hulme, C. (2012). Annual research review: the nature and classification of reading disorders – a commentary on proposals for DSM-5. *The Journal of Child Psychology and Psychiatry*, **53**(5), 593–607.

Stein, C. L., Cairns, H. S., & Zurif, E. B. (1984). Sentence comprehension limitations related to syntactic deficits in reading-disabled children. *Applied Psycholinguistics*, **5**, 305–322.

Stella, M., & Engelhardt, P. E. (2021). Comprehension and eye movements in the processing of subject- and object-relative clauses: evidence from dyslexia and individual differences. *Brain Sciences*, **11**(7), 915.

Stevenson, H. W., Stigler, J. W., Lucker, G. W., Lee, S.-Y., Hsu, C.-C., & Kitamura, S. (1982). Reading disabilities: the case of Chinese, Japanese, and English. *Child Development*, **53**(5), 1164–1181.

Tanaka, N., Lau, E., & Lee, A. L. F. (2024). On the universality of the subject preference in the acquisition of relative clauses across languages. *First Language*, **44**(3), 219–243.

Tunmer, W. E., & Chapman, J. W. (2012). The simple view of reading redux: vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities*, **45**(5), 453–466.

Van Dyke, J. A., Johns, C. L., & Kukona, A. (2014). Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition*, **131**(3), 373–403.

Wang, X.-L., & Tao, B.-P. (1996). *Chinese character recognition test battery and assessment scale for primary school children*. Shanghai: Shanghai Education Press.

Wang, Z., Cheng-Lai, A., Song, Y., Cutting, L., Jiang, Y., Lin, O., . . . Zhou, X. (2014). A perceptual learning deficit in Chinese developmental dyslexia as revealed by visual texture discrimination training. *Dyslexia*, **20**(3), 280–296.

Wiseheart, R., & Altmann, L. J. P. (2018). Spoken sentence production in college students with dyslexia: working memory and vocabulary effects. *International Journal of Language & Communication Disorders*, **53**(2), 355–369.

Wiseheart, R., Altmann, L. J. P., Park, H., & Lombardino, L. J. (2009). Sentence comprehension in young adults with developmental dyslexia. *Annals of Dyslexia*, **59**(2), 151–167.

Zhang, C., Zhang, J., Yin, R., Zhou, J., & Chang, S. (1996). Experimental research on the reading disability of Chinese students. *Psychological Science*, **19**(4), 222–226.

Zhang, H.-C., & Wang, X.-P. (1985). *Raven standard progressive matrices: Chinese city revision*. Beijing: Beijing Normal University Press.

Zhao, J., Bi, H.-Y., & Coltheart, M. (2017). Different relationship of magnocellular-dorsal function and reading-related skills between Chinese developing and skilled readers. *PLoS ONE*, **12**(7), e0179712.

Zhao, J., Liu, H., Li, J., Sun, H., Liu, Z., Gao, J., . . . Huang, C. (2019). Improving sentence reading performance in Chinese children with developmental dyslexia by training based on visual attention span. *Scientific Reports*, **9**, 18964.

Zhou, J., Guo, L., & Chen, Y. (2013). Reliability and validity of the Chinese version of Swanson, Nolan, and Pelham version IV rating scale-parent form for attention-deficit/hyperactivity disorder. *Chinese Mental Health Journal*, **27**(6), 424–428.