Article

Cognitive and linguistic effects of narrative-based language intervention in children with Developmental Language Disorder

Laura J Pauls and Lisa MD Archibald ⓘ

School of Communication Sciences and Disorders, Faculty of Health Sciences, The University of
Western Ontario, Canada

Abstract

Background and aims: Narrative-based language intervention provides a naturalistic context for targeting overall story structure and specific syntactic goals in children with Developmental Language Disorder (DLD). Given the cognitive demands of narratives, narrative-based language intervention also has the potential to positively impact related abilities such as working memory and academic skills.

Methods: Ten children (8–11 years old) with DLD completed 15 sessions of narrative-based language intervention.

Results: Results of single subject data revealed gains in language for five participants, four of whom improved on a probe tapping working memory. An additional four participants improved on a working memory probe only. On standardized measures, clinically significant gains were noted for one additional participant on a language measure and one additional participant on a visuospatial working memory. Carry over to reading was noted for three participants and to math for one participant. Across measures, gains in both verbal and visuospatial working memory were common. A responder analysis revealed that improvement in language may be associated with higher verbal short-term memory and receptive language at baseline. Those with working memory impairments were among those showing the fewest improvements across measures.

Conclusions: Narrative-based language intervention impacted verbal skills in different ways across individual children with DLD.

Implications: Further research is needed to gain an understanding of who benefits most from narrative-based language intervention.

Keywords

Narrative intervention, language impairment, Developmental Language Disorders, working memory impairment, near transfer, far transfer

Introduction

The ability to tell a story is particularly important for school age children; narrative ability has been linked to better outcomes both socially (Davidson et al., 2017) and academically (Griffin et al., 2004). The cognitive demands of generating or retelling a narrative are quite high, requiring support from a range of cognitive-linguistic resources (Duinmeijer et al., 2012).

One population that has particular difficulty with narratives is children with Developmental Language Disorder (DLD), who demonstrate a persistent language disorder with a significant impact on educational or social outcomes not attributable to other biomedical conditions (Bishop et al., 2017). Children with DLD have demonstrated difficulty with many aspects of narration, such as making logical connections between

Corresponding author:
Lisa MD Archibald, School of Communication Sciences and Disorders, The University of Western Ontario, London, ON N6C 1J1, Canada.
Email: larchiba@uwo.ca

story events (e.g. Reilly et al., 2004), establishing a sense of continuity (e.g. Liles, 1985), or describing characters' feelings or intentions (e.g. Klecan-Aker & Kelty, 1990). Because of the importance of narratives in both social and academic realms, recent research has explored various narrative interventions for children with DLD. Results, however, have not always been favourable (e.g. Green & Klecan-Aker, 2012), possibly due in part to heterogeneity among children with DLD. The present study addressed this problem by testing the effectiveness of a narrative-based language intervention for school age children with DLD with a single-subject design and examining effects on language, working memory, and academics as well as factors influencing response to the intervention.

The role of narratives has been well documented in both social and academic realms of childhood. Narratives make up the majority of conversations among young children (Preece, 1987). Narrative ability is a critical skill for maintaining friendships and fitting in with peers (Davidson et al., 2017), and predicts later reading comprehension even after controlling for non-verbal intelligence and initial reading ability (Botting et al., 2006). In fact, narrative skill has been found to predict academic outcomes up to seven years later including reading comprehension (Griffin et al., 2004), vocabulary (Dickinson & McCabe, 2001), and involvement in academic remediation (Fazio et al., 1996).

Key components of a well-crafted story can be categorized broadly as either macrostructure or microstructure elements (Liles et al., 1995). Macrostructure, also called story grammar, refers to the global framework of the narrative, or the way the content of the story is organized (van Dijk & Kintsch, 1983). Understanding the typical macrostructural framework for narratives facilitates not only generation of stories but also comprehension of oral narratives. Microstructure refers to the word- and sentence-level components of a story, such as the variety of vocabulary, clarity of cohesion or pronominal references (e.g. Liles, 1985), or complexity of syntax (e.g. Liles et al., 1995).

In addition to linguistic knowledge, narrative ability also relies on other cognitive resources, such as short-term memory, the temporary storage of information, and working memory, which additionally involves manipulation of the information briefly held in short-term memory (Baddeley & Hitch, 1974). Working memory supports complex language processing such as narrative discourse by maintaining activation of verbal information until semantic, grammatical, planning, and inferencing processes can be completed (Acheson & MacDonald, 2009; Yeari, 2017). It has been suggested that working memory may be important for encoding and incorporating components of a story into an integrated mental representation of the narrative (Montgomery et al., 2009). This hypothesis has broad support from other studies reporting correlations between short-term and working memory abilities and narrative comprehension and recall (e.g. Duinmeijer et al., 2012).

Children with DLD (also known as specific language impairment) typically demonstrate simpler syntax (e.g. Nippold et al., 2009), higher rates of grammatical error (e.g. Owen & Leonard, 2006), and greater difficulty acquiring new vocabulary (e.g. Kan & Windsor, 2010). Although a number of generalizations can be asserted about DLD, it is important to note the heterogeneity among children with DLD. One factor contributing to the heterogeneity among children with DLD is working memory capacity. It is well-established that children with DLD demonstrate limited verbal short-term memory (Graf Estes et al., 2007); however, there is evidence that only some children with DLD show deficits in working memory capacity (Archibald & Joanisse, 2009). This variation in presentation is likely to affect performance on tasks known to correlate with working memory, such as narrative tasks (e.g. Duinmeijer et al., 2012).

Not surprisingly, children with DLD have demonstrated many weaknesses in narrative ability. In regard to content and story structure, narratives by children with language impairment include fewer complete episodes (Merritt & Liles, 1987), poorer coherence (Liles, 1985), and more off-topic comments and disordered sequences of events (Miranda et al., 1998) relative to peers. Children with DLD tend to produce shorter narratives (Colozzo et al., 2011) with little elaboration (Ukrainetz & Gillam, 2009) using fewer cognitive state terms (Bishop & Donlan, 2005) and fewer elaborated noun phrases (Greenhalgh & Strong, 2001). In addition, narratives of children with DLD are often grammatically weaker than their peers' narratives, as demonstrated by shorter sentences (Scott & Windsor, 2000), fewer dependent clauses (Bishop & Donlan, 2005), less variety of complex syntactical structure (Reilly et al., 2004) and fewer instances of combining different complex forms (Gillam & Johnston, 1992). As well, narratives of children with DLD have been judged to be of poorer quality even when rated by laypersons or teachers (Newman & MacGregor, 2006). Finally, poor narrative ability among children with DLD has been shown to persist into adulthood (Wetherell et al., 2007).

In a recent systematic review, Petersen (2010) defined narrative intervention as interventions using oral narratives as a medium whereby language-related features are modeled by a clinician and practiced by the participant. Generally, stories connected to a series of

pictures are modeled, and supports are gradually withdrawn so participants independently retell the story by the end of the intervention steps (e.g. Gillam et al., 2012; Swanson et al., 2005). Intervention strategies include focused stimulation, scaffolding, and dialogic reading to target micro- and macrostructure elements (e.g. Gillam et al., 2012; Swanson et al., 2005). The majority of research on narrative intervention is aimed at children up to eight years old; however, older children with DLD continue to have difficulties with narrative abilities (Wetherell et al., 2007) and benefit from language intervention (Ebbels et al., 2017). Given the association between narrative language and other abilities, such as working memory and academic performance, it is possible that improvement in linguistic ability may lead to carry over gains in related domains. Therefore, the present study aimed to test the effectiveness of a narrative-based language intervention on language and related abilities among children with DLD aged 8–11 years.

Narrative-based language intervention with school age children with language impairment has been examined in a few studies with some reporting improvements on both story grammar and linguistic outcome measures (Davies et al., 2004; Gillam & Gillam, 2016; Petersen et al., 2010) and others demonstrating macro- but no microstructure gains (Green & Klecan-Aker, 2012; Fey et al., 2010). Reasons for the discrepancy in these findings are multifaceted. For example, the intervention reported by Gillam and Gillam was substantially longer (43 sessions compared to 18 or fewer sessions in the other cited studies). In addition, Davies et al. included a bespoke measure of language structures targeted while others used a free-form narrative generation task (e.g. Green & Klecan-Aker, 2012; Fey et al., 2010). Finally, the Petersen et al. study included participants with severe language impairments, which may have impacted response to intervention. These findings clearly point to the need to examine factors associated with narrative-based language intervention outcomes in detail. For instance, it is possible that some variation in intervention effectiveness is associated with heterogeneity among children with DLD. Although the heterogeneity in DLD is not well understood, a variety of factors may influence profiles known to characterize and vary amongst children with DLD, and could be associated with response to intervention. One possible factor is related to language skills. Specifically, language impairment with a receptive language component has been found to be more resistant to intervention than expressive-only impairments (Boyle et al., 2009). In the present study, we examine language and other factors associated with both the response to intervention and intervention outcomes.

In addition to linguistic skills, narratives have been shown to tap other cognitive mechanisms such as memory (Montgomery et al., 2009). Therefore, it is plausible that intervention targeting narrative ability might affect memory or other academic skills that share similar cognitive demands, such as reading and math, a question that was examined in the present study. Broad support for working memory effects following language intervention is provided by studies showing transfer to verbal short-term and working memory after phonological awareness interventions (Park et al., 2014; van Kleeck et al., 2006). However, the effect of narrative intervention on working memory is seldom measured. One study (Swanson et al., 2005) found that narrative-based language intervention had no effect on verbal short-term memory as measured by a nonword repetition task, however more research including multiple memory measures is needed.

Findings of associations between reading comprehension and both oral language (Nation et al., 2004) and narrative ability (Roth et al., 1996) have prompted researchers to advocate for the use of oral narrative language intervention as a strategy to support reading (Perfetti et al., 2005; Scott, 2009). Earlier studies repeatedly demonstrated that explicit instruction in story grammar led to improvements in reading comprehension of narratives among children with learning disabilities (see Gersten et al., 2001 for review). A more recent study (Clarke et al., 2010) found that an oral language intervention targeting story grammar alongside other language goals (e.g. vocabulary, figurative language) led to better long term reading comprehension gains than a parallel text-based intervention among children with poor reading comprehension (8–9 years).

Studies of language and math have shown associations between language ability and performance on a wide variety of mathematical tasks (Kleemans et al., 2018), and between language and word problems in particular (Fuchs et al., 2006). These associations are reinforced by findings of poor math skills among children with DLD (Cowan et al., 2005). As well, higher math scores have been found in children demonstrating strength in narrative ability relative to syntax or vocabulary (Feagans & Appelbaum, 1986), suggesting a unique link between narrative ability and math. This link is strengthened by findings that narrative ability in preschool was related to math performance two years later (O'Neill et al., 2004). The strength of the association between math and language suggests that a language intervention may lead to improvement in math ability.

Using a multiple-baseline single subject design, the present study tested the effectiveness of a narrative-based language intervention for school age children

with language impairment and a range of working memory abilities including those with working memory impairment. All children were offered language intervention following the same basic structure and using the same story books; however, the intervention was individualized by adjusting the targeted level of sentence complexity to suit each child's abilities. Intervention effects were measured using probes completed throughout the baseline, intervention, and follow-up phases and designed to be sensitive to change in language and/or working memory or neither skill. Additionally, an assessment battery was administered before, immediately after, and three months after completion of the intervention to measure language, working memory, reading, and math abilities. The first aim was to examine the impact of the intervention on language and related domains. We expected positive changes for all participants in the specific skills targeted by the intervention, namely knowledge and use of story grammar and complex syntax. Considering the cognitive demands of narrative retell and the importance of narrative ability to later academic success, we anticipated that cross over effects could occur in the related domains of working memory, reading, and math abilities. We hypothesized that those who showed greater cross over benefits to working memory would be those with the lowest working memory at baseline, and perhaps only those with working memory impairments. Our second goal was to investigate participant response to intervention by examining patterns between responders and their baseline ability in language or working memory. We were interested in whether baseline language or working memory abilities could account for any observed heterogeneity in response to intervention in a multiple-baseline design with several participants.

Methods

Participants

Participants were 10 children (8 male) recruited from a database of children from a previous study ($n = 383$; Archibald, Oram Cardy, Joanisse, & Ansari, 2013), for which children completed an assessment battery on two occasions approximately one year apart. The battery included standardized measures of working memory, language, nonverbal intelligence, reading, and math at both time points, and parent and teacher reports at the first time point. A total of 42 children in the database met criteria for either a language or working memory impairment (as outlined below), of which 29 could be contacted and 16 agreed to participate in this and a companion study (Pauls & Archibald, 2021, 2021). Of the 16 who agreed to participate, 8 met

criteria for a language impairment only and were included in the present study and 4 met criteria for a working memory impairment only and were included in the companion study. The remaining four participants met criteria for both a language and working memory impairment, two of whom were selected randomly to participate in the current study.

Participants in the present study were recruited based on meeting the following criteria for having a language impairment: (a) standard score below 86 on the Composite Language Score (CLS) of the *Clinical Evaluation of Language Fundamentals—4* (CELF–4; Semel et al., 2003) at the second time point in the previous study; (b) reported teacher concern for language or reading ability at the second time point, and (c) evidence of impairment at the first time point, as indicated by two or more of the following: a low score (≤ 87) on the CELF–4, reported parent or teacher concern for language or reading ability, or a low score (≤ 87) on one or more measures of reading or math. We adopted a cutoff score for language impairment of 1 *SD* or more below the standardized language test mean, which is consistent with many previous studies but more lenient than some (Nitido & Plante, 2020). Given this, our additional criteria ensured that we had evidence of a persistent language difficulty with a functional impact for all participants (Bishop et al., 2017).

Children were also required to demonstrate average nonverbal intelligence (standard scores at or above 85) at both time points, as obtained from either the *Wechsler Abbreviated Scale of Intelligence* (WASI; Wechsler, 1999) or the *Wechsler Preschool and Primary Scale of Intelligence – Third edition* (WPPSI-III; Wechsler, 2002) according to participant age. An additional working memory score was obtained by averaging performance on three subtests from the *Automated Working Memory Assessment* (AWMA; Alloway, 2007) found to load on a working memory factor separate from language in a previous study (Archibald, 2013): two visuospatial working memory tasks (Odd One Out, Spatial Recall) and one verbal working memory task (Listening Span). No requirements were set on working memory performance for study enrollment. Of the 10 children completing the study, 2 (DLD-9, DLD-10) were considered to have a working memory impairment as indicated by low working memory scores (< 87). Table 1 summarizes demographic and baseline measures for all participants. The time span between the most recent assessment in the previous study and the initial measures for the present study ranged from 10 to 23 months. As well, the purpose of the previous study was to describe learning profiles and did not involve any intervention. It is unlikely that participation in the previous study had any impact on the current study.

Table 1. Descriptive statistics for participant demographics and inclusion measures.

	n	Male	Age (yrs)	CLS	WM comp	PIQ
DLD	8	7	10.24 (0.97)	77.25 (3.24)	99.83 (7.06)	101.88 (13.51)
LWMI	2	1	9.83 (1.41)	78.50 (0.71)	76.67 (14.14)	103.50 (20.51)
All participants	10	8	10.16 (0.99)	77.5 (2.92)	95.19 (12.51)	102.20 (13.75)

DLD: Developmental Language Disorder; LWMI: language and working memory impairment; CLS: Composite Language Score on the CELF-4; WM comp: Working Memory composite; PIQ: performance intelligence quotient.
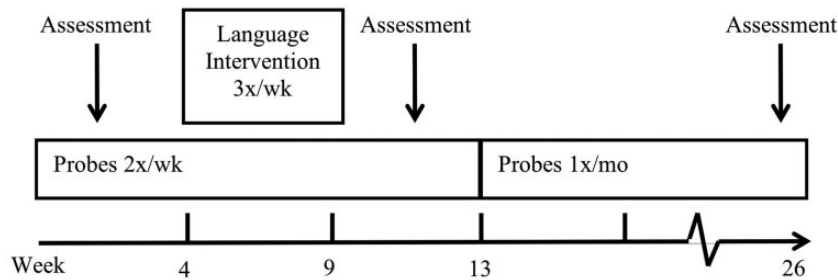
Figure 1. Study timeline.

General procedures

Study timeline. This study was one in a pair of concurrent studies with single subject designs aimed at evaluating language intervention (this study) and working memory training (Pauls & Archibald, 2021). Both studies consisted of three phases: baseline, intervention, and follow up (see Figure 1). In keeping with the single subject design, four probe measures were completed two times per week throughout the baseline phase, intervention phase, and for the first four weeks of the follow-up phase. Eight repetitions of the probe measures were completed during baseline in order to capture any change due to repetition of the probe alone. For the final three months of the follow-up phase, probe measures were administered monthly. During the five-week intervention phase, children completed three 40-minute intervention sessions each week. In addition to the probe measures, an assessment battery consisting of standardized tests of language, short-term and working memory, reading and math was completed at the beginning of the study, immediately following completion of the intervention phase and at the end of the follow up phase. Participants completed all intervention and assessment sessions individually in a quiet room in their school or home. All research sessions were completed by trained research assistants. Different research assistants completed the assessment, probe measures, and intervention sessions. All research assistants were blinded to the working memory status of the participant, and those administering the assessment and probe measures were blinded additionally to the purpose of the study.

Intervention

Initial goal selection. The narrative-based language intervention targeted both macrostructure and microstructure goals. The macrostructure goals were the same for all participants, namely to promote understanding and use of story grammar components. In contrast, microstructure goals were individualized based on a dynamic assessment consisting of narrative retell (Lost in Space; Warr-Leeper, 1990) and expository language samples (Nippold et al., 2005) and a bespoke complex syntax task. For the bespoke measure, children were prompted to produce sentences with increasingly complex syntax, ranging from structures with simple infinitive to those with multiple instances of embedding (based on Covington et al., 2006 and Steffani, 2007). Following failure to produce a complete sentence with appropriate use of the given target structure, children were provided up to two additional prompts (i.e. a sentence starter, then a model). Structures were identified as suitable intervention targets when a child showed difficulty with them across the 3 measures, but demonstrated readiness by responding to extra prompts.

Intervention materials. The narrative intervention was similar to that described in previous studies (e.g. Gillam et al., 2012; Swanson et al., 2005) and incorporated materials from published children's books: Small Saul (Spires, 2011); Stanley's Party (Bailey, 2003); The Boy Who Loved Bananas (Elliott, 2005); Purple, Green, and Yellow (Munsch, 1992); and Willow's Whispers (Button, 2010).

Intervention procedure

Each week (three sessions) focused on a different story book and followed the same basic pattern of activities. Each session was comprised of an introductory discussion of the theme, interactive readings and retellings of the story, and additional activities to promote deeper understanding of vocabulary and story structure. Each session ended with the child providing spontaneous language samples, which were recorded and later transcribed.

Intervention activities for each day are outlined in Figure 2. On Day 1, the theme was introduced and existing knowledge activated through a discussion of concepts related to the theme. Relevant vocabulary was highlighted with discussion of meaning and phonological features. Where possible, images and sketched drawings were used to support comprehension. Additional activities on Day 1 provided further introduction to the story. The research assistant engaged in dialogic reading of the story. In the third activity, the research assistant and child collaboratively retold the story using visuals depicting the story characters and settings. Throughout the retelling, the research assistant offered scaffolding by using story grammar terms, pointing out new vocabulary, and recasting the child's comments into complete complex sentences using grammatical structures at the child's microstructure goal level. Next, the child was asked to recall pertinent vocabulary from the story based on given semantic and phonological clues.

Activity	Day One	Day Two	Day Three
Introduction of Theme	• Introduce main theme • Brainstorm known concepts related to theme • Highlight related vocabulary, with visual support	• Review main theme • Introduce secondary theme using strategies from Day One	• Review themes, key vocabulary • Discuss concepts learned in previous sessions
Interactive Story Reading	• RA leads interactive reading of story • While reading, clarify new vocabulary, discuss characters' feelings, make predictions and personal connections • Child answers comprehension questions	• RA leads interactive reading using scripts adapted for child's syntax targets • While reading, discuss implied meanings or character motivations, imagine alternative outcomes	• Child retells story • RA prompts with sentence starters, probes for elaboration where necessary
Retell	• Child retells story using paint program • RA offers scaffolding for story grammar terms and complex syntax	• Child retells story from perspective of secondary character, using paint program	• Discuss conflicts in story: how they were addressed, related personal experiences, alternate solutions
Comprehension Activity	• Child solves riddles targeting relevant story vocabulary	• Given a story event, child indicates when it occurred and which events preceded and followed it	• Child points to details in illustrations based on given clue
Independent Retell	• Unaided story retell • Spontaneous expository sample on related topic	• Unaided story retell • Spontaneous expository sample on related topic	• Unaided story retell • Spontaneous expository sample on related topic • Retell new story

Figure 2. Intervention session structure.

Finally, the child provided an unaided retelling of the story as well as an expository sample related to the theme.

Day 2 began with a review and introduction of the secondary theme. For the interactive story reading, the research assistant read a version of the script adapted to include more exemplars of the child's syntax targets. Additional activities on Day 2 served to draw attention to implicit elements of the story. Throughout the reading, the research assistant asked the child about aspects of the story not explicitly stated in the text, such as the characters' motivations or the meaning of idiomatic phrases, and engaged the child in imagining possible alternative events to those in the story. Then, the child used visuals depicting the story characters and settings to recount the story from the perspective of a character other than the main character. Next, the child was given an event from the story and indicated whether it was from the beginning, middle, or end of the story. The session ended with another expository and narrative sample.

Day 3 began with a review of key vocabulary and themes. For the interactive story reading, the research assistant prompted the child to complete sentences using starter phrases targeting the child's microstructure goal level. Additional activities on Day 3 focused on story elaborations and connecting story elements. Children were asked to elaborate on the story by adding further details about the settings, the characters' feelings, or minor events as prompted by the illustrations in the book. Next, children and research assistants discussed each of the problems or conflicts in the story, attempts to address the conflicts in the story, possible alternate solutions to the conflict, and any related personal experiences. For the fourth activity, children were asked to point to details in the illustrations based on clues from the research assistant. The final spontaneous speech samples once again included expository and narrative retell samples as well as a retell of a new story, which had a plot structure similar to the theme story.

Treatment fidelity

Intervention sessions were conducted by six different coaches, including two speech-language pathologists (SLPs), three masters students in an SLP program, and one research assistant. All coaches completed rigorous training with the first author, which involved instruction in complex syntax structures, viewing videotapes of sessions, and role playing aspects of the sessions. In addition, 19% of the sessions were observed by the first author, and monitored for essential criteria using a bespoke fidelity checklist. In these observations, 98% of the intervention elements were executed as intended. When elements were missed, the observing author offered further education to the coaches through discussion and modeling.

Frequent school absences affected data collection for two participants. As a result, one participant (DLD-9) received the intervention over the course of seven weeks instead of the prescribed five weeks. Follow-up data collection for another participant (DLD-10) was limited to a single time point.

Outcome measures

Probe measures. The probes were designed to measure change in the independent variables of interest, namely language and working memory. Tasks were adapted from tasks well described in the literature or often used in relevant standardized tests. The tasks were also chosen because they could be administered consistently in a short period of time and reliably scored. The probes measured language (sentence combining with strategies to reduce memory demands), verbal short-term memory (nonword repetition), or visuospatial working memory (puzzle completion). A control probe (number comparison) not expected to place demands on language or working memory was also included. In a single subject design, change on a control probe would suggest a general response rather than a specific response to the intervention.

The *Sentence Combining probe* was a measure of language. In this probe, the research assistant read two simple sentences aloud (e.g. "Selena flies her kite" and "It is not very windy") and asked the child to combine them into a single sentence twice over (e.g. "Selena flies her kite even though it is not very windy" and "It is not very windy but Selena still tries to fly her kite"). In each session, the child completed three new trials drawn from a pool of 80 pairs of simple sentence with vocabulary considered to be familiar to young children. The 80 sentence pairs were grouped by the types of sentences they might elicit, and pairs were taken from each group throughout all intervention phases (e.g. pairs with the potential to elicit combining two things about a subject; combining a subject and object; using mental state verbs, using -ing complements, prepositional phrases, wh-clauses, and conjunctions). Sentences were transcribed and scored by calculating two measures of sentence productivity: words per sentence and propositions per sentence. Propositions (i.e. ideas) per sentence provides a measure of language richness not entirely captured by number of words spoken (Smolík et al., 2016). For example, "Jason cleans up his toys at lunchtime" has 7 words and 3 propositions (cleans up, his, at) whereas "Her favourite dress is the one that looks like it has big polka dots" has 14 words and 8 propositions (her,

favourite, is, that, looks, like, has, big). The Sentence Completion probe was designed to tap syntactical knowledge. Memory demands were minimized by using short individual sentences and providing repetitions of the verbal material as necessary.

The *Nonword Repetition probe* was designed to place demands on verbal working memory due to the requirements to selectively attend or ignore selected stimuli. In this probe, children listened over headphones to audio recorded trials of four three-syllable nonwords (e.g. da-moy-cho, tay-chee-dow, tow-doy-foo, voo-ta-yee), some of which were spoken by a male and some by a female voice. Children were instructed to listen for the 1–3 nonwords spoken by that day's targeted voice and recall those words at the end of each trial. Each session, 4–8 of the 12 nonwords presented had to be recalled for a total of 12–24 target syllables per session. The percent of target syllables correctly recalled was scored.

The *Puzzle Completion probe* was designed to tap visuospatial working memory. Children were shown a design for five seconds and were provided with seven plastic shapes to recreate the design from memory. Children were timed as they recreated the design using three or four of the provided shapes. Children were asked to recreate three designs each session. The score for each session was calculated by dividing the total number of shapes selected correctly by the total time required to recreate all three designs.

Finally, for the *Number Comparison probe* was included as a control probe. As such, no improvements were expected on this probe. Children were shown 56–60 pairs of dot arrays on a worksheet (see Figure 3), and were timed as they crossed out the array in each pair that contained the greater number of dots. Percent correct items was scored.

Assessment battery. The assessment battery included two subtests from the CELF-4: *Concepts and Following Directions*, a receptive language task in which children pointed to objects as indicated by increasingly lengthy verbal instructions, and *Recalling Sentences*, an expressive language task in which children repeated sentences read aloud by the examiner. As measures of working memory, children completed three subtests from the AWMA: *Digit Recall*, *Counting Recall,* and *Spatial*

Recall. In the verbal short-term memory task, Digit Recall, children repeated lists of numbers of increasing length. In the verbal working memory task, Counting Recall, children first counted red circles in arrays of mixed shapes, and at the end of the trial recalled their tallies. In the visuospatial working memory task, Spatial Recall, children recalled locations of a red dot after first completing a mental rotation task on a shape associated with the red dot. Reading ability was assessed with the *Test of Word Reading Efficiency* (TOWRE; Torgensen et al., 1999). In the *Phonemic Decoding Efficiency* (PDE) subtest, children were given 45 seconds to read as many nonwords as possible. In the *Sight Word Efficiency* (SWE) subtest, children were given 45 seconds to read as many words as possible. The *Reading Fluency* subtest from the *Woodcock-Johnson-III Tests of Achievement* was also completed (WJ-III; Woodcock et al., 2001), in which children read sentences and made truth judgments about them, completing as many as possible in three minutes. For math measures, the *Math Fluency* subtest from the WJ-III was completed, in which children were given three minutes to solve simple addition, subtraction, and multiplication questions, as was the *Calculations* subtest from the WJ-III, in which children solved increasingly difficult arithmetic problems.

Analysis

For the probe data, statistically significant change was tested using the proportion/frequency approach (Bloom et al., 2006). Briefly, a 2 standard deviation band was calculated from baseline data points, which then served as a benchmark for determining whether data points in the intervention or follow-up phases were successes (exceeding the 2 *SD* band) or failures (falling below the 2 *SD* band). The principles of binomial probability were used to determine whether a child's rate of success in the intervention or follow-up phase (i.e. the ratio of success to all data points in the phase) was significantly different from the rate of success in the baseline phase. For the Sentence Combining probe, it was necessary to adopt a more lenient benchmark of 1 *SD* in order to capture any reliable changes, which is consistent with the subtle changes commonly found following language intervention. As a second analysis of intervention effects, effect sizes were
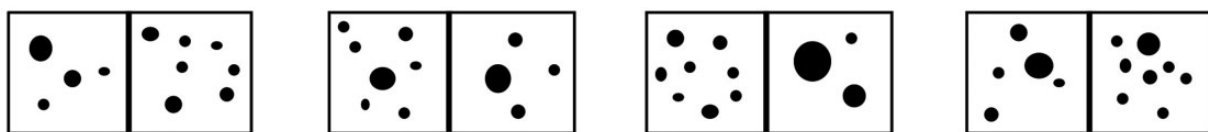

Figure 3. Example of dot array pairs from the number comparison probe. Child draws a line through the array in each pair with more dots working as quickly and accurately as possible.

Table 2. Effect sizes of probe measures.

| Participant | Sentence combining | | | | Nonword repetition | | Puzzle completion | | Number comparison | |
| | Words | | Propositions | | | | | | | |
	I	F	I	F	I	F	I	F	I	F
DLD-1	0.10	0.12	−0.06	0.04	1.86[a]	2.01[a]	1.76[a]	1.40[a]	0.13	0.40
DLD-2	−0.29	0.67	0.15	0.77	0.97[a]	0.34	−0.85	−0.07	−2.09	−6.46
DLD-3	0.20	−0.92	0.16	−0.59	−0.28	−0.54	0.68	2.11[a]	0.09	−0.42
DLD-4	1.12[a]	1.26[a]	1.04[a]	0.57	−0.06	−0.35	0.63	2.57[a]	−1.34	−2.60
DLD-5	−0.48	−0.59	−0.17	−0.81	−0.74	−0.91	0.01	0.59	−0.66	−0.25
DLD-6	1.06[a]	1.14[a]	0.72	0.77	1.81[a]	1.02[a]	−0.40	−0.51	−0.28	−0.12
DLD-7	−0.09	−0.13	0.13	−0.03	−0.67	−0.53	1.72[a]	0.54	−0.90	−1.28
DLD-8	−0.23	0.19	−0.23	−0.33	0.05	0.37	1.20[a]	0.63	0.04	−0.02
DLD-9	−0.64	−0.59	−0.68	−0.81	−0.21	−0.40	1.52[a]	2.24[a]	−0.47	−0.60
DLD-10	2.73[a]	3.00[a]	2.97[a]	2.21[a]	−1.10	−1.07	−0.72	0.09	−1.57	−0.36

I: intervention phase; F: follow-up phase.
[a]Large effect sizes ($d \geq 0.8$).

calculated as standard mean differences (SMD; Busk & Serlin, 1992), an output broadly comparable to Cohen's *d* (Cohen, 1988) and employed in other intervention studies with children with language impairment (e.g. Ebert et al., 2012). An SMD of 0.8 or greater was interpreted as a clinically significant treatment effect. For measures standardized around a mean of 100, this translated to a minimum increase of 12 standard points. For scaled measures standardized around a mean of 10, a minimum increase of 3 points was required. Additional analyses examined the possible influence of baseline abilities on training effects.

Results

Probe measures

Results from the probes indicating improvement according to the proportion/frequency approach and effect size calculations are summarized in Tables 2 and 3, respectively. Studying the results for the Sentence Combining probe (Figure 4) reveals intervention effects for 50% of participants (DLD-1, DLD-2, DLD-4, DLD-6, DLD-10). Of these five, two participants (DLD-4, DLD-10) showed improvements for both words and propositions per sentence at intervention and follow-up as measured by effect size and the 1 *SD* bandwidth method. According to the 1 *SD* bandwidth method, a third participant (DLD-6) demonstrated large significant increases in words per sentence (and large effect sizes) and significant but moderate improvements in propositions. Two additional participants (DLD-1, DLD-2) showed gains on the 1 *SD* bandwidth method at follow-up only. DLD-1 showed significant but small increases in propositions

per sentence, and DLD-2 showed significant moderate increases in both word and propositions per sentence. Performance on the Nonword Repetition probe (Figure 5) revealed intervention effects for three participants (30%). DLD-1 showed large effect sizes and significant 2 *SD* bandwidth change in both the intervention and follow-up phases, and DLD-6 showed large effect sizes (only) in both phases. In contrast, DLD-2 demonstrated a large treatment effect during the intervention phase only, although not significant according to the 2 *SD* bandwidth analysis. Results from the Puzzle Completion probe (Figure 6) showed large effects and significant 2 *SD* bandwidth changes for 50% of participants (DLD-1, DLD-3, DLD-4, DLD-7, DLD-9). Of these five, one (DLD-7) showed improvement during intervention only, and two (DLD-3 and DLD-4) showed improvements at follow-up only. DLD-1 showed a large effect with significant change in intervention but only a large effect at follow-up. DLD-9 showed a large effect in intervention and a large effect with significant change at follow-up. One additional participant (DLD-8) showed a large effect size during the intervention phase only, although not significant according to the 2 *SD* bandwidth analysis. On the Number Comparison probe (Figure 6), the 2 *SD* band exceeded 100% accuracy for all participants; therefore, the 2 *SD* limit was set to 100%. Despite high accuracy scores and a lenient cutoff, none of the participants showed ceiling effects. In addition, no participants (0%) showed gains on the Number Comparison probe according to either the proportion/frequency approach or effect size calculations.

In summary, according to the stringent criteria of positive results across the two methods of analysis and

Table 3. Summary of results from probes and standardized measures of language, working memory, reading, and math.

	Probes				Standardized measures			
	Sent Comb	Nwd Rep	Puzz Comp	Num Comp	Language	WM	Reading	Math
DLD-1	✓[F]	✓[IF]	✓[IF]		CFD[I]	CR[IF], SR[IF]	PDE[F]	
DLD-2	✓[IF]	✓[I]			RS[IF]	CR[I], SR[I]		MF[I]
DLD-6	✓[IF]	✓[IF]				CR[I]		
DLD-4	✓[IF]		✓[F]					
DLD-10	✓[IF]					SR[F]		
DLD-5								
DLD-7			✓[I]				PDE[F]	
DLD-8			✓[I]			SR[I]	PDE[I], RF[IF]	
DLD-9			✓[IF]					
DLD-3			✓[F]		CFD[F], RS[I]	DR[F], SR[F]		

✓: Improvement in probes according to either proportion/frequency or effect size calculations; [I]: improvement during or post-intervention; [F]: improvement during or at follow-up; Sent Comb: sentence combining probe; Nwd Rep: Nonword Repetition probe; Puzz Comp: Puzzle Completion probe; Num Comp: Number Comparison probe; CFD: Concepts and Following Directions; RS: Recalling Sentences; CR: Counting Recall; DR: Digit Recall; SR: Spatial Recall; PDE: Phonemic Decoding Efficiency; RF: Reading Fluency; MF: Math Fluency.

at both the intervention and follow up, score increases were noted for 30% of participants on the sentence combining (language) probe, 20% on the nonword repetition (verbal working memory) probe, and 20% on the puzzle completion (visuospatial working memory) probe. These results include one participant who improved on all probes, and one who improved on two probes (language and verbal working memory). Considering a more lenient criteria of positive results according to either method of analysis at follow up, score increases were noted for 50% of participants on the sentence combining probe, 20% on the Nonword Repetition probe, and 40% on the puzzle completion probe. These results include one participant who improved on all probes, and two who improved on two probes ($n = 1$: language and verbal working memory; $n = 1$: language and visuospatial working memory).

Standardized measures

Results from standardized measures of language, working memory, reading, and math are summarized in Table 3, which shows the subtests for which clinically significant changes were observed for the assessment at the immediate post intervention (I) or follow up phase (F). Improvements on language subtests from the CELF-4 were noted at either intervention or follow-up for 30% of participants (DLD-1, DLD-2, DLD-3). In one case (DLD-2), improvement following intervention was maintained at follow-up. In the other two cases, increases were seen at either post-intervention or follow-up only.

Working memory measures showed gains for 60% of participants (DLD-1, DLD-2, DLD-3, DLD-6, DLD-8, DLD-10). Of these, one showed improvement both post-intervention and at follow-up (DLD-1).

Three participants scored significantly higher at post-intervention testing only (DLD-2, DLD-6, DLD-8) and two showed increases at follow-up only (DLD-3, DLD-10). Performance on reading measures showed positive change for 30% of participants, DLD-1, DLD-7, DLD-8. Scores of 2 of these participants (DLD-1, DLD-7) showed an upward trajectory throughout all three testing sessions, reaching a clinically significant change at follow-up testing. The third (DLD-8) demonstrated large improvements at both post-intervention and follow-up testing. Performance on math measures showed a clinically significant change for only one participant (DLD-2).

To summarize across all measures, positive probe (either analysis method) or standardized test results at follow up were observed for 60% of participants in the area of language (sentence combining, recalling sentences, or Concepts and Following Directions), 30% in the area of verbal short-term or working memory (nonword repetition, counting recall, or digit recall), and 50% in the area of visuospatial working memory (puzzle completion or spatial recall). These results include two participants who improved in all areas, and three who improved in two areas ($n = 1$: language and verbal working memory; $n = 2$: language and visuospatial working memory). With regards to academic outcomes based on standardized tests, 30% of participants showed an improvement at follow up, all in reading. Across all measures, 80% of participants had increased scores on a language or literacy measure at follow up.

Responder analysis

In examining responders, five participants (50%) showed a response on the language probe (Sentence
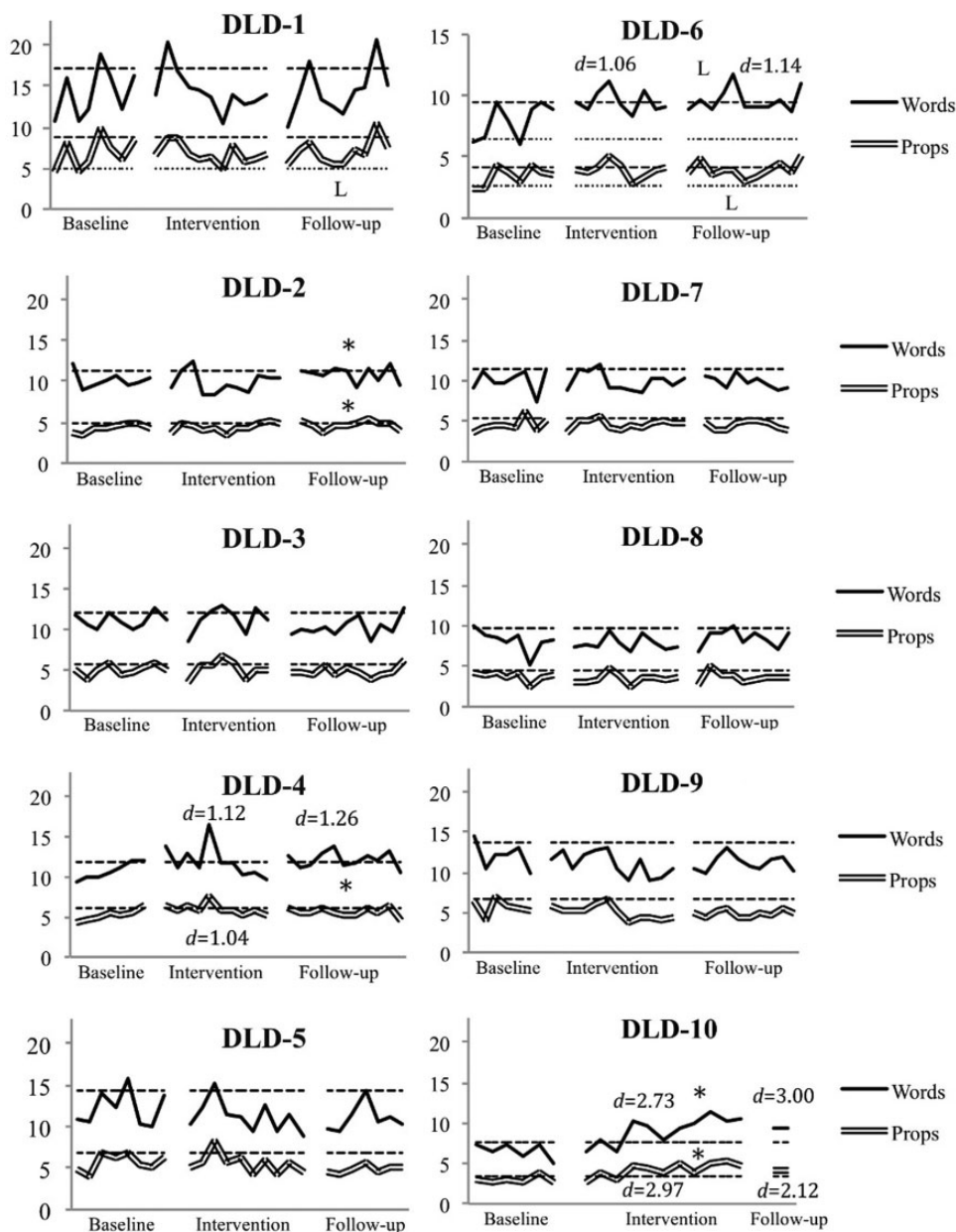
Figure 4. Sentence combining probe. Graphs present the number of words per trial and propositions per trial. Dashed line represents 1 SD above mean baseline performance. Dotted line represents 1 SD below mean baseline performance. Asterisks indicate significance according to +1 SD limit. L indicates significance according to −1 SD limit. All unmarked effect sizes d < 0.8.

Combining). Descriptively, these 'Language Responders' were differentiated from 'Language Nonresponders' by higher Digit Recall (verbal short-term memory) and Concepts and Following Directions (receptive language) scores at baseline, but did not differ on the remaining baseline measures (expressive language: Recalling Sentences; working memory: Counting Recall; Spatial Span). Of the Language Responders, all showed an improvement on at least one working memory measure. Notably, the two children with working memory impairments (DLD-9,

DLD-10) improved on one or two outcome measures (respectively), while five of the remaining eight participants improved on three or more outcomes.

Discussion

The primary purpose of this study was to evaluate the effectiveness of a narrative-based language intervention that targeted both story grammar and complex syntax for school aged children. Additionally, carry over effects were tested in related domains, including
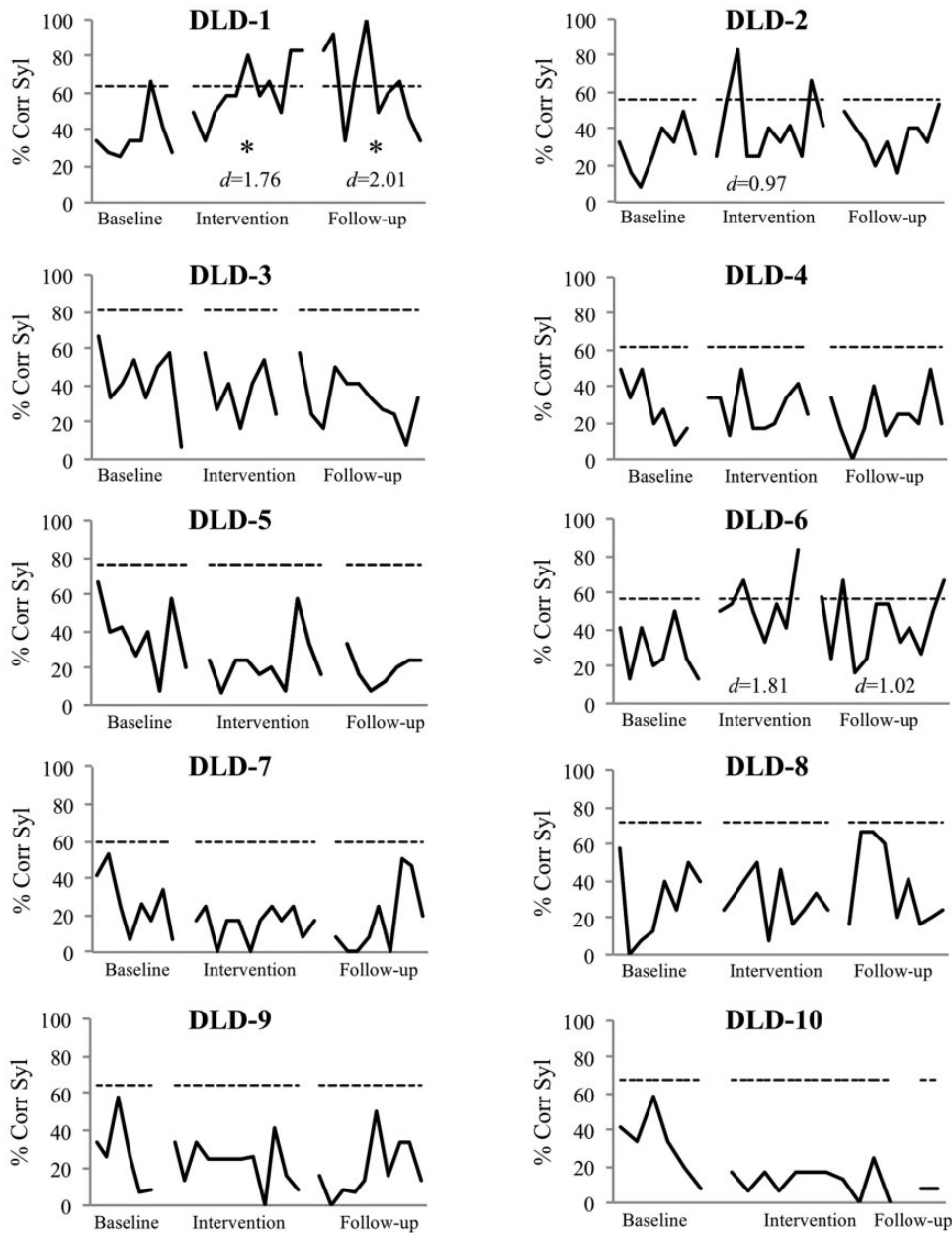
Figure 5. Nonword Repetition probe. Graphs present the percent of syllables correctly recalled in each session. Dashed line represents 2 SD above the mean baseline score. Asterisks indicate significant improvement over baseline using 2 SD limit. All unmarked effect sizes d < 0.8.

working memory, reading, and math. Language gains on the Sentence Combining probe were evident for five participants, of which four also improved on a verbal ($n = 3$) or visuospatial working memory probe ($n = 2$). Four additional participants improved on the visuospatial working memory probe (Puzzle Completion) only. Clinically significant standardized test gains were noted for one additional participant on a language measure and one additional participant on a visuospatial working memory. Carry over to reading was noted for 3 participants and to math for 1

participant. Overall, at follow up, 80% of participants improved on a language or literacy measure, and 60% on a working memory measure. Of those improving on a language or literacy measure at following up, 62% (5 of 8) also improved on a working memory measure (verbal: $n = 3$; visuospatial: $n = 3$; note: one participant improved on both). Only one participant improved on the visuospatial working memory probe with no other improvements in a verbal working memory or language measure. One additional participant did not improve on any measure. A responder analyses revealed that
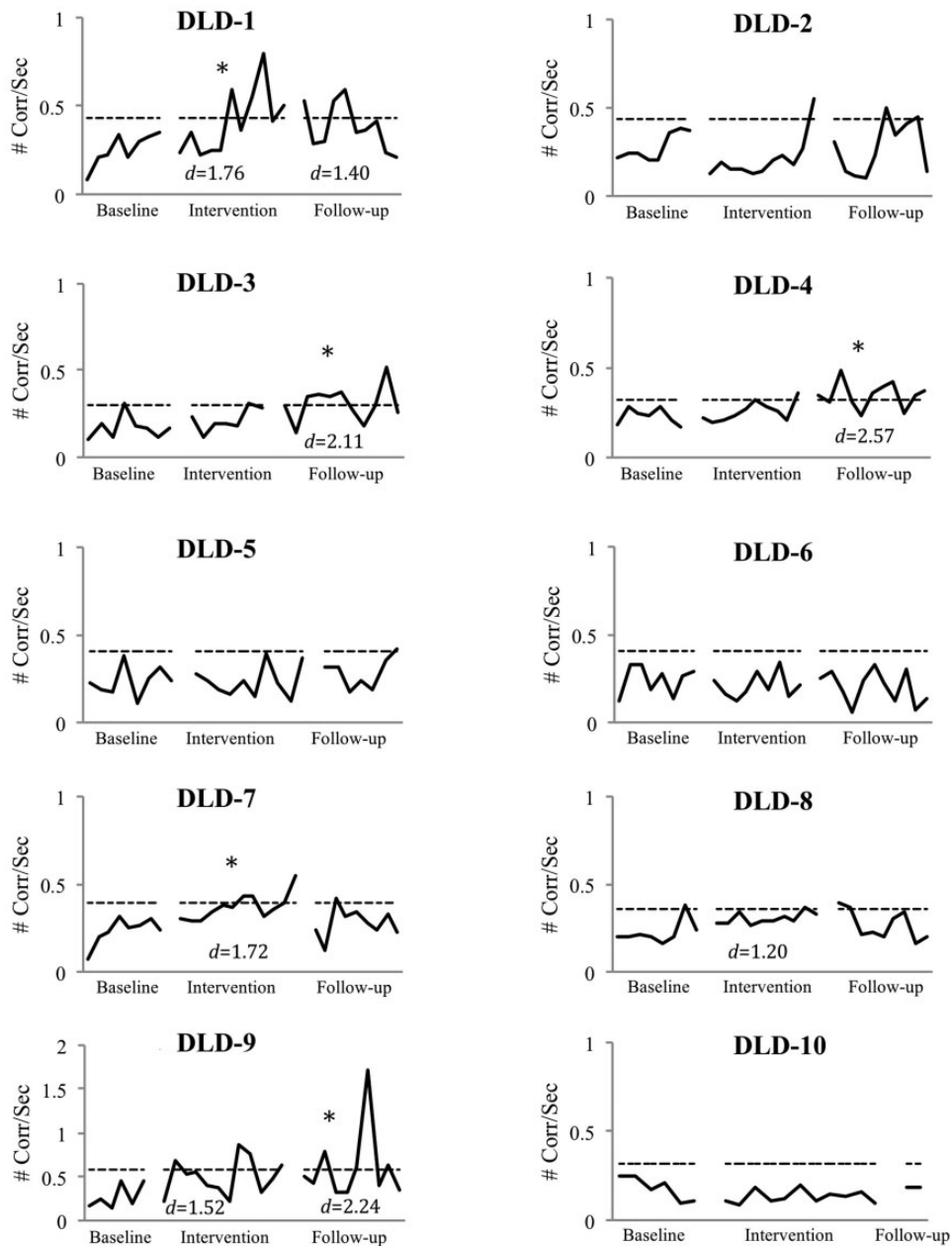
Figure 6. Puzzle Completion probe. Graphs present the correct number of shapes selected per second averaged over all three trials for each session. Dashed line represents 2 SD above mean score at baseline. Asterisks indicate significant improvement using 2 SD limit. All unmarked effect sizes d < 0.8.

improvement in language may be associated with higher verbal short-term memory and receptive language at baseline, and that those with working memory impairments were among the poorest responders.

Consistent with some previous findings (Gillam & Gillam, 2016; Petersen et al., 2010), we observed oral language gains for 6 of our 10 school age participants with DLD who received narrative-based language intervention in the present study, and all of these

gains were present at follow up. Nevertheless, the remaining participants failed to show convincing oral language change (see also, Fey et al., 2010; Swanson et al., 2005). A responder analysis based on the Sentence Combining Probe indicated that higher verbal short-term memory and receptive language baseline abilities were associated with better language gains. In contrast, those with a working memory impairment showed limited response to intervention. These results are consistent with previous studies documenting the
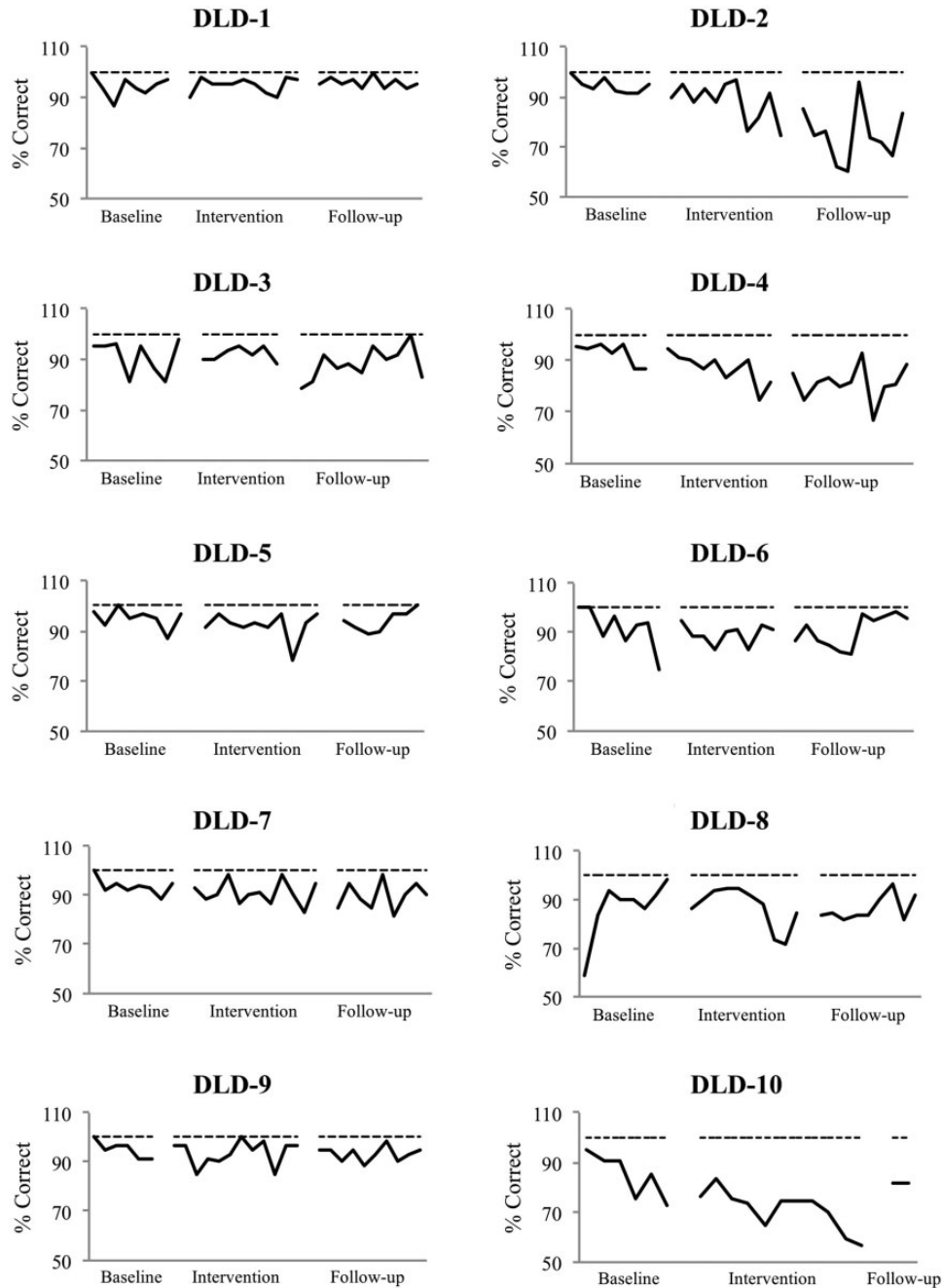
Figure 7. Number comparison probe. Graphs present percent items correct from each session. Dashed line indicates 100% items correct in place of 2 SD limit.

involvement of verbal short-term memory in vocabulary acquisition (Gathercole, 2006) and language ability in general (e.g. Baddeley, 2003), and lend further evidence to the importance of verbal short-term memory for learning language. These findings would suggest that acquisition of new linguistic forms, in this case complex syntax, is facilitated by short-term retention of verbal material. Better short-term retention allows the listener sufficient time to process the

stimuli, which in turn leads to better long-term retention of the stimuli. In contrast, poor verbal short-term memory restricts the potential for fully processing a verbal signal. In such cases, a listener may need repeated exposures to the same stimuli in order to adequately process the signal (e.g. Gray, 2004). It is possible that poor verbal short-term memory may be particularly restrictive when attempting to process lengthy verbal stimuli, as is the case with learning complex syntax.

Similarly, positive associations between baseline receptive language and response to intervention are in line with other research reporting greater gains for children with expressive-only language impairment (e.g. Boyle et al., 2009). Stronger language ability likely facilitated language gains in a number of ways. For instance, higher language abilities could have aided comprehension of the stories employed in the intervention, which, in turn, would have supported comprehension of the target complex syntax structures. In addition, existing language knowledge is known to support short-term retention of verbal material (e.g. Hulme et al., 2003), allowing greater opportunity for processing the target structures.

This study also examined transfer effects in domains related to language including working memory, reading, and math. Improvements on a visuospatial working memory probe or standardized measure unrelated to the intervention activities were observed for 90% of participants. Of the six participants (60%) showing working memory gains at follow up, one improved on only one measure and the remaining five were mixed in their verbal or visuospatial domain of improvement. More conservatively, only five participants showed increases on at least two working memory measures at any study phase, which included a verbal working memory measure for four of them. Across participants, then, working memory gains (when they occurred) crossed verbal and visuospatial domains. These findings contrast with those from working memory training studies in which persistent gains are observed for visuospatial rather than verbal domains (Melby-Lervåg & Hulme, 2013; AUTHORS, submitted). Taken together, these related findings suggest that the way in which working memory is impacted differs for narrative language vs. working memory training. It may be that the increased verbal facility afforded by narrative language intervention improves working memory efficiency across domains. This suggestion is in keeping with reports that nonverbal executive function is mediated by language (Botting et al., 2017). Overall, the results provide preliminary evidence for verbal working memory gains from narrative language intervention.

Transfer effects to reading and math were not widespread in this study, and only the reading effects were observed at follow up. On the one hand, these effects could be considered far transfer because they represent improvement in a different domain (i.e. academic) from that targeted in the intervention (i.e. language). However, the fact that only reading effects persisted at follow up suggests that these effects may be better interpreted as near transfer due to increased verbal facility. Review of phonological features for words discussed was an inherent part of the intervention.

Resultant improvements in phonological awareness could account for the improved nonword reading and reading fluency recorded in this study (see also, Park et al., 2014; van Kleeck et al., 2006).

This study employed a single subject design in order to evaluate individual responses to narrative-based language intervention. We employed four probe measures to evaluate outcomes including probes expected to respond to language intervention directly, to changes in working memory, and one control probe not expected to respond to training. One limitation of the study was the lack of stable baselines for many participants (see Figures 4 to 7), general variability throughout the data, and small number of participants. Variability in the baseline, however, results in large 2 SD bands, which would make it more difficult to achieve 'improbable improvements' in our proportion/frequency statistical analysis. As a result, effects are likely under- rather than overestimated in the present study. We also employed multiple measures across the study, however these were necessary to assess outcomes across different domains (i.e. language, working memory, academic). For our sentence combining probe, we used two analytic approaches and relaxed our bandwidth to 1 SD to assess change. Given our interest in cross over effects to working memory, we included participants with a range of working memory abilities but only two who met our definition for a working memory impairment. It must be noted, however, that the majority of working memory training studies have included participants with no demonstrated impairment in working memory and reported near transfer effects (Melby-Lervåg & Hulme, 2013). Given these limitations, caution is warranted in interpreting and generalizing the present results. The range of response patterns observed across participants in the current study, however, point to the need for continued investigation of intervention effects in children with DLD.

Taken together, the results from this study have important implications for clinicians. In particular, the results demonstrate the variability both in impact of narrative-based language intervention and in response to intervention for children with DLD with different characteristics. Specifically, narrative-based language intervention was associated with a positive change on oral language measures for 60% of our participants. Measures of reading captured gains for an additional 20% of children. Verbal working memory impacts were noted for about half of those showing language or literacy changes. As well, children with DLD who showed pre-intervention strengths in verbal short-term memory, working memory, and receptive language were more likely to benefit. These findings point to the need for further research to

explore individual differences in how children with DLD respond to interventions such as narrative-based language intervention.

Acknowledgements

Declaration of conflicting interests

Funding

ORCID iD

Lisa MD Archibald https://orcid.org/0000-0002-2478-7544

References

Acheson, D. J., & MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin*, 135, 50–68. doi: 10.1037/a0014411

Alloway, T. P. (2007). *Automated working memory assessment*. Pearson Assessment.

Archibald, L. M. D., & Joanisse, M. F. (2009). On the sensitivity and specificity of nonword repetition and sentence recall to language and memory impairments in children. *Journal of Speech, Language, and Hearing Research: JSLHR*, 52, 899–914. doi:10.1044/1092-4388(2009/08-0099)

Archibald, L.M.D., Oram Cardy, J., Joanisse, M.F., & Ansari, D.. (2013). The language, working memory, and other cognitive demands of verbal tasks. *Topics of Language Disorders*, 33, 190–207. DOI: 10.1097/TLD.0b013e31829dd8af

Archibald. (2013). Language, reading, and math learning profiles in an epidemiological sample of school age children. *PLoS ONE*, 8, e77463. doi:10.1371/journal.pone.0077463

Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, 36, 189–208. doi:10.1016/S0021-9924(03)00019-4

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (*Vol. 8*, pp. 47–90). Academic Press.

Bailey, L. (2003). *Stanley's party*. Kids Can Press.

Bishop, D. V. M., & Donlan, C. (2005). The role of syntax in encoding and recall of pictorial narratives: Evidence from specific language impairment. *British Journal of Developmental Psychology*, 23, 25–46. doi:10.1348/026151004X20685

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., & Greenhalgh, T.; the CATALISE 2 Consortium. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *The Journal of Child Psychology and Psychiatry*, 58, 1068–1080. doi:10.1111/jcpp.12721

Bloom, M., Fischer, J., & Orme, J. G. (2006). *Evaluating practice: Guidelines for the accountable professional* (5th ed.). Pearson.

Botting, N., Jones, A., Marshall, C., Denmark, T., Atkinson, J., & Morgan, G. (2017). Nonverbal executive function is mediated by language: A study of deaf and hearing children. *Child Development*, 88, 1689–1700. DOI: 10.1111/cdev.12659

Botting, N., Simkin, Z., & Conti-Ramsden, G. (2006). Associated reading skills in children with a history of specific language impairment (SLI). *Reading and Writing*, 19, 77–98. doi:10.1007/s11145-005-4322-4

Boyle, J. M., McCartney, E., O'Hare, A., & Forbes, J. (2009). Direct versus indirect and individual versus group modes of language therapy for children with primary language impairment: Principal outcomes from a randomized controlled trial and economic evaluation. *International Journal of Language & Communication Disorders*, 44, 826–846. doi: 10.1080/13682820802371848

Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research designs and analysis: New directions for psychology and education* (pp. 187–212). Lawrence Erlbaum Associates.

Button, L. (2010). *Willow's whispers*. Kids Can Press.

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science*, 21, 1106–1116. doi:10.1177/0956797610375449

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Colozzo, P., Gillam, R., Wood, M., Schnell, R. D., & Johnston, J. R. (2011). Content form in the narratives of children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 54, 1609–1627. doi:10.1044/1092- 4388(2011/10-0247)

Covington, M. A., He, C., Brown, C., Naçi, L., & Brown, J. (2006). *How complex is that sentence? A proposed revision of the Rosenberg and Abbeduto D-Level Scale*. (CASPR Research Report 2006-01). University of Georgia.

Cowan, R., Donlan, C., Newton, E. J., & Llyod, D. (2005). Number skills and knowledge in children with specific language impairment. *Journal of Educational Psychology*, 97, 732–744. doi:10.1037/0022-0663.97.4.732

Davidson, A. J., Walton, M. D., Kansal, B., & Cohen, R. (2017). Narrative skill predicts peer adjustment across elementary school years. *Social Development*, 26, 891–906. doi:10.1111/sode.12219

Davies, P., Shanks, B., & Davies, K. (2004). Improving narrative skills in young children with delayed language

development. *Educational Review*, 56, 271–286. doi:10.1080/0013191042000201181

Dickinson, D. K., & McCabe, A. (2001). Bringing it all together: The multiple origins, skills, and environmental supports of early literacy. *Learning Disabilities Research and Practice*, 16, 186–202. doi:10.1111/0938-8982.00019

Duinmeijer, I., de Jong, J., & Scheper, A. (2012). Narrative abilities, memory and attention in children with a specific language impairment. *International Journal of Language & Communication Disorders*, 47, 542–555. doi:10.1111/j.1460- 6984.2012.00164.x

Ebbels, S. H., Wright, L., Brockbank, S., Godfrey, C., Harris, C., Leniston, H., Neary, K., Nicoll, H., Nicoll, L., Scott, J., & Marić, N. (2017). Effectiveness of 1:1 speech and language therapy for older children with (developmental) language disorder. *International Journal of Language & Communication Disorders*, 52, 528–539. doi:10.1111/1460-6984.12297

Ebert, K. D., Rentmeester-Disher, J., & Kohnert, K. (2012). Nonlinguistic cognitive treatment for bilingual children with primary language impairment. *Clinical Linguistics & Phonetics*, 26, 485–501. doi:10.3109/02699206.2012.660226

Elliott, G. (2005). *The boy who loved bananas*. Kids Can Press.

Fazio, B. B., Naremore, R. C., & Connell, P. J. (1996). Tracking children from poverty at risk for specific language impairment: A 3-year longitudinal study. *Journal of Speech and Hearing Research*, 39, 611–624. https://doi.org/10.1016/S0165-5876(97)81324-9

Feagans, L., & Appelbaum, M. I. (1986). Validation of language subtypes in learning disabled children. *Journal of Educational Psychology*, 78(5), 358–364. doi:10.1037/0022-0663.78.5.358

Fey, M. E., Finestack, L. H., Gajewski, B. J., Popescu, M., & Lewine, J. D. (2010). A preliminary evaluation of fast ForWord-Language as an adjuvant treatment in language intervention. *Journal of Speech, Language, and Hearing Research: JSLHR*, 53, 430–449. doi:10.1044/1092-4388 (2009/08-0225)

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., Christopher, S., & Fletcher, J. M. (2006). The cognitive correlates of third-grade skills in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, 98, 29–43. doi:10.1037/0022-0664.98.1.29

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27, 513–543. doi.org/10.1017/S0142716406060383

Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research*, 71, 279–320. doi:10.3102/00346543071002279

Gillam, S. L., & Gillam, R. B. (2016). Narrative discourse intervention for school-aged children with language impairment. *Topics in Language Disorders*, 36, 20–34. DOI: 10.1097/TLD.0000000000000081

Gillam, S. L., Gillam, R. B., & Reece, K. (2012). Language outcomes of contextualized and decontextualized language intervention: Results of an early efficacy study. *Language, Speech, and Hearing Services in Schools*, 43, 276–291. doi:10.1044/0161-1461(2011/11-0022)

Gillam, R. B., & Johnston, J. R. (1992). Spoken and written language relationships in language/learning-impaired and normally achieving school-age children. *Journal of Speech and Hearing Research*, 35, 1303–1315. doi:10.1044/jshr.3506.1303

Graf Estes, K., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research: JSLHR*, 50, 177–195. https://doi.org/10.1044/1092-4388(2007/015

Gray, S. (2004). Word learning by preschoolers with specific language impairment: Predictors and poor learners. *Journal of Speech, Language, and Hearing Research: JSLHR*, 47, 1117–1132. doi: 10.1044/1092-4388(2004/083)

Green, L. B., & Klecan-Aker, J. S. (2012). Teaching story grammar components to increase oral narrative ability: A group intervention study. *Child Language Teaching and Therapy*, 28, 263–276. doi:10.1177/0265659012456029

Greenhalgh, K. S., & Strong, C. J. (2001). Literate language features in spoken narratives of children with typical language and children with language impairments. *Language, Speech, and Hearing Services in Schools*, 32, 114–125. doi:10.1044/0161-1461(2001/010)

Griffin, T. M., Hemphill, L., Camp, L., & Wolf, D. P. (2004). Oral discourse in the preschool years and later literacy skills. *First Language*, 24, 123–147. doi:10.1177/0142723704042369

Hulme, C., Stuart, G., Brown, G. D. A., & Morin, C. (2003). High- and low-frequency words are recalled equally well in alternating lists. Evidence for associative effects in serial recall. *Journal of Memory and Language*, 49, 500–518. doi:10.1016/S0749-596X(03)00096-2

Kan, P. F., & Windsor, J. (2010). Word learning in children with language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research: JSLHR*, 53, 739–756. doi:10.1044/1092-4388(2009/08-0248)

Klecan-Aker, J. S., & Kelty, K. R. (1990). An investigation of the oral narratives of normal and language-learning disabled children. *Journal of Childhood Communication Disorders*, 13, 207–216. doi:10.1177/152574019001300207

Kleemans, T., Segers, E., & Verhoeven, L. (2018). Role of linguistic skills in fifth-grade mathematics. *Journal of Experimental Child Psychology*, 167, 404–413. doi:10.1016/j.jecp.2017.11.012

Liles, B. A., Duffy, R. J., Merritt, D. D., & Purcell, S. L. (1995). Measurement of narrative discourse ability in children with language disorders. *Journal of Speech and Hearing Research*, 38, 415–425. doi:10.1044/jshr.3802.415

Liles, B. Z. (1985). Cohesion in the narratives of normal and language disordered children. *Journal of Speech and Hearing Research*, 28, 123–133. doi:10.1044/jshr.2801.123

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, 49, 270–291. doi:10.1037/a0028228

Merritt, D. D., & Liles, B. Z. (1987). Story grammar ability in children with and without language disorder: Story generation, story retelling, and story comprehension. *Journal of Speech and Hearing Research*, 30, 539–552. doi:10.1044/jshr.304.539

Miranda, A. E., McCabe, A., & Bliss, L. S. (1998). Jumping around and leaving things out: A profile of the narrative abilities of children with specific language impairment. *Applied Psycholinguistics*, 19, 647–667. doi:10.1017/S0142716400010407

Montgomery, J. W., Polunenko, A., & Marinellie, S. A. (2009). Role of working memory in children's understanding of spoken narrative: A preliminary investigation. *Applied Psycholinguistics*, 30, 485–509. doi:10.1017/S0142716409090249

Munsch, R. (1992). *Purple, green and yellow*. Annick Press.

Nation, K., Clarke, P., Marshall, C. M., & Durand, M. (2004). Hidden language impairment in children: Parallels between poor reading comprehension and specific language impairment? *Journal of Speech, Language, and Hearing Research: JSLHR*, 47, 199–211. doi:10.1044/1092-4388(2004/017)

Newman, R. M., & McGregor, K. K. (2006). Teachers and laypersons discern quality differences between narratives produced by children with or without SLI. *Journal of Speech, Language, and Hearing Research: JSLHR*, 49, 1022–1036. doi:10.1044/1092- 4388(2006/073)

Nippold, M. A., Hesketh, L. J., Duthie, J. K., & Mansfield, T. C. (2005). Conversational versus expository discourse: A study of syntactic development in children, adolescents, and adults. *Journal of Speech, Language, and Hearing Research*, 48, 1048–1064. doi:10.1044/1092-4388(2005/073)

Nippold, M. A., Mansfield, T. C., Billow, J. L., & Tomblin, J. B. (2009). Syntactic development in adolescents with a history of language impairments: A follow-up investigation. *American Journal of Speech-Language Pathology*, 18, 241–251. doi:10.1044/1058-0360(2008/07-0049)

Nitido, H., & Plante, E. (2020). Diagnosis of developmental language disorder in research studies. *Journal of Speech, Language, and Hearing Research*, 63, 2777–2788. doi:10.1044/2020_JSLHR-20-00091

O'Neill, D. K., Pearce, M. J., & Pick, J. L. (2004). Preschool children's narratives and performance on the Peabody individualized achievement test – Revised: Evidence of a relation between early narrative and later mathematical ability. *First Language*, 24, 149–183. doi:10.1177/0142723704043529

Owen, A. J., & Leonard, L. B. (2006). The production of finite and nonfinite complement clauses by children with specific language impairment and their typically developing peers. *Journal of Speech, Language, and Hearing Research: JSLHR*, 49, 548–571. doi:10.1044/10902-4388 (2006/040)

Park, J., Ritter, M., Lombardino, L., Wiseheart, R., & Sherman, S. (2014). Phonological awareness intervention for verbal working memory skills in school-age children with specific language impairment and concomitant word reading difficulties. *International Journal of Research Studies in Language Learning*, 3, 3–22. https://doi.org/10.5861/ijrsll534 doi: 10.5861/ijrsll.2013.534

Pauls, L. J., & Archibald, L. M. D. (2021). Cognitive and linguistic effects of working memory training in children with corresponding deficits. The University of Western Ontario.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–247). Wiley-Blackwell.

Petersen, D. B. (2010). A systematic review of narrative-based language intervention with children who have language impairment. *Communication Disorders Quarterly*, 32, 207–220. doi.org/10.1177/1525740109353937

Petersen, D. B., Gillam, S. L., Spencer, T., & Gillam, R. B. (2010). The effects of literate narrative intervention on children with neurologically based language impairments: An early stage study. *Journal of Speech, Language, and Hearing Research*, 53, 961–981. doi:10.1044/1092-4388 (2009/09-0001)

Preece, A. (1987). The range of narrative forms conversationally produced by young children. *Journal of Child Language*, 14, 353–373. doi:10.1017/S0305000900012976

Reilly, J., Losh, M., Bellugi, U., & Wulfeck, B. (2004). "Frog, where are you?" Narratives in children with specific language impairment, early focal brain injury and Williams syndrome. *Brain and Language*, 88, 229–247. doi:10.1016/S0093- 934X(03)00101-9

Roth, F. P., Speece, D. L., Cooper, D. H., & Paz, S. D. L. (1996). Unresolved mysteries: How do metalinguistic and narrative skills connect with early reading? *The Journal of Special Education*, 30, 257–277. doi:10.1177/002246699603000303

Scott, C. M. (2009). A case for the sentence in reading comprehension. *Language, Speech, and Hearing Services in Schools*, 40, 184–191.

Scott, C. M., & Windsor, J. (2000). General language performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. *Journal of Speech, Language, and Hearing Research: JSLHR*, 43, 324–339. doi:10.1044/0161- 1461(2008/08-0042)

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical evaluation of language fundamentals* (4th ed.). The Psychological Corporation.

Smolík, F., Stepankova, H., Vyhnálek, M., Nikolai, T., Horáková, K., & Matêjka, S. (2016). Propositional density in spoken and written language of czech-speaking patients with mild cognitive impairment. *Journal of Speech, Language, and Hearing Research: 59*, 1461–1470. doi.org/10.1044/2016_JSLHR-L-15-0301

Spires, A. (2011). *Small saul*. Kids Can Press.

Steffani, S. A. (2007). Identifying embedded and conjoined complex sentences: Making it simple. *Contemporary Issues in Communication Science and Disorders*, *34*, 44–54. doi: 1092-5171/07/3401-0044

Swanson, L. A., Fey, M. E., Mills, C. E., & Hood, L. S. (2005). Use of narrative-based language intervention with children who have specific language impairment. *American Journal of Speech-Language Pathology*, *14*, 131–143. doi:10.1044/1058- 0360(2005/014)

Torgensen, J. K., Wagner, R. K., & Rashotte, C. S. (1999). *Test of word reading efficiency*. AGS Publishing.

Ukrainetz, T. A., & Gillam, R. B. (2009). The expressive elaboration of imaginative narratives by children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *52*, 883–898. doi:10.1044/1092-4388(2009/07-0133)

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.

van Kleeck, A., Gillam, R. B., & Hoffman, L. M. (2006). Training in phonological awareness generalizes to phonological working memory: A preliminary investigation. *The Journal of Speech and Language Pathology – Applied Behavior Analysis*, *1*, 228–243. doi:10.1037/h0100201

Warr-Leeper, G. (1990). *Lost in space* [Unpublished normative data]. The University of Western Ontario

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. The Psychological Corporation.

Wechsler, D. (2002). *Wechsler preschool and primary scale of intelligence* (3rd ed.). Pearson Assessments.

Wetherell, D., Botting, N., & Conti-Ramsden, G. (2007). Narrative in adolescent specific language impairment (SLI): A comparison with peers across two different narrative genres. *International Journal of Language & Communication Disorders*, *42*, 583–605.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson tests of achievement* (3rd ed.). Riverside Publishing. doi:10.1080/13682820601056228

Yeari, M. (2017). The role of working memory in inference generation during reading comprehension: Retention, (re)activation, or suppression or verbal information? *Learning and Individual Differences*, *56*, 1–12. doi.org/10.1016/j.lindif.2017.04.002