RUNNING HEAD: ORAL NARRATIVE SKILLS

A Longitudinal Investigation of Oral Narrative Skills
in Children with Mixed Reading Disability

Marleen F. Westerveld

Gail T. Gillon

Catherine Moran

University of Canterbury

Correspondence to:
Prof G. T. Gillon
Department of Communication Disorders
College of Science
University of Canterbury
Private Bag 4800
Christchurch
New Zealand
Email: gail.gillon@canterbury.ac.nz
Tel: +64 3 366 7001
Fax: +64 3 364 2760

Abstract

This two-year longitudinal study investigated oral narrative ability in 14 children with mixed reading disability and their age-matched peers with typical development. The children were aged between 6;4 and 7;8 at the commencement of the study and assessments were administered individually to the children on three occasions over a two-year period. Oral narratives were elicited in a personal narrative context (i.e., the child was encouraged to relate a personal experience in response to a photo prompt) and a story retelling context. Oral narrative comprehension was assessed in a fictional story context through questions relating to story structure elements. Results indicated that the children with mixed reading disability demonstrated inferior oral narrative production and oral narrative comprehension performance compared to children with typical reading development at each assessment occasion. To further explore these children's difficulties in oral narrative ability, their performance was compared to a reading comprehension-age match control group at the third assessment trial. The results suggested the children with mixed reading disability had a specific deficit in oral narrative comprehension.

Key Words: Mixed reading disability, Oral narrative ability, Personal narratives, Story retelling

Introduction

Oral narrative ability has received increasing attention over the past three decades and the importance of children's oral narrative skills to academic achievement has been well established. More specifically, close links have been found between oral narrative abilities (such as the ability to tell a well-structured story) and reading comprehension performance (e.g., Cain, 2003; Feagans & Short, 1986; Oakhill, Cain, & Yuill, 1998). It is hypothesized that skilled readers comprehend written material by building a "mental model" of the situation described in a story (see Westby, 2005). Skilled readers' oral language competencies at word, sentence, and text level combined with their knowledge of the typical structure of a story help create this mental model (or 'picture' in their mind) of what they are reading. Weak story structure knowledge may limit the formation of these mental models, not only restricting children's ability to derive meaning from written text, but also constraining their ability to produce well-structured oral narratives (or stories). Older children with specific reading disability¹ are known to demonstrate difficulties in oral narrative ability (e.g., Roth & Spekman, 1986; Snyder & Downey, 1991). However, little is known about how children with specific reading disability develop oral narrative skills over time and whether the difficulties in oral narrative performance are a consequence of their limited reading experience. This study begins to explore this issue by comparing oral narrative skills in young school-age children with specific reading disability to their age-matched peers with typical development, over a two-year period. The oral narrative performance of the poor readers will be also be compared to a younger reading comprehension-age matched control group of children with typical development at the final assessment trial.

Classifying children with specific reading disability

Children with specific reading disability are a heterogeneous group and much debate has focused on the classification of subgroups of children with specific reading disability (e.g., Catts & Kamhi, 2005). One approach to poor reading classification is to distinguish between poor readers who have deficits in word recognition and those who have deficits in both word recognition and listening comprehension (the Reading Component Model: Aaron, Joshi, & Williams, 1999; Catts, Hogan, & Fey, 2003). This distinction is based on the “Simple View of Reading”, which considers reading comprehension the product of word recognition and listening comprehension (Gough & Tunmer, 1986). According to this model, four subgroups of children with specific reading disability can be identified: 1) children who have difficulties in word recognition alone (often referred to as dyslexia), 2) children who have difficulties in listening comprehension but not in word recognition (referred to as having a specific comprehension deficit), 3) children with a mixed reading disability who have deficits in both word recognition and listening comprehension, and 4) children with a non-specified reading disability who show good word recognition and listening comprehension skills.

Children with mixed reading disability (MRD) are the focus of this study. This subgroup of poor readers forms a sizeable portion of children with reading difficulties. Catts et al. (2003) found that from a cohort of 183 second-grade poor readers, 35% presented with dyslexia. Children with dyslexia have been the focus of much research, with phonological processing deficits a main contributing factor to their word recognition and subsequent reading comprehension deficits (see Gillon, 2004). However, children with MRD also accounted for 35% of the poor readers in Catts et al.’s study. The language development of children with MRD has been less well described (Catts, Hogan, & Adlof, 2005). Given the association of poor listening comprehension (e.g., the difficulty in answering questions in response to an

orally presented text; Hoover & Gough, 1990) to reading comprehension deficits, understanding the development of oral language in this group becomes critical. This study examines whether children who initially present with a mixed reading disability show persistent weakness in aspects of their oral language, namely oral narrative ability.

Assessment of oral narrative ability

Oral narrative ability can be assessed in two modalities (comprehension and production) and analyzed at two levels (macrostructure and microstructure). At macrostructure level, story structure knowledge is thought to play an important role in both the understanding and the generation of oral narratives (or stories). This includes understanding and identifying structural parts of the story, such as story grammar elements (e.g., identifying the characters and setting), and perceiving the main theme of a story. Analysis of oral narrative ability at microstructure level provides detail of the child's expressive semantic and morphosyntactic abilities. Oral narrative discourse requires higher-level formulation and organization skills to integrate language skills at word, sentence, and text level (Hughes, McGillivray, & Schmidek, 1997). An oral narrative elicitation context is therefore appropriate in revealing a child's oral language weaknesses in semantics and morphosyntax (Hadley, 1998). Language deficits in these domains impact upon reading development by restricting children's ability to use linguistic context for both word recognition and reading comprehension (Nation & Snowling, 2004). Although phonological awareness is critical to word recognition acquisition and subsequent reading comprehension performance (Gillon, 2004), oral language skills outside of the phonological domain can predict reading comprehension, both concurrently and longitudinally (Bishop & Adams, 1990; Catts, Fey, Tomblin, & Zhang, 2002).

To analyze children's oral narrative production to gain insight into semantic, syntactic, and morphological development it is commonly accepted that at least 50 complete and intelligible utterances of spontaneous language are required (Miller, 1996). However, most research into oral narrative language abilities of children with language and/or reading impairment has used fictional story (re)tellings (e.g., Feagans & Short, 1984; Fey, Catts, Proctor-Williams, Tomblin, & Zhang, 2004; Paul, Hernandez, Taylor, & Johnson, 1996; Snyder & Downey, 1991), which typically do not yield a sufficient number of utterances for this type of microstructure analysis. Another oral narrative elicitation context that is appropriate for this age group and potentially provides a higher number of utterances is personal narratives (see also Westerveld, Gillon, & Miller, 2004). One study that has investigated personal narrative abilities in school-age children with language impairment found these children produced personal narratives that were significantly less coherent (on measures of topic maintenance, event sequencing, and explicitness) compared to their typically developing peers (Miranda, McCabe, & Bliss, 1998). Unfortunately, the length of the language samples was not reported and no microstructure analysis data were presented. Further research is required to examine whether microstructure analysis of oral narrative production samples of sufficient length differentiates children who are poor readers from their peers with typical reading development.

Oral narrative development in children with reading disability

Snyder and Downey (1991) used a cross-sectional age comparison design to investigate the oral language skills (including oral narrative ability) and reading comprehension performance over time in groups of children with differing reading profiles, by comparing the performance of 93 children with specific reading disability and 93 children with typical development, from

8 to 14 years of age. Results indicated significant differences between the two groups of children on oral narrative production (number of “ideas” recalled in a story retelling task) and oral narrative comprehension (ability to answer questions about fictional stories) measures. Furthermore, the older children within each group performed significantly better on the story retelling task than the younger children, but there was no significant effect for age on the story comprehension task. Multiple regression analyses, however, indicated that the oral narrative retell measure accounted for significant variance in the reading comprehension scores of all typically developing children but not for the children with reading disability. Rather, performance on the oral narrative comprehension task accounted for significant variance in the reading comprehension performance of the older children with reading disability. These results not only suggest persistent oral narrative difficulties in children with specific reading disability but also indicate a markedly different relationship between oral narrative ability and reading comprehension performance for children with specific reading disability compared to their typically developing peers. Longitudinal results obtained from a study by Feagans and Short (1984) into the oral narrative comprehension and production abilities of children with specific reading disability provide support to these assumptions. It was found that the children with reading disability demonstrated persistent difficulty retelling a script-like oral narrative compared to their typically developing peers over a 3-year period. Correlational analyses indicated moderate relationships between oral narrative measures and reading achievement for the children with specific reading disability but not for the children with typical development. It is not clear, however, if the findings from Snyder and Downey and Feagans and Short pertaining to a broad group of children with specific reading disability can be extended to a

more specific group of children with a mixed reading disability who may show a different oral language profile.

Oral narrative ability and reading comprehension performance

The influence of differing oral language profiles on reading comprehension over time has been highlighted in recent years by longitudinal studies of the relationship between oral language ability and reading comprehension (Catts et al., 2002; Roth, Speece, & Cooper, 2002). Catts et al. explored this relationship by examining reading outcomes in 302 children identified with language impairment in kindergarten. Children were classified as good or poor readers in second and fourth grades based on their performance on a reading comprehension composite score. Results indicated stable mild to moderate concurrent correlations between language measures (i.e., receptive and expressive vocabulary, grammar, and oral narrative measures) and reading comprehension in second and fourth grades. However, stepwise multiple regression analyses established that word recognition was the best predictor of reading comprehension achievement with kindergarten grammar ability being the only language measure to account for any unique variance in fourth grade reading comprehension performance (6.5% of the variance). In contrast, Roth, Speece, and Cooper's investigation of oral narrative ability and reading comprehension development in children with typical development yielded different results to Catts et al.'s findings. In their study, 39 children with typical language skills were assessed in kindergarten, grade one, and grade two, on measures of oral language (structural language, metalinguistic skill, and narrative) and reading. Oral narrative ability was measured in two modalities, story comprehension and story production (analyzed for number of propositions and episodes). Regression analyses showed that after reading related skills in kindergarten (i.e., print awareness) had been entered into the regression analysis, semantic

abilities in kindergarten were most predictive of first and second grade reading comprehension. Oral narrative production (episode analysis) in kindergarten accounted for unique variance in reading comprehension performance in grade one ($R^2 = 0.6$), while oral narrative comprehension in kindergarten accounted for unique variance in reading comprehension in grade two ($R^2 = 0.5$).

Issues of causality

Another important issue regarding the relationship between reading performance and oral narrative ability concerns the direction of this association. Poor readers are likely to read much less frequently than good readers (Stanovich, 1986), which in turn could inhibit oral language development in the areas of semantics and morphosyntax and restrict the development of story structure knowledge. It is possible, therefore, that the oral narrative difficulties exhibited by older poor readers (e.g., Roth & Spekman, 1986) are a reflection of their more limited exposure to text and opportunities to comprehend more complex text. Alternatively, underlying oral language weaknesses in children with specific reading disability may contribute to their reading comprehension difficulties by restricting their ability to use semantic and/or syntactic context for both word recognition and reading comprehension (Nation & Snowling, 2004).

Cain and Oakhill (1996) hypothesized that weak story structure knowledge in children with reading disability may be causally related to their reading comprehension difficulties. Cain and Oakhill examined story structure knowledge using an oral narrative production task. Sixteen children (aged 7 and 8) with specific reading comprehension difficulties were asked to generate oral narratives using topic prompts (e.g., the farm or the circus) and picture sequences (containing six pictures with a story title). Their stories were compared to a group of 12 skilled readers (matched for age, reading accuracy and sight vocabulary), and a group of 15 six- to

seven-year-old children, matched for reading comprehension skill. The results showed that the children with specific reading comprehension difficulties told significantly less well-structured oral narratives (that lacked cohesion and showed poor sequencing of events) compared to both the age-matched and the younger reading comprehension age-matched groups. These results suggested that the poor story structure knowledge demonstrated by the children with specific comprehension difficulties was likely to be one of the causes of poor reading comprehension skill (p. 200). To address this issue of causality, the present study compares the oral narrative performance of the group of children with a mixed reading disability to that of a group of younger children with typical development who are matched to the poor readers on reading comprehension ability.

The present study

The current study extends previous research through 1) the longitudinal nature of the design and the comparison of oral narrative ability in children identified as having a mixed reading disability to their peers with typical development, 2) a description of semantic and morphosyntactic skills of good and poor readers through microstructure analysis of personal narrative language samples, and 3) a reading comprehension-age match design comparing oral narrative performance in children with a mixed reading disability to younger average readers at the third assessment trial. Specifically the study aimed to address the following questions:

- 1) Do children with mixed reading disability demonstrate difficulty in oral narrative ability compared to their age matched peers with typical development at three assessment trials over a two-year period during the early school years?

- 2) How do children with mixed reading disability perform on oral narrative production and comprehension tasks compared to their reading comprehension-age matched peers with typical development?

Experiment 1: Longitudinal Study

Method

Participants

Fourteen children with a mixed reading disability (MRD; nine boys and five girls aged between 6;4 and 7;8 years) and 14 children with typical development (aged between 6;8 and 8;2), all of whom spoke standard New Zealand English as their only language, participated in the study (see Table 1). Class teachers were invited to refer children to the study who were aged between 6;0 and 8;0 years and were identified as poor readers through curriculum based reading assessments (e.g., Marie Clay reading assessment battery; Clay, 1993). Most children in New Zealand commence school on or near the day of their 5th birthday and therefore have received a full year of formal education by 6 years of age. From a group of 59 children with poor reading ability referred to the study, children were selected based on a mixed reading disability profile. That is, both word recognition and listening comprehension deficits contributed to their reading comprehension difficulty. Thus, children remained in the study if they scored below the average range in both reading comprehension and reading accuracy performance as described in the Neale Analysis of Reading Ability, 3rd edition, manual (NARA; Neale, 1999) and below age expectation (at or below standard score 7) on the Listening to Paragraphs subtest of the Clinical Evaluation of Language Fundamentals, 3rd

edition (CELF-3; Semel, Wiig, & Secord, 1995). To ensure these children had a specific reading disability they were also required to score within the average range (i.e., standard score between 85 and 115) on a test of nonverbal intelligence as measured by the Test Of Nonverbal Intelligence, 3rd edition (TONI; Brown, Sherbenou, & Johnsen, 1997) and to have no history of physical, neurological, sensory, or intellectual impairments as indicated by teacher report and school records. In addition, to confirm this group of children had no global receptive language delay, a standard score of at least 80 was required on the Peabody Picture Vocabulary Test, 3rd edition (Dunn & Dunn, 1997) and children were excluded if they had a history of speech and language difficulty that required referral to speech and language therapy services (as evidenced by teacher report and parent questionnaire). Fourteen children met these criteria. From this group, 29% were from schools in lower socio-economic areas, with the remaining 71% from mid or high socio-economic areas, as determined by the Ministry of Education school classification system. Ten children were of New Zealand European descent and four children were of Maori descent.

Following this selection process the children with MRD were assessed on the six core subtests of the Clinical Evaluation of Language Fundamentals 3rd Edition (Semel et al., 1995) to provide further detail on their oral language abilities. Results indicated that, as a group, the children with MRD scored at the low end of the normal range (Mean Total Language Score 85.0, SD 8.1).

Fourteen children were recruited from the same schools as the participants in the poor reading group for the control group of children with typical reading development. These participants were selected through class teacher referral of children who displayed grade appropriate reading skills in curriculum based assessments and who matched children in the

experimental group on age (+/- 6 months), gender, ethnicity, and year of schooling. These children were included in the control group if they scored above the 23rd percentile rank on the comprehension and reading accuracy measures of the NARA, above a standard score of 85 on nonverbal intelligence (TONI-3), and above a standard score of 7 on the CELF-3 subtest Listening to Paragraphs. Table 1 details the group performance on the selection tasks. As indicated in Table 1, there were no significant group differences on chronological age or nonverbal intelligence. In contrast, the two groups differed significantly on reading accuracy, [$F(1,26) = 66.3, p < .001$], reading comprehension, [$F(1,26) = 63.607, p < .001$], and listening comprehension ability [$F(1,26) = 86.345, p < .001$].

Insert Table1

Materials and Procedure

Utilizing a longitudinal repeated measures design, the children were assessed on three occasions (Time 1, Time 2, and Time 3). Assessments were approximately eight months apart to ensure the children received at least two full terms of classroom instruction between the assessment sessions. The English curriculum the children were exposed to during this time included specific goals to advance children's oral narrative ability. For example in the curriculum oral language achievement objectives it states that "students should listen to texts and recall and respond to the main ideas in an organized way relating them to personal and wider experience" (Ministry of Education, 1994). Assessment measures administered at each assessment trial comprised a standardized reading test, oral narrative production tasks, and an oral narrative comprehension measure. The first author assessed all children during morning sessions, to counteract effects of tiredness, and conducted the assessments in a quiet room in the children's school setting. Assessment sessions lasted between 1 and 2 hours.

Reading ability NARA (Neale, 1999): This test consists of a series of graded passages of fictional narratives that are to be read aloud by the child, yielding a reading accuracy score. If the child reads a word incorrectly the examiner prompts with the correct word. The children are then required to answer comprehension questions following each passage, generating a reading comprehension score. The NARA contains two parallel forms (Form 1 and Form 2) with an average reliability coefficient for accuracy scores of 0.97 and comprehension scores of 0.93. To counteract the possible effects of familiarity with the test passages, Form 1 was used at Time 1 and Time 3, whereas Form 2 was used at Time 2.

Oral narrative production tasks At Time 1 the personal narrative and the story retelling sections of the Language Sampling Protocol developed by Westerveld and Gillon (2002) were used. Details on the development and piloting of the protocol are reported in Westerveld and Gillon (1999/2000, 2001). A New Zealand database has been created comprising oral narrative language samples elicited with the Language Sampling Protocol from more than 250 children with typical development, aged between 4;6 and 7;6 years. The protocol proved a reliable tool to accurately describe children's strengths and weaknesses in speaking situations relevant to school, family, and social routine. Analyses of the oral narrative language samples (story retelling and personal narrative language samples) contained in the database revealed a clear developmental trend of increasing syntactic complexity, semantic diversity, and verbal productivity with increasing age of the participants (Westerveld, Gillon, & Miller, 2004). Time 2 and Time 3 utilized the same elicitation techniques but different materials at each assessment to maintain the children's interest in the oral narrative tasks. These procedures and materials are outlined below.

Personal narratives. The personal narrative protocol was based on a narrative elicitation procedure developed by Peterson and McCabe (1983). This technique uses short narrative prompts, embedded in conversation, to encourage children to share personal experiences with the examiner. In adapting this technique for the current investigation, photo prompts were used to encourage the child to share one of his or her personal experiences. A pocketsize photo album with a series of carefully selected photos was used for the stimulus items. Topics depicted in the photos included: a bee sting, a car accident, a playground accident, a birthday party, a school trip, and a doctor's visit. These topics were found to be successful in eliciting personal narratives in 6- to 8-year-old children (Peterson & McCabe, 1983; Westerveld & Gillon, 2001). Different photos depicting similar topics were used at each assessment time to maintain the children's interest in the task. Each photo was presented individually in separate sleeves of the photo album. The examiner provided a short prompting narrative with each photo followed by the question "Did anything like that ever happen to you"? If the child responded "no," the examiner turned the page of the photo album to the next photo. If the child responded "yes," a follow-up question was asked, "Can you tell me about it?" The child was encouraged to tell as much as possible by using as many neutral sub-prompts as needed, such as "Can you tell me more?" or "and then (what happened)?" Personal narratives were excluded from the analysis if the child responded "no" or if the child only provided one utterance in response to the examiner's request to "tell me about it."

Story retelling. In the story retelling context the child was required to listen twice to an audio recording of an unfamiliar story (while looking at pictures in a story book, which either had text in a language other than English or was wordless). After listening to the story for a second time the child was asked to retell the story without the use of the pictures. At Time 1

the child listened to an English translation of a Tokelauan story “Ko au na galo” (Anna gets lost; Swan, 1992). The story is about a Pacific Islands girl who gets lost in the city while looking for her mum and dad. At Time 2 the children listened to an English translation of a Dutch story “Kikker is een held” (Frog is a hero; Velthuijs, 1995). This is a story about four friends whose houses get flooded. At Time 3 the children listened to a story accompanying the wordless picture book “A boy, a dog, and a frog” (Mayer, 1967). The story is about a boy who decides to catch a frog for a new pet, which proves difficult. The stories were selected for several reasons: presenting a wordless picture book or a text in an unknown language prevented the children from reading the text while they heard the story and thus removed any reading advantage. Having no text or a text written in another language also provided a convincing reason for listening carefully to the tape recording (of the English version) of the text. Children from different cultures living in New Zealand were expected to be familiar with the stories’ content and vocabulary translation. Different books were used at each assessment time to maintain the children’s interest and avoid children becoming too familiar with the story. Therefore at Time 2 and Time 3 books were chosen that were comparable in story structure to the story used at Time 1. All three stories involved goal-directed behaviour carried out by animate characters and included six common story structure elements as summarized by Hughes et al. (1997, p. 118-119), i.e., setting, problem, attempts (actions to solve the problem), consequence, resolution, and ending.

Oral narrative comprehension After hearing the fictional story used for the oral narrative production task for the first time, the children were asked comprehension questions tapping underlying story structure elements, such as characters, setting, problem, attempts, and resolution (Hughes et al., 1997). At Time 1, the children were asked eight questions, at Times

2 and 3, ten questions were asked. To ensure all children had access to the same information prior to the children listening to the story for a second time, children were provided with the correct information after answering the questions if they did not respond to the question or if their answers were clearly incorrect. This method was used to make sure that if children generated weak stories, they were likely to be the result of the children's difficulty in applying story structure knowledge when retelling the story, rather than the result of the children's failure to remember and/or pay attention to important elements of the story.

Transcription and Analysis

All oral narrative language samples were tape-recorded, using an Olympus DM-1 Digital Recorder. Samples were transcribed by the first author using standard Systematic Analysis of Language Transcripts conventions (SALT; Miller & Chapman, 2003). Utterance segmentation was based on communication units (CU), using Loban's (1976) rules. Only complete and intelligible (C&I) utterances were used for analysis; interrupted and abandoned sentences were excluded, as well as utterances containing unintelligible segments.

Microstructure Analysis The personal narrative samples were cut after the first 50 C&I utterances and analyzed at microstructure level. At Time 3, one experimental child and one control child only produced 31 and 42 utterances, respectively. In these cases their full personal narrative language transcripts were included in the analyses. Quantitative measures of language ability that have been shown to distinguish between children with language impairment and children with typical language development were selected and calculated automatically using SALT.

Grammatical competence was measured as 1) grammatical complexity: the mean length of CU in morphemes (MLCU-M), and 2) grammatical accuracy (GA): the percentage of

grammatical CU's (Fey et al., 2004). Utterances that were not considered grammatically accurate were coded in SALT during the transcription process.

Semantic diversity was based on the number of different words (NDW). Several studies have indicated that NDW derived from 50 C&I's is a promising quantitative indicator of expressive vocabulary (e.g., Miller, 1996; Watkins, Kelly, Harbers, & Hollis, 1995).

Macrostructure Analysis The full story retelling transcripts were used and analyzed at macrostructure level. To evaluate the child's ability to apply story structure knowledge when retelling a story, the story retellings were scored on a story quality rubric. The rubric was adapted from Jones and Lodholz (1999) and assessed inclusion of six text structure elements (introduction, main character/s, supporting character/s, conflict, resolution, and conclusion) as well as a measure of holistic coherence. It also investigated whether the child included the theme ("the overall coherent topic of the text and its essential points;" Westby, 2005, p. 162) of the story. The story quality rubric used at Time 3 is included in Appendix A. The child was awarded points for each characteristic, 5 points if the child proficiently included the characteristic, 3 points if the skill was emerging, and 1 point if the child provided minimal or no information. The rubric included specific scoring examples to promote easy and reliable scoring by other examiners. The scores were totalled to yield a Story Quality score. Consequently the minimum score was 8 and the maximum score was 40. At Time 2 and Time 3, the same rubric was used but the scoring examples were changed to reflect the different stories.

Reliability

Twenty percent of the transcripts were analyzed by an independent examiner, experienced in language transcription, who was blind to the children's group status. The percentage of

agreement between the two examiners at each assessment time was as follows. Transcription reliability (including utterance segmentation) in % utterance agreement: stories T1: 98.4%, T2: 100%, T3: 97.8%; Personal narratives T1: 97.2%; T2: 99%; T3: 99.5%. Error coding reliability in % error-code agreement: Stories T1: 86.75%, T2: 94%, T3: 93.3%; Personal narratives T1: 89.7%, T2: 97%, T3: 96.8%. With regards to the story quality rubric, the independent examiner was trained in scoring the stories by the first author. Five stories were then independently scored by the second examiner and any disagreements were discussed and/or the wording of the rubric was revised. Following this training period another eight stories were independently scored and reliability of the total scores on the story quality rubric in Cronbach's Alpha was: T1: 0.97; T2: 0.98; T3: 0.96. The total scores (of all individual stories) awarded by the two examiners differed by 4 points or less.

Results

Two-way (group and time) repeated measures analyses of variance (ANOVA) were used to determine group differences, time differences, and interactions between group and time. Effect sizes were calculated using effect size index f with the conventional interpretation of f index as follows: small effect size $f = 0.10$; medium effect size $f = 0.25$; large effect size $f = 0.40$. Effect sizes are unaffected by sample size. A large effect size suggests that the differences in group performance would be clinically observable since a high degree of separation between the groups' performances is indicated (Portney & Watkins, 2000 p. 710-711)

Microstructure Analysis All personal narrative transcripts were cut after the first 50 C&I's and analyzed on measures of semantic diversity and grammatical competence. Table 2 reports the means, standard deviations, and effect size indices at Times 1, 2, and 3.

Grammatical competence: 1) grammatical complexity (in MLCU-M) - the children with TD consistently outperformed the children with MRD, and there was a significant main effect for group [$F(1,26) = 12.909, p < .05, f = 0.5$]. There was a significant main effect for time [$F(2,52) = 6.837, p < .05, f = 0.26$], indicating MLCU-M improved significantly over time. The interaction group x time was not significant ($p = .731$), indicating the two groups of children made similar progress. Follow-up one-way ANOVA's indicated significant differences between the two groups at all three assessment times: Time 1 [$F(1,26) = 5.932, p < .05$], Time 2 [$F(1,26) = 9.733, p < .05$], and Time 3 [$F(1,26) = 10.417, p < .05$]; 2) Grammatical accuracy: the children with TD produced a higher percentage of grammatically correct sentences than the children with MRD (group effect [$F(1,26) = 5.803, p < .05, f = 0.22$]). The main effect for time was not significant ($p = .68$), nor was the interaction group x time ($p = .491$). Follow-up one-way ANOVA's indicated no significant differences between the two groups at Time 1 ($p = 0.140$) and Time 3 ($p = 0.460$) but a significant difference at Time 2 [$F(1,26) = 7.849, p < .05$].

Semantic diversity – NDW: results from the two-way repeated measures ANOVA showed a significant main effect for group [$F(1,26) = 8.918, p < .05, f = 0.34$] and a significant effect for time [$F(2,52) = 5.64, p < .05, f = 0.22$], but the interaction group x time was not significant ($p = .618$). Follow up one-way ANOVA's showed that the TD group outperformed the MRD group at each assessment time: Time 1 [$F(1,26) = 4.315, p < .05$], Time 2 [$F(1,26) = 5.060, p < .05$], and Time 3 [$F(1,26) = 8.627, p < .05$].

Insert Table 2

Macrostructure Analysis The full story retelling transcripts were used and scored on the story quality rubric (see Table 3). Results from the two-way repeated measures ANOVA indicated a

significant main effect for group, with the children with TD outperforming the children with MRD [$F(1,26) = 15.524, p < .05, f = 0.60$]. The main effect for time was not significant ($p = .231$), nor was the interaction group $\times$ time ($p = .793$). Follow-up one-way ANOVA's indicated the TD group outperformed the MRD group at every assessment time: Time 1 [$F(1,26) = 6.060, p < .05$], Time 2 [$F(1,26) = 10.841, p < .05$], Time 3 [$F(1,26) = 9.212, p < .05$].

Insert Table 3

Oral narrative comprehension For ease of comparison, scores were converted to percentage of questions answered correctly. A repeated measures ANOVA showed a significant group effect [$F(1,26) = 42.664, p < .001$] with the TD group answering more comprehension questions correctly (see Table 4). No comparisons over time were made because different questions were used at each assessment time. As indicated in Table 4 the TD group outperformed the MRD group at every assessment time and these differences were significant: Time 1 [$F(1,26) = 9.253, p < .05$], Time 2 [$F(1,26) = 20.926, p < .001$], and Time 3 [$F(1,26) = 14.174, p = .001$].

Insert Table 4

Reading comprehension Results from a two-way repeated measures ANOVA (using the raw scores derived from the Neale Analysis of Reading Ability- revised; NARA) showed a significant main effect for group [$F(1,26) = 52.710, p < .001, f = 2.03$] with the TD group outperforming the MRD group. There was a significant effect for time [$F(2,52) = 42.398, p < .001, f = 1.63$], but the interaction group $\times$ time was not significant ($p = .507$), indicating the two groups of children made similar significant progress over time.

To further evaluate the MRD group's oral narrative skills, their performance at Time 3 was compared with younger average readers.

Method

Participants

At the completion of the third assessment trial in the longitudinal study there continued to be significant differences between the group of children with mixed reading disability (MRD) and their peers with typical development (TD) on reading accuracy, [$F(1,26) = 32.644, p < .001, f = 1.26$], and reading comprehension performance, [$F(1,26) = 33.684, p < .001, f = 1.3$]. Visual inspection of individual data revealed that 12 of the 14 children with MRD continued to perform below average on the reading comprehension measure of the NARA (two children's reading comprehension scores had improved to within normal range: percentile ranks of 23, and 28, respectively). This remaining group of eight boys and four girls (referred to hereafter as persistent poor readers, PPR) was then matched to younger average readers for reading comprehension ability (reading comprehension age $\pm$ 6 months on the NARA). T-tests showed there were no significant differences between the younger reading comprehension-age matched group of average readers (RMTD) and the PPR group on reading comprehension age ($p = .97$) or reading accuracy age ($p = .80$). Table 5 reports the mean reading accuracy and reading comprehension scores of the PPR group and the two control groups of children with typical development.

The reading comprehension-age matched participants (RMTD) were recruited from two of the schools the poor readers attended. Class teachers were invited to refer children who demonstrated grade appropriate reading and language skills and were aged between 6;0 and 7;6 (the current reading comprehension age of the persistent poor readers group). Children were

included in experiment 2, if their reading accuracy and reading comprehension were age-appropriate (on the NARA), and if their reading comprehension age (derived from the NARA) could be matched to that of one of the experimental children (i.e., no more than six months' difference between the scores). All children spoke standard New Zealand English as their only language and were from mid socio-economic areas as determined by the Ministry of Education school classification system.

Insert Table 5

Procedure

The oral narrative tasks used in the longitudinal study were administered to the RMTD participants. The same testing procedures and conditions that the PPR group and the group with typical development (TD) received at Time 3 were applied. Similar reliability procedures to experiment one were followed. Twenty percent of the transcripts were analyzed by an independent examiner who was blind to the children's group status. Transcription reliability in % utterance agreement was 95% for the stories and 98% for the personal narratives. Error coding reliability in % error-code agreement: stories 93%, personal narratives 96%. Five stories were independently scored and the reliability of the total scores on the story quality rubric in Cronbach's Alpha was 0.96.

Results

The third assessment trial results of the oral narrative analyses for the PPR group (n=12) were compared to those of the 12 RMTD children and the 12 children with TD who participated in the longitudinal experiment. For ANOVA, effect size indices *f* were calculated (see experiment 1). However, effect sizes for post-hoc comparison findings were calculated as the difference

between the means divided by the root mean square of the groups' standard deviations and reported as Cohen's d . Conventional interpretations of d were used: small effect size $d = 0.20$, medium effect size $d = 0.50$, large effect size $d = 0.80$ (Portney & Watkins, 2000).

Microstructure analysis One-way ANOVA's indicated significant group differences in syntactic competence, MLCU-M: [$F(2,33) = 4.509, p < .05, f = 0.27$], and semantic diversity, NDW: [$F(2,33) = 3.935, p < .05, f = 0.24$], but not in syntactic complexity (GA: $p = .73$). Table 6 lists the means and standard deviations for the three groups across the three productive language measures derived in the personal narrative context. Post hoc analyses (Bonferroni) indicated the following:

Syntactic competence: the TD group produced significantly longer sentences than the PPR group ($p < .05, d = 1.09$) but no significant differences were found between the PPR group and the RMTD group ($p = .696$). Although the TD group used longer sentences than the RMTD group, this difference was not statistically significant ($p = .259$).

Semantic diversity: the TD group used significantly more different words than the PPR group ($p < .05, d = 1.04$). No significant differences in semantic diversity were found between the PPR group and the RMTD group ($p = .171$), or the TD group and the RMTD group ($p = 1.00$).

Macrostructure analysis One-way ANOVA showed a significant effect for group [$F(2,33) = 4.261, p < .05, f = 0.26$]. Post-hoc analysis indicated that the TD group produced significantly better stories than the PPR group ($p < .05, d = 1.19$). No significant differences in performance were found between the PPR group and the RMTD group ($p = .651$) or the RMTD group and the TD group ($p = .324$). Group scores are reported in Table 6.

Oral narrative comprehension One-way ANOVA indicated a significant effect for group [$F(2,33) = 9.116, p = .001, f = 0.55$]. Post-hoc analysis (Bonferroni) indicated that the PPR group scored significantly below both the TD group ($p = .002, d = 1.43$) and the RMTD group ($p = .002, d = 1.36$). There were no significant differences between the TD group and the RMTD group ($p = 1.00$). Means and standard deviations for all three groups are reported in Table 6.

Insert Table 6

Discussion

This study explored the oral narrative abilities of a group of 14 children with mixed reading disability (MRD) using two experiments. First, the longitudinal examination compared the oral narrative abilities of the children with MRD to a group of 14 age-matched peers with typical development on three occasions over a two-year period. Second, the reading comprehension-age matched experiment compared the MRD group's oral narrative performance at the third assessment trial to a younger reading comprehension-age matched control group.

The first question the study addressed was whether the children with MRD would show difficulty in oral narrative ability compared to their chronological-age matched peers with typical reading development over a two-year period. Results of the longitudinal experiment clearly showed that the children with MRD demonstrated difficulty in oral narrative comprehension and production ability compared to the good readers on all three assessment trials. This is consistent with results from previous studies that have investigated aspects of oral narrative ability in children with reading disorders, such as oral narrative comprehension (i.e., the ability to answer questions about a story; Snyder & Downey, 1991) and oral narrative production at macrostructure level (i.e., the number of propositions or action units recalled;

Feagans & Short, 1984; Roth & Spekman, 1986; Snyder & Downey, 1991). Results from the current study extend previous research by demonstrating that children with MRD also exhibited weaknesses in microstructure measures of oral language ability derived in an oral narrative context, such as grammatical complexity and semantic diversity. In summary, the present study demonstrated that young children with MRD performed poorly in all aspects of oral narrative ability, and that these difficulties persisted over a two-year period.

The oral narrative comprehension difficulties experienced by the children with MRD were not surprising given these participants were selected based on listening comprehension deficits at the commencement of the study. Similarly, it was expected that the children with MRD would demonstrate difficulties in their ability to tell a well-structured story, containing all the critical story elements (including setting, characters, problem, attempts, resolution, and ending), as captured in the story quality rubric. However, results from the longitudinal investigation indicated that these difficulties were persistent, despite a national curriculum that aims to develop these skills. In contrast, the finding that the children with MRD demonstrated inferior performance on measures of semantic diversity and grammatical competence compared to their chronological-age matched peers with typical development, was unexpected. Expressive language impairment had not previously been identified by the children's teachers or parents and none of the children had been referred nor were on any waiting list for speech and language assessment. The findings from the microstructure analysis for these 14 children with MRD are however consistent with the unidentified language difficulties of other populations who have been portrayed as having a specific reading disorder, described in previous research (Gillon & Dodd, 1994; Nation, Clarke, Marshall, & Durand, 2004), and

highlight the need for comprehensive assessment of both spoken and written language competency in children who present with reading difficulties.

Results from this study provided some evidence that classroom instruction and school-based reading interventions were effective in developing the children's reading comprehension performance over time, but that they did not facilitate the accelerated progress necessary for the children with MRD to catch up to the level of their peers with typical development. When comparing the groups' development in reading comprehension skills over the two years, it was found that both groups made significant, but similar progress in reading comprehension ability. This was of some surprise and appears to go against the commonly accepted Matthew effects in reading phenomenon (Stanovich, 1986). Based on this theory, different rates of progress in reading comprehension would have been expected between the group of children with MRD and the TD group. A plausible explanation for this finding relates to the extra attention most children received from their schools, teachers, and/or parents during the course of this investigation. All of these children had already been identified by their teachers as struggling in their ability to learn to read. Subsequently some children received school-based remedial reading assistance. These interventions may have ensured that the poor readers did not fall further behind their peers over time. However, the results of this study also suggest that these interventions did not result in accelerated growth in reading comprehension required by the poor readers. Moreover, despite their progress in reading comprehension, the children with MRD continued to show difficulty in the ability to comprehend and produce oral narratives.

The second research question posed whether there would be differences in oral narrative performance between the children with MRD and a control group of reading comprehension-age matched children with typical development (RMTD). Results from the

second experiment showed that the subgroup of children with MRD who had shown persistent reading comprehension difficulties (86% of the original group) performed significantly below the reading comprehension-age matched group on oral narrative comprehension but not on the macrostructure measure of oral narrative ability (i.e., story quality). Furthermore, microstructure analyses of the oral narrative language samples revealed that the children with MRD performed much like their younger reading comprehension-age matched peers with typical development. These findings are summarized and their implications discussed, in turn.

The finding that the children with MRD performed significantly below the reading comprehension-age matched control group of children on the oral narrative comprehension measure indicates a specific oral narrative comprehension deficit. It suggests that the children with MRD may have problems forming an accurate mental model of the situation described in the story. As Bishop described (1997, p. 169) “information that is not integrated into a mental model is much more fragile and prone to be forgotten”, resulting in poor identification and/or memory of the underlying structural elements of the story (e.g., setting, characters, problem, and resolution). Although these difficulties in oral narrative comprehension experienced by the children with MRD could be the result of limited exposure to text, this seems unlikely for two reasons. First, it seems doubtful that the children in the RMTD control group, who are on average two years younger than the children in the MRD group, would have had more exposure to print. Second, the RMTD group and the MRD groups demonstrated similar word reading ability and previous research has shown strong links between word reading ability and exposure to print (see Cain, Oakhill, & Bryant, 2000, for a discussion on this topic).

When investigating the story retellings at macrostructure level, results indicated that the children with MRD performed at a similar level compared to their reading comprehension-age

matched peers. This was unexpected in view of the poor readers' weak oral narrative comprehension performance. That is, inferior performance in understanding the story compared to the younger RMTD group could be expected to lead to inferior performance in retelling the same story. The methodology used in the story comprehension task may account for this unexpected finding. The opportunity to listen to the story a second time and the provision of the correct answer to the story comprehension questions may have provided the extra support needed for the poor readers to reach the level of the young typical readers (see also Cain & Oakhill, 1996). Further research is needed to explore methods that enhance the story retelling performance of children with reading disability.

The microstructure analyses of the oral narrative language samples revealed that the poor readers performed much like their younger reading comprehension-age matched peers with typical development, suggesting a pattern of language delay. The only oral narrative language measure that failed to show sensitivity to age or reading ability profile was the grammatical accuracy measure, suggesting that the children with MRD employed in this study did not show specific morpho-syntactic deficits as have been reported in oral narrative development of children with language impairment (Fey et al., 2004). Rather, the oral narratives of the poor readers in this study were characterized by relatively short but grammatically correct sentences.

Implications, limitations, and future research

The present study investigated oral narrative ability in a group of poor readers who demonstrated a mixed reading disability using a classification system based on the Simple View of Reading (Gough & Tunmer, 1986). The results from this study underline the importance of conducting a comprehensive oral language assessment in this group of poor

readers, none of whom had previously been identified with oral language difficulties (see also Roberts & Scott, 2006). Although results from standardized testing (on the CELF-3) suggested their oral language impairments were mild, results from this longitudinal investigation showed that their oral narrative difficulties were persistent. Moreover, the children with MRD appeared to have a specific oral narrative comprehension deficit that could not be explained by lack of reading experience alone. Taken together these results lend support to the claim that intervention for children with MRD will need to address listening comprehension as well as word recognition abilities (Catts & Kamhi, 2005).

Although the present study focused on one subgroup of children with specific reading disability, there was still considerable variation in oral narrative performance within the group and across the tasks. To determine if this classification system is useful clinically in differentiating between subgroups of poor readers, further research is needed comparing the oral narrative abilities in children from differing reading subgroups, namely dyslexia, specific reading comprehension deficits, and mixed reading disability.

Acknowledgements

The authors would like to sincerely thank the children, the children's parents, and class teachers for their co-operation in this research project. We would also like to acknowledge the Tertiary Education Commission for their financial assistance. Our very grateful thanks to Professor Jon Miller and Ann Nockerts from the Language Analysis Lab at the University of Wisconsin, for their expertise and help with the SALT analyses. Our thanks are also extended to Dr Emily Lin for her assistance with the statistical analyses.

¹ The term specific reading disability is used throughout this paper to refer to poor readers who have a reading disability in the absence of sensory, emotional, neurological, physical, or cognitive disorder.

References

- Aaron, P. G., Joshi, M., & Williams, K. A. (1999). Not all reading disabilities are alike. *Journal of Learning Disabilities, 32*(2), 120-137.
- Bishop, D. V. M. (1997). *Uncommon understanding. Development and disorders of language comprehension in children*. East Sussex: Psychology Press Ltd. Publishers.
- Bishop, D. V. M., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders, and reading retardation. *Journal of Child Psychology and Psychiatry, 31*, 1027-1050.
- Brown, L., Sherbenou, R., & Johnsen, S. (1997). *Test of nonverbal intelligence: A language-free measure of cognitive ability* (3rd ed.). Austin, TX: Pro-ed.
- Cain, K. (2003). Text comprehension and its relation to coherence and cohesion in children's fictional narratives. *British Journal of Developmental Psychology, 21*, 335-351.
- Cain, K., & Oakhill, J. (1996). The nature of the relationship between comprehension skill and the ability to tell a story. *British Journal of Developmental Psychology, 14*, 187-201.
- Cain, K., Oakhill, J., & Bryant, P. (2000). Investigating the causes of reading comprehension failure: The comprehension-age match design. *Reading and Writing: An Interdisciplinary Journal, 12*, 31-40.
- Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research, 45*(6), 1142-1157.
- Catts, H. W., Hogan, T. P., & Adlof, S. M. (2005). Developmental changes in reading and reading disabilities. In H. W. Catts & A. G. Kamhi (Eds.), *Connections Between Language and Reading Disabilities* (pp. 25 - 40). Mahwah, NJ: Lawrence Erlbaum

Associates.

Catts, H. W., Hogan, T. P., & Fey, M. E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities*, 36(2), 151-164.

Catts, H. W., & Kamhi, A. G. (2005). Classification of reading disabilities. In H. W. Catts & A. G. Kamhi (Eds.), *Language and reading disabilities* (2nd ed.) (pp. 72-93). Boston, MA: Pearson Education, Inc.

Clay, M. M. (1993). *An observation survey of early literacy achievement*. Auckland, NZ: Heinemann.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test - III*. Circle Pines, MN: American Guidance Service.

Feagans, L., & Short, E. J. (1986). Referential communication and reading performance in learning disabled children over a 3-year period. *Developmental Psychology*, 22, 177-183.

Feagans, L., & Short, E. J. (1984). Developmental differences in the comprehension and production of narratives by reading-disabled and normally achieving children. *Child Development*, 55, 1727-1736.

Fey, M. E., Catts, H. W., Proctor-Williams, K., Tomblin, J. B., & Zhang, X. (2004). Oral and written story composition skills of children with language impairment. *Journal of Speech, Language, and Hearing Research*, 47(6), 1301-1318.

Gillon, G., & Dodd, B. J. (1994). A prospective study of the relationship between phonological, semantic and syntactic skills and specific reading disability. *Reading and Writing: An Interdisciplinary Journal*, 6, 321-345.

- Gillon, G. T. (2004). *Phonological awareness: From research to practice*. New York: The Guilford Press.
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7(1), 6-10.
- Hadley, P. A. (1998). Language sampling protocols for eliciting text-level discourse. *Language, Speech, and Hearing Services in Schools*, 29, 132-147.
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127-160.
- Hughes, D., McGillivray, L., & Schmidek, M. (1997). *Guide to narrative language: Procedures for assessment*. Eau Claire, WI: Thinking Publications.
- Jones, M. A., & Lodholz, C. (1999). Story Grammar Rubric. Madison Metropolitan School District and the Language Analysis Lab, Waisman Center, University of Wisconsin, Madison: Unpublished document created by the SALT working group, Jon Miller Chair.
- Loban, W. (1976). *Language development: Kindergarten through grade twelve*. Urbana, IL: National Council of Teachers of English.
- Mayer, M. (1967). *A boy, a dog, and a frog*. New York: Penguin Putnam Inc.
- Miller, J. F. (1996). Progress in assessing, describing, and defining child language disorder. In K. N. Cole, P. S. Dale & D. J. Thal (Eds.), *Assessment of communication and language* (6th ed.) (pp. 309-324). Baltimore, MD: Paul H. Brookes.
- Miller, J. F., & Chapman, R. S. (2003). *SALT: Systematic Analysis of Language Transcripts*. Madison, WI: Language Analysis Laboratory, Waisman Center, University of Wisconsin.

Ministry of Education (1994). *English in the New Zealand curriculum*. Wellington, NZ:

Learning Media.

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(4), 647-667.

Nation, K., Clarke, P., Marshall, C. M., & Durand, M. (2004). Hidden language impairments in children: Parallels between poor reading comprehension and specific language impairment? *Journal of Speech, Language, and Hearing Research*, 47(1), 199-211.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, 27(4), 342-356.

Neale, M. D. (1999). *Neale Analysis of Reading Ability* (3rd ed.). Melbourne, VIC: Australian Council for Educational Research Ltd.

Oakhill, J., Cain, K., & Yuill, N. (1998). Individual differences in children's comprehension skill: Toward an integrated model. In C. Hulme & R. M. Joshi (Eds.), *Reading and spelling development and disorders* (pp. 343-367): Lawrence Erlbaum Associates Inc.

Paul, R., Hernandez, R., Taylor, L., & Johnson, K. (1996). Narrative development in late talkers: Early school age. *Journal of Speech and Hearing Research*, 39(6), 1295-1303.

Peterson, C., & McCabe, A. (1983). *Developmental psycholinguistics: Three ways of looking at a child's narrative*. New York: Plenum.

Portney, L. G., & Watkins, M. P. (2000). *Foundations of clinical research: Applications to practice* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall Inc.

Roberts, J. A., & Scott, K. A. (2006). The simple view of reading: Assessment and

- intervention. *Topics in Language Disorders*, 26(2), 127-143.
- Roth, F. P., Speece, D. L., & Cooper, D. H. (2002). A longitudinal analysis of the connection between oral language and early reading. *Journal of Educational Research*, 95(5), 259-272.
- Roth, F. P., & Spekman, N. J. (1986). Narrative discourse: Spontaneously generated stories of learning-disabled and normally achieving students. *Journal of Speech and Hearing Disorders*, 51(1), 8-23.
- Semel, E. M., Wiig, E. H., & Secord, W. (1995). *CELF: Clinical Evaluation of Language Fundamentals* (3rd ed.). San Antonio, TX: Psychological Corp., Harcourt Brace Jovanovich.
- Snyder, L. S., & Downey, D. M. (1991). The language-reading relationship in normal and reading-disabled children. *Journal of Speech and Hearing Research*, 34, 129-140.
- Stanovich, K. E. (1986). Matthew Effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360-407.
- Swan, E. (1992). *Ko au na galo (Ana gets lost)*. Wellington, NZ: Learning Media, Ministry of Education.
- Velthuijs, M. (1995). *Frog is a hero*. London: Andersen Press Ltd.
- Watkins, R. V., Kelly, D. J., Harbers, H. M., & Hollis, W. (1995). Measuring children's lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech and Hearing Research*, 38(6), 1349-1355.
- Westby, C. E. (2005). Assessing and remediating text comprehension problems. In H. W. Catts & A. G. Kamhi (Eds.), *Language and reading disabilities* (2nd ed.) (pp. 157-232). Boston, MA: Pearson Education Inc.

Westerveld, M., & Gillon, G. (1999/2000). Narrative language sampling in young school-age children. *New Zealand Journal of Speech-Language Therapy*, 53/54, 34-41.

Westerveld, M., & Gillon, G. (2001). Oral language sampling in 6-year-old New Zealand children from different cultural backgrounds. *New Zealand Journal of Speech - Language Therapy*, 56, 5-17.

Westerveld, M., & Gillon, G. (2002). A New Zealand Spontaneous Language Sampling Protocol. Christchurch, New Zealand: University of Canterbury: Department of Speech and Language Therapy.

Westerveld, M. F., Gillon, G. T., & Miller, J. F. (2004). Spoken language samples of New Zealand children in conversation and narration. *Advances in Speech- Language Pathology*, 6(4), 195-208.

Appendix A
Story Quality Rubric “A boy, a dog, and a frog”.

CHARACTERISTIC	PROFICIENT	EMERGING	MINIMAL / IMMATURE
Introduction to story Where - forest why – to catch frog what - A frog is spotted.	Setting stated: At least 3 setting info points are mentioned	Setting stated incompletely, i.e. no more than 2 setting info points are provided. Or setting info points not provided at the beginning.	Only 1 info point provided. OR launches into the story with no attempt to provide setting or story theme.
Story Theme Frog is lonely	Story theme is stated explicitly, e.g., <i>Frog does not want to be left alone in the pond.</i>	Mentions that the frog is sad, but not why. OR says that <i>frog is alone</i> (not lonely).	No story theme stated
Main Character (s) Boy and Dog	Boy and Dog Main characters correctly introduced to listener. All further references appropriate and clear.	Main character/s referred to by <u>a</u> boy or <u>a</u> dog after initial mention. OR boy or dog referred to by ‘he’ or ‘they’ later in the story, not clear if this is the boy, dog etc.	No mention about the dog OR the dog is mentioned much later. OR main characters predominantly referred to by pronoun.
Supporting Character /s Frog	Supporting character correctly introduced to listener. All further references appropriate and clear.	Frog introduced by “the” too soon. Consistent use of “the” frog. Not clear if referred to frog, or dog.	Inconsistent use of “a” and “the”. See above.
Conflict wants to catch a frog	Includes rationale for character’s behaviour. Provides the relationship connecting events and actions.	Vague or incomplete rationale or statement for the character’s behaviour. <i>E.g., looking for a frog</i> , not why. Or looking for pet frog, not “catch”.	No rationale for character’s behaviours OR no attempt to provide a relationship connecting events and actions. E.g., only mentioned at the end: <i>‘you’re hard to catch’</i> .
Resolution The frog wants to play with them They all become friends	Clear resolution regarding characters, conflicts and events. Both resolutions are mentioned.	<u>Some</u> resolution provided for characters, conflicts or events – e.g., all having fun. Only one mentioned.	No resolution provided.
Conclusion Frog leaps into the bath. They all have fun / play happily	Smooth transition to conclusion. Both “endings” are mentioned. Might provide insight into the character’s feeling or effects of the ending.	Abruptly states: the end – that’s all. Story just ends with <i>the frog lands on the dog’s head</i> / or different plausible ending.	Stops talking and listener may need to ask if that is the end. Story finishes halfway.
Coherence	Events follow a logical order. Critical events are included and minor events ignored. Smooth transitions provided between events.	Events follow a logical order. BUT Inconsistent provisions of transitions between events. AND/OR 1 critical event missing.	Story is missing 2 or more critical events. Events are provided in random order. Minimal or no connection between events. Transitions between events are lacking.
TOTAL	X 5 =	X 3 =	X 1 =

Table 1. Group performance on the selection measures.

	Reading				
	Age (months)	Nonverbal intelligence	Listening comprehension ^c	Accuracy percentile rank ^c	Comprehension percentile rank ^c
MRD^a					
Mean	84.7	98.3	6.4	10.7	9.1
SD	5.9	6.8	0.7	6.5	6.7
Range	76 - 92	88 - 111	5 - 7	1 - 21	1 - 22
TD^b					
Mean	86.8	102.9	10.3	53.2	51.4
SD	6.2	7.0	1.9	13	12.7
Range	78 - 98	93 - 118	8 - 14	27 - 86	34 - 73

Note. ^aMRD: Mixed reading disability. ^bTD: Typically developing. Nonverbal intelligence: performance on the TONI, Test of Nonverbal Intelligence, 3rd edition, standard score; Listening comprehension: Listening to Paragraphs subtest from the Clinical Evaluation of Language Fundamentals, 3rd Edition, standard score; Reading: Performance on the NARA, Neale analysis of reading ability, 3rd Edition.

^cGroups differ at $p < .001$

Table 2. Group performance on the microstructure measures.

	NDW			MLCU-M			GA		
	Time1	Time2	Time3	Time1	Time2	Time3	Time1	Time2	Time3
MRD									
Mean	110	120.3	114.7	5.57	5.9	6.08	90.1	89.3	91.6
Range	86 - 134	94 - 145	85 - 140	4.2 – 6.7	4.5 – 7.0	3.7 – 7.8	82 - 100	78 - 98	70 - 100
SD	16.9	16.9	18.8	0.99	0.87	1.0	5.4	5.7	7.9
TD									
Mean	123.4	134.1	133.9	6.47	6.91	7.26	92.7	94.3	93.4
Range	91 - 159	113 - 162	107 - 162	4.7 – 8.2	5.4 – 8.7	5.1 – 8.3	88 - 98	90 - 100	82 - 100
SD	17.3	15.7	15.5	0.96	0.83	0.94	3.4	3.5	4.8
Effect size index f^a	.17	.19	.33	.23	.37	.40		.30	

Note. Transcripts were cut after the first 50 complete and intelligible utterances. NDW: Number of different words;

MLCU-M: Mean length of communication unit (CU) in morphemes; GA: Grammatical accuracy (percentage of grammatical CU's).

^aValues followed by an effect size index (f) within each measure are statistically significant ($p < .05$).

Table 3. Group performance on the story quality rubric.

	Score		
	Time1 ^a	Time2 ^a	Time3 ^a
MRD			
Mean	24.3	22.0	24.4
SD	6.3	4.6	5.8
Range	16 - 36	14 - 30	14 - 32
TD			
Mean	29.4	28.7	30.1
SD	4.7	6.1	4.0
Range	20 - 36	22 - 38	22 - 34
Effect size index <i>f</i>	.23	.42	.35

Note. Minimum score = 8; Maximum score = 40.

^aGroups differ at $p < .05$.

Table 4. Group performance on the oral narrative comprehension task..

	Score		
	Time1 ^a	Time2 ^b	Time3 ^a
MRD			
Mean	70.5	59.3	57.1
SD	18.1	16.9	16.4
Range	37.5 - 100	30 - 90	30 - 80
TD			
Mean	88.4	84.3	75.7
SD	12.5	11.6	8.5
Range	62.5 - 100	60 - 100	60 - 90
Effect size index <i>f</i>	.36	.80	.55

Note. Scores are presented as percentage of questions answered correctly.

^aGroups differ at $p < .05$. ^bGroups differ at $p < .001$

Table 5. Group characteristics on the reading measures at Time 3.

	Age	Reading Accuracy		Reading Comprehension	
		Age equivalent	Raw score	Age equivalent	Raw score
PPR (n=12)					
Mean (SD)	8;7 (5.8)	6;8 (7.0)	22 (10.6)	6;8 (6.31)	7.8 (3.3)
Range	7;10 – 9;2	5;11 – 7;9	6 - 38	6;0 – 7;5	3 - 12
TD (n=12)					
Mean (SD)	8;8 (6.0)	9;4 (18.9)	51.1 (12.4)	8;8 (14)	18.3 (4.8)
Range	7;11 – 9;7	7;7 – 13;1	35 - 82	7;1 – 11;9	11 - 28
RMTD (n=12)					
Mean (SD)	6;9 (5.6)	6;7 (3.9)	21.2 (4.2)	6;9 (5.4)	7.8 (3.0)
Range	6;0 – 7;5	6;1 – 7;2	14 - 29	6;0 – 7;5	2 - 12

Note. Ages are in years;months. Reading measures based on the NARA: Neale Analysis of

Reading Ability, 3rd Edition; PPR: Persistent poor readers; TD: Typically developing; RMTD:

Reading comprehension-age matched group of children with typical development.

Table 6: Group performance on the oral narrative measures at Time 3.

	PPR	TD	RMTD
Age in years;months	8;7 (5.8)	8;7 (6.0)	6;9 (5.6)
Oral narrative comprehension	5.67 (1.56) ^{ab}	7.5 (0.91) ^a	7.5 (1.09) ^b
Macrostructure			
Story quality score	23.67 (5.9) ^c	29.83 (4.3) ^c	26.33 (5.2)
Microstructure			
MLCU-M	6.16 (1.05) ^d	7.29 (1.01) ^d	6.62 (0.69)
GA	0.91 (0.08)	0.93 (0.05)	0.91 (0.05)
NDW	115.58 (19.07) ^e	133.92 (16.12) ^e	128.92 (14.08)

Note. Oral narrative comprehension: Number of correct answers out of 10 questions.

Microstructure measures of oral language were derived in the personal narrative condition;

Transcripts were cut after the first 50 complete and intelligible utterances.

Values with the same superscript letter for each variable are statistically different ($p < .05$). For example, analysis of the oral narrative comprehension task showed that the persistent poor readers answered significantly fewer questions correctly than both the typically developing children (^a) and the reading comprehension-age matched children (^b).