8-28-2024 10:00 AM

Impact of explicit whole class oral narrative intervention on oral and written language in Grade 1 students

Diya Nair, *University of Western Ontario*

Supervisor: Archibald, Lisa M.D, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Health and Rehabilitation Sciences

Abstract

Oral language is one skill known to support written language development. Story-telling or narrative skills are important oral language skills that have been found to improve with explicit teaching and predict academic outcomes. This study investigated impact of the oral narrative ~~i~~intervention on the oral and written language skills of first-grade children. In a cross-over design with 63 participants from two schools, whole class intervention in one class per school was conducted in twelve 25–30-minute sessions over three weeks either during phase 1 or phase 2 of the study while the other class in each school completed their business-as-usual curriculum. Children completed oral language measures before and after each study phase and written language samples throughout the study. Results indicated no significant improvements in oral or written language measures that were specifically associated with intervention phases. The two schools differed at baseline and showed different patterns of increases across repeated testing times. The findings are not consistent with many previous studies showing positive oral and written responses to narrative intervention.

Keywords

narrative, intervention, Story Champs™, children, oral language, written language

Summary for Lay Audience

Oral language is one skill known to support written language development. Story-telling or narrative skills are an important oral language skill that have been found to improve with clear teaching and further predicts academic outcomes. The study investigates the impact of the oral narrative intervention, on the oral and written language skills of first-grade children. Conducted across two schools with 63 participants, the research aimed to assess the changes in oral language, using the Listening Retell and Recalling Sentences and written language through several word-level and text-level measures. Story Champs™ is an oral narrative intervention program that focuses on enhancing children's academic language abilities through the foundation of storytelling. Oral narratives are seen as foundational not only for

oral language but also for written language, as children often translate their spoken stories into written form. This connection is crucial as proficiency in oral narratives can lead to better writing skills, including the use of more complex sentence structures and coherent storylines. In this study, the intervention consisted of twelve 25-30-minute sessions over three weeks. Children were divided into two groups based on who received intervention first. Results showed that there were no improvements in either oral or written language measures in either of the two groups. The study's findings indicate that Story Champs™ did not impact the children's written narratives suggesting that oral narrative instruction is not beneficial and that its transfer to written language may require more specific intervention. There were significant baseline differences between the two schools, and this significantly influenced the results. The findings are not consistent with many previous studies showing positive oral and written responses to narrative intervention. The results might indicate that factors such as socioeconomic status and practice effects need to be considered in more detail.

Acknowledgments

I would like to express my deepest appreciation to my supervisor, Professor Lisa Archibald, whose guidance, patience and expertise have been the guiding force behind the completion of my thesis. Professor Archibald has continually conveyed an incredible spirit of research and scholarship which inspired me to work on my thesis.

I would like to thank my advisory committee members, Professor Janis Oram and Stephanie Somers, whose guidance helped shape my study.

In addition, I would like to extend my gratitude to my lab members for constantly supporting and encouraging me. I would also like to thank my family for their encouragement, support and appreciation in my academic achievements.

Table of Contents

List of Tables

List of Figures

Chapter 1

(i) Introduction

It is widely acknowledged that storytelling or narratives are a crucial milestone of children's language development especially oral language development (Spencer and Slocum, 2010). This skill carries significant academic and social relevance prompting the development of intervention around narratives and their frameworks. There is considerable evidence that oral narrative instruction in young children positively impacts their narrative skills especially in overall story structure (Spencer and Peterson, 2018; Spencer and Slocum, 2010; Hayward and Schneider, 2000). Story champs is one such intervention that focusses on targeting narrative production by working on story elements such as setting and problem, and cohesion elements such as conjunctions and pronouns (Spencer and Peterson, 2018). Given the links between oral and written language, it is reasonable to expect that the positive effects of oral narrative intervention would transfer to written language outcomes. However, considerably less research has focused on this question (Spencer & Peterson, 2018). The purpose of the present study was to expand this evidence base by investigating written language outcomes following oral narrative instruction in kindergarten children.

Proficiency in language holds great significance in nearly all facets of a child's development, this includes communicating effectively, making and maintaining peer relationships and academic development (Dickinson and Snow, 1987). The role of language in academic performance is crucial and central and one that must not be downplayed (Bishop and Edmundson, 1987). It has been found that children entering school with lower than typical language experiences such as less exposure to environments with more social cues or less language stimulation at home, tend to run into notable challenges with reading and writing (Whitehurst & Lonigan, 2001).

One key oral language skill is oral narratives, or the ability to tell a story or an event or answer questions about a story (Paris & Paris, 2003). Children start to tell stories as early as 2 years of age (Kemper, 1984). They then learn to expand their stories significantly

between the ages of 2 to 5 years (Kemper, 1984). Initially, 2-year-old children construct stories without a relevant cause and/or consequence and may switch between topics frequently (Ripley, 2012). As they grow older, around 4 years of age, they learn to incorporate more macrostructure and microstructure elements which helps build a stronger story plot called primitive narratives (Peg, 2001). They learn to follow a topic and a central theme with its characters, settings and ending. As they come closer to 7 years, their stories additionally include the solution in a logical sequence in their narratives (Ripley, 2012). By the age of 8, they further learn to include more about the characters, story setting, opening and closing of a story and relevant story episodes in turning their narratives into true narratives (Peg, 2001).

Narratives consist of two broad elements: macrostructure and microstructure (Silva and Cain, 2019). The story macrostructure refers to the overall story organization and includes the episodes or series of events that help formulate a sound story (Spencer and Peterson, 2018) and individual elements such as story characters, problem in a story, story setting, solution and ending. These elements provide coherence across the story creating a story arch. The story microstructure refers to the grammatical sentence structure such as mental verbs, conjunctions, causal ties, dialogue and pronouns (Westerveld & Billon, 2009). These structures provide cohesion across the story (Silva & Cain, 2019). When assessing and providing intervention for oral narratives, it is important to examine both micro- and macrostructure elements. Narrative skill can be assessed either through storytelling tasks where a child describes a story or event either with or without prompts or story retelling where the child listens to the story narrated first and then repeats it to the best of their abilities. These tasks allow for the assessment of story coherence and cohesion through macro and microstructure analyses, and can give an overall indication of grammatical skills, vocabulary, and word finding (Nation & Snowling, 2004) as well as age-appropriate syntax, semantics, sequencing and cognition (Vandewalle et al., 2012).

Children who have language difficulties have been found to have weak narrative skills. In a study conducted by Manhardt and Rescorla (2002), children who were late talkers showed weaker macrostructure and microstructure skills with lower scores in the latter

resulting in an overall decrease in narrative production quality. Children with Developmental Language Disorder (DLD) are well known to produce narratives with fewer macrostructure elements which in turn leads to poorly structured stories, and in terms of microstructure they formulate shorter sentences to produce their narratives and use less complex sentences in terms of microstructure (Bishop & Edmunson, 1987).

Oral language skills, particularly oral narrative skills, have been linked to academic success including vocabulary, reading comprehension, and narrative writing quality (National Reading Panel, Kim et al., 2015; Heilmann et al., 2006; Botting, Simkin, & Conti-Ramsden, 2006, Griffin; Hemphill, Camp, & Wolf, 2004; Hatcher et al. 2004; Bishop & Adams, 1990). In a study of 4-year-old children involving numerous language measures, narrative skills were a better predictor of academic success at 5.5 years of age than measures of syntax and/or semantics (Bishop & Edmunson, 1987).

It is said that 'students cannot write what they cannot say' (Snow and Katz, 2010). Oral language has time and again been linked to written language (Dyson 2006; Kim et al., 2015; Snow and Katz, 2010; Spencer and Peterson 2018). Oral language lays the groundwork for a child's ability to produce written language, thus acting as a valid precursor. Proficiency in oral language abilities has been shown to be foundational to written language development (Ukrainetz, 2006). Young children often articulate their ideas vocally prior to writing them and they frequently verbalize text during the writing process (Dyson, 2009). The links between oral and written language are clearly represented in current models of reading comprehension and writing. According to the most influential model of reading comprehension, the Simple View of Reading (Farell et al, 2014), reading comprehension is supported by two basic yet crucial components: word recognition and oral language. Indeed, as children begin to master word recognition, it is their oral language skills that predict reading comprehension (Farell et al., 2014). Similarly, in the Not-So-Simple View of Writing (Berninger and Graham, 1998), one of the crucial components for writing is text generation, which includes the main goals of writing at the word, sentence and text levels. Thus, both of these models, with the inclusion of oral language in the Simple View of Reading and text generation in the Not-So-Simple View of Writing show that oral language supports written language.

In addition to the overall link between oral and written language, oral narratives in particular are foundational for written narratives (Fey et al., 2004). Narratives involve the use of structurally complex and detailed language that is extremely similar to the written language students encounter in school (Dickinson & McCabe, 2001; Gillam & Johnston, 1992) and have such commonalities as story grammar structure and complex literate language like casual and temporal ties and dialogue (Greenhalgh & Strong, 2001; Roth & Spekman, 1986; Westby, 1985; Westby, 1984). Children with well-developed oral narrative skills produce longer written narratives (Dyson, 2009, Greenhalgh & Strong, 2001, Hudson & Shapiro, 1991, Ukrainetz, 2006). Additionally, knowledge of story grammar supports the production of skilled narratives, which aids in the successful development of literacy (Hudson & Shapiro, 1991). Indeed, curriculum standards indicate that kindergarten children should be able to narrate events and stories with an appropriate order of story events such as the ability to start with the characters followed by talking about the setting, problem, solution and finally describing the ending of the story (Ontario Language Curriculum, 2023).

Given the importance of oral narratives to oral language development and potentially for written narratives, oral narrative instruction has been investigated in a number of studies. Explicit teaching of language structures has been found to effectively scaffold children's productions (Snow & Katz, 2010). When a teacher effectively models and instructs a particular sentence frame to use in narratives and uses clear and concise instructions for children (Williams and Roberts, 2011), the child's understanding of what is expected of them increases, which leads to better narrative quality, promotes better learning and further increases the likelihood of generalizing these learnt practices to everyday oral and written language situations. Children are also able to show an advancement in their narrative abilities and further write their own personal stories when given clear instructions (Bradbury, 2010). Oral language narrative instruction can also help children learn to interact with their peers and adults purposefully, productively and meaningfully (Nelson, 2019). A number of evidence based narrative intervention programs such as Let's Know (LAARC et al., 2019), Supporting Knowledge in Language and Literacy (SKILL; Gilliam et al., 2024) and Story Champs™ (Spencer and Peterson, 2012) have incorporated 'softly scripted or loosely scripted lesson plans' (Gilliam et al, 2024) in

order to ensure consistent and carefully delivered rich instruction for children with positive effects.

Story Champs™ (Spencer and Peterson, 2012) is one of the previously mentioned oral language interventions' that focuses on narrative instruction in story retelling and personal story generation in either small or large groups of children (Spencer and Peterson, 2022). This intervention primarily focusses on eliciting stories with specified story structure and grammar through the following tiered steps: active narration where a story is narrated by the examiner, reconstructing the story supplemented by visual aids such as icons for specific story elements all while receiving guidance, children retelling the story while decreasing the visual aids and finishing with a personal story generation based on the story practiced earlier (Spencer and Peterson, 2016). Story Champs™ includes 12 repeatable stories that are created based on common childhood themes such as injuring oneself, misplacing an object or having a quarrel with a sibling or peer, and that lay the base for intervention. Each story also has five major story grammar elements (Stein and Glenn, 1979; Spencer and Peterson, 2012) namely: character (individuals involved in the story), problem (initial event to begin the plot of a story), feeling (response to the problem), action (event following the problem in a story to initiate resolving the problem) and ending (resolution of the story).

There has been considerable research on the efficacy of Story Champs™ with several studies by the authors and colleagues reporting improvements in oral story retelling (Petersen et al., 2014; Spencer & Slocum, 2010; Spencer et al., 2013), generalization to personal story generation (Spencer & Slocum, 2010; Weddle et al., 2016), acquisition of new vocabulary in children with or without language impairments in oral storytelling, and formulation of significantly more coherent stories through the inclusion of more macrostructure elements (Spencer & Slocum, 2010; Spencer et al., 2013). Children as young as 4 years of age (Spencer et al, 2014; Spencer and Slocum, 2010) showed a significant improvement in their oral narratives in that they included more complex story grammar elements and formulated well defined oral narratives after receiving intervention using Story Champs™ for at least 15 minutes over a period of 3 weeks. Weddle et al. (2015) documented a similar effect on oral retell and personal story

generation using Story Champs™ in seven 3-and-4-year-old children. On receiving intervention, children's narrative retell skills and personal stories showed a significant improvement in terms of a higher inclusion of story grammar elements specifically macrostructure, similar to the previous studies mentioned.

The impact of oral narrative intervention on written language has been investigated in only a few studies, and all have been reported by the Story Champs' author group using Story Champs intervention. Oral narrative language intervention has been found to impact writing in children positively (Griffin et al., 2004; Kirby et al., 2021; Nelson et al., 2021; Peterson et al., 2022, Spencer et al., 2020; Spencer and Peterson 2018). A pilot study by Nelson et al. (2021) examined the impact of Story Champs™ on language and reading comprehension and writing in second and third grade Indian children. Participants were divided into three groups: a treatment group (Story Champs™ intervention), an alternate treatment group (whole-class shared story book lessons with greater stress on vocabulary instruction) and a no treatment group. Writing samples collected were scored based on the authors' Narrative Language Measure, which examines quality of content and complexity: for quality, story grammar was scored for characters, setting, emotion, plan, attempt, consequence, ending, ending emotion. Stories were also scored for episodes, language complexity, and the inclusion of conjunctions. Results indicated that participants from the treatment group showed significantly higher scores on all measures in question when compared to the alternate treatment and no-treatment group. For the writing samples, children who received the Story Champs™ intervention had better scores for their writing samples than those in the alternate treatment and no-treatment groups.

In a similar study, Peterson et al., (2022) examined the impact of Story Champs™ on oral and written language in kindergarteners. The participants were divided into three groups based on them receiving intervention in a large group, small group, or no intervention. The intervention groups completed six sessions over 2 weeks. Written samples were collected pre- and post-intervention and scored using the Narrative Language Measure. Results revealed a significant increase in story grammar and language complexity for the children from the large intervention group. Although there was a trend towards an

increase in story length, this increase was not significant. In another study examining effects for kindergarten children receiving 6 sessions of Story Champs intervention over 2 weeks in small groups, Kirby et al. (2021) observed significantly higher scores for story grammar, language complexity and episodes immediately after intervention in written stories. The scores for story grammar, language complexity and episodes continued to increase as the participants received more oral narrative instruction in the successive sessions over a period of 2 weeks.

The present study examined the impact of oral narrative instruction on oral and written language outcomes in a cross over trial of grade 1 children who either did or did not receive intervention. The Story Champs™ intervention program was provided to the whole class, and intervention effects were measured using an oral narrative measure administered pre and post intervention or control phases and written narratives collected throughout the study. The first aim was to examine the impact of the oral narrative intervention on oral language. Replication of positive effects on the re-telling of macrostructure elements was anticipated. In keeping with past research, use of microstructure elements was not expected to change. The second goal was to explore the impact of oral narrative intervention on written language outcomes. Specifically, both word and text level production were examined, although the intervention was considered to address the latter to a greater degree. It was hypothesized that text level productions, in particular, would improve with word level effects potentially not being observed.

Chapter 2

(ii) Methods

(i) Participants

Participants included 63 children (35 males, 28 females; mean age 6.6 years, SD 3.13, range 6.2 to 7) in four grade 1 classrooms, two classes in each of two schools (School 1: class 1, n= 17; class 2, n= 18; School 2: class 1, n= 15; class 2, n= 13). According to StatsCanada, the annual income of families of children from School 1 was $50,000 to $80,000 and that from School 2, $120,000 to $140,000 (Government of Canada, Statistics Canada, 2022). Each class was taught by one teacher (4 teachers total). English was reported to be the primary language for the majority of children (90%). For research purposes, only those children whose parents consented to being in the study were included in the current report. Of the children invited, 78% participated in the study.

(ii) Procedures

This study had a between groups, cross-over design in which one randomly chosen class in each school received the intervention in the first study phase and no intervention in the second phase of the study (i.e., acted as the control group) and the other class in the school was the control group in the first study phase and received intervention in the second study phase. In the intervention classrooms, the study speech-language pathologist provided twelve 25 to 30-minute whole class sessions over 3 weeks (intervention phase) using Spencer and Peterson's (2010) *Story Champs* curriculum while the teacher observed. Teachers then reinforced this learning throughout the week in their classrooms. In the non-intervention classrooms, teachers continued with their business-as-usual teaching as per their language arts curriculum. Outcome measures included assessments completed at 4 time points: pre-intervention 1, post-intervention 1, pre-intervention 2, and post-intervention 2, and writing probes completed throughout the study.

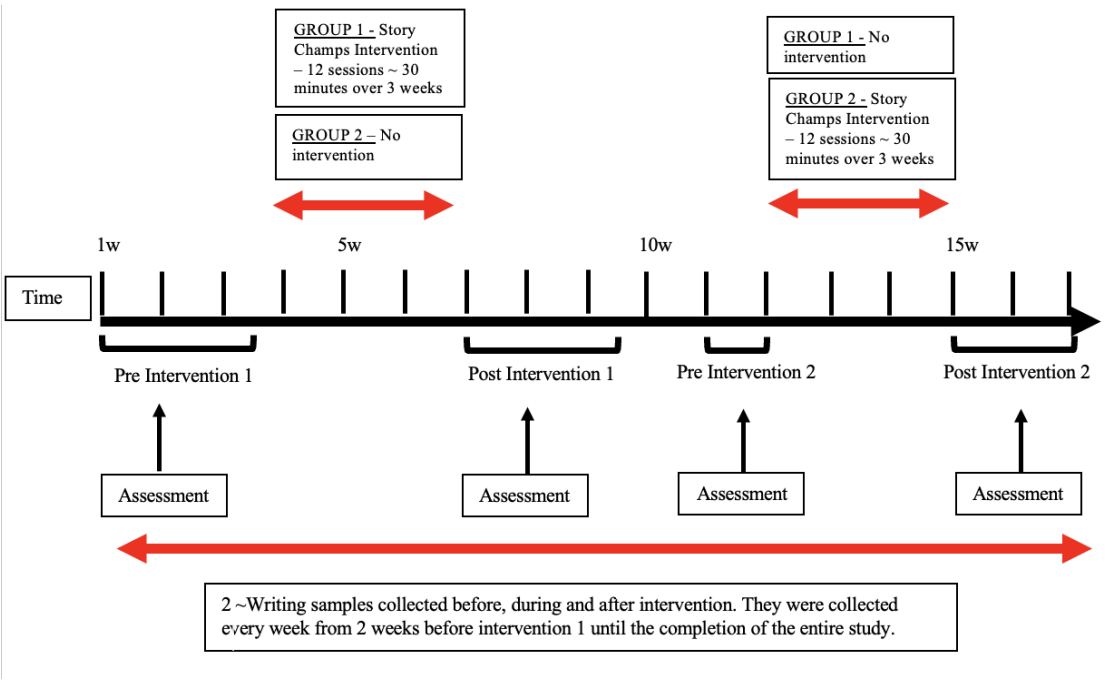


Figure 1:Study Timeline

Assessments were completed in one 30-minute individual session with each child in a quiet room in the school by a trained graduate student in speech-language pathology who was otherwise unfamiliar with the study. Assessment tasks included the CUBED *Narrative Language Measure (*NLM; Petersen & Spencer, 2016*)*, and the *Recalling Sentences* substest from the *Clinical Evaluation of Language Fundamentals*, 5th ed (CELF-5; Semel et al., 2013). In addition, teachers were asked to complete *writing probes* with the children twice weekly throughout the study. Writing probes were administered by the teacher as part of classroom activities.

(iii) Assessment tasks

CUBED Narrative Language Measure (NLM). Each child completed the NLM CUBED test (Petersen & Spencer, 2016) as a measure of narrative language. In this test, a short story is read to the child and the child is asked to retell the story, answer questions about the story, and generate a personal story related to the story just told. Each story is scored on Story Grammar (2 points maximum for 8 of the elements and 1 point maximum for 1 of the elements, total maximum score of 17), Language Complexity (total score is based on the total number of conjunctions used, maximum score of 9), and

Episode Complexity (total score of 5). These elements are summed to give a Listening Retell Score with a maximum score of 31. Raw scores were converted into scaled scores based on published norms (Peterson and Spencer, 2020). The NLM has multiple equivalent test forms available and the first four stories were employed with a different story used (in order) at each testing time point for all students. All test components were administered to all participants.

Recalling Sentences. The Recalling Sentences subtest of the CELF-5 (Semel et al., 2013) was completed by each child as a measure of expressive language. In this test, the child listens to spoken sentences of increasing length and complexity and is asked to repeat each sentence verbatim. A total of 26 sentences are presented. Scoring is based on a 4-point system (3 – no errors; 2 – 1 error; 1 – 2 or 3 errors; 0 - 4+ errors) giving a maximum raw score of 78. Errors considered are words changed, added, substituted, omitted, or change in the sequence of phrases containing 4 or more words. Raw scores are then converted to scaled scores.

(iv) Probes

Our goal was to collect writing samples throughout the study with a minimum of 2 samples per week in the month before intervention, during intervention, and the month after each intervention phase, and at least in alternate weeks in the remainder of the school year. These probes were administered by the classroom teacher as a whole class activity. Each student was given a lined paper and a photo story prompt and asked to ‘Please write a story on your own. Use this picture to write your story. Do the best you can’. No other prompts were given. Students were given a maximum of five minutes to complete their story. Photo story prompts were selected by the teachers in the study from pictures arising from a google search 'photo story prompts for young children'. All pictures were chosen at the teachers discretion. Teachers were asked to choose pictures with one activity and pictures that were not visually cluttered. Only photos available for free use were used. De-identified coded copies of written stories by students were transcribed. Written stories were scored based on their content only. In order to remove bias due to handwriting along with spelling, punctuation or capitalization errors, all samples were typed into their gloss by trained research assistants. A gloss is a corrected

version or summary of a passage in order to make it recognizable (Illinois Wesleyan Writing Center, 2021). Unintelligible words were defined as XXX. The study SLP reviewed all glosses, and any disagreements were discussed and resolved. The gloss samples were used for the following measures: Story grammar and Holistic measures. Whereas the original written stories were used for the quantitative measures in order to catch the spelling errors in particular.

The written samples were scored on story grammar quality and transcription quality following the written story scoring protocol of Peterson and Spencer (Peterson and Spencer, 2010). Briefly for text production quality, a holistic scores was derived based on the following elements, which were scored on a scale from 3 to 0: content/elements of Story Grammar (including all elements of a story such as character, feelings, action, problem and ending), organization (if story had a distinct beginning, middle and end to the flow), quality of response, transition or sequence words and descriptive words giving a maximum score of 15 for text production quality. The transcription quality on the other hand, assessed the writing probes on the basis of number of words written, correct letter sequences, words spelled correctly, and given as the following measures: (Kim et al., 2015): number of words spelled correctly (WSC), total number of words produced (TOTW), % words spelled correctly (PerWS), number of words at 5 minute mark (W5), total letters in 5 different longest words spelled correctly (LET), mean length of 5 different longest words spelled correctly (ML5).

A bespoke measure of overall substantive quality of writing samples was developed for this study based on the work of Harris et al. (2006). This quality was given a holistic score of 9 based on narratives, organization of words, sentence structure, punctuation, and capitalization. (Harris et al., 2006). The scoring team identified and selected written samples to represent anchor texts for low (scores ranging from 0-3), medium (scores ranging from 4-6) and high (scores ranging from 7-9) quality samples (Harris et al., 2006; Kim et al., 2015). Two trained research assistants scored all of the samples. The study SLP reviewed all scores. In cases of rater differences of 3 or more, the study SLP would provide a corrected score for that sample. Inter-rate reliability was 92% or greater for all writing sample measures.

(v) Intervention

Intervention involved the implementation of the manualized oral narrative instruction program, Story Champs (Spencer & Petersen, 2012). Briefly, 12 sessions of about 30 minutes were completed. Each session followed a set sequence involving modeling the story, playing story gestures, team retell, partner retell or partner personal story. For the first 6 lessons the children retold the story to each other, for the latter 6 lessons the children told a personal story related to the lesson story (e.g., going to the doctor, falling down, fighting with a brother or sister or friend). Materials for each session included story illustration cards (5), story grammar icons depicting character, setting, problem, feeling, action, ending and ending feeling and partner checklists ('champ check boards'). Every effort was made to involve all children in a class.

A minimum of two fidelity checks were completed during each intervention phase for each class using the relevant story champs lesson plan for the day. In all cases, there was 100% fidelity to the lesson.

ii.v.1 Modeling the story

With illustrations displayed, the study SLP told the story and placed the relevant icons next to the illustrations as the story unfolded. The study SLP then reviewed each icon by name, provided a definition and asked the children to repeat the name of each icon after her model and then name the icon as she pointed to each one.

ii.v.2 Playing Story Gestures

The study SLP repeated the story again and this time included a corresponding gesture along with each icon (see Table 1). Then the children were prompted to repeat the gestures along with the name of each icon. The study SLP then retold the story using the gestures and encouraged children to make the gestures along with her. The study SLP commented on how the story part connected to the gesture.

Story Grammar	Story Gesture
Character	Hand on Head
Setting	Two hands with fingertips together forming a roof for a house
Problem	Thumbs down
Feeling	Finger pointing to side of eye moving downward like a tear falling or right hand over the heart
Action	With right hand in a closed fist on left shoulder, move hand across body opening hand; you can do several times to make it stand out
Ending	Thumbs up
End Feeling	Both pointers at side of smiling mouth

Table 1: Story gestures used along with elements of story grammar

ii.v.3 Team Retell

During the team retell, the study SLP asked a question about each story grammar element. A child was selected to answer the question and then retell that part of the story with prompts if needed. The study SLP then modelled the answer for all students to repeat. Finally, the whole class repeated the sentence together. This process was completed for additional questions. At the end the study SLP summarized the entire story.

ii.v.4 Partner Retell

Finally, children worked in pairs to complete a partner retell. Within their pairs, one child retold the story while their partner used their ‘champ check board’ to

check off the elements of the story and then the children switched roles. The study SLP and teacher walked around and guided the children as needed.

ii.v.5 Partner Personal Retell

Still working in pairs, the children were then encouraged to tell their partner a personal story on a theme similar to the story of the day.

(vi) Data analysis

A preliminary analysis was planned to investigate potential baseline non-equivalence between order and school for each of the outcome measures. Separate 2 (School: 1, 2) by 2 (Order: Intervention-first; Control-first) one-way analysis of variances (ANOVAs) were completed on each outcome measure at time point 1. In cases of baseline differences based on School, school was retained as a variable in our analyses. In order to evaluate the effects intervention, an ANOVA was planned with School (1, 2) and order (Intervention-first; Control-first) as between group factors and intervention phase (time 1, time 2, time 3, time 4) as a within groups factor. Even though scaled scores are used for descriptive purposes, given the constricted age range, analyses were completed on raw scores for the outcome measures due to the greater variability in these scores.

Chapter 3

(iii) Results

(i) Oral Language Outcomes – Preliminary Analysis

Table 2 shows descriptive statistics for classes receiving intervention-first or second for each school and phase of the study for the oral outcome measures including Recalling Sentences and Listening Retell. In a preliminary analysis, baseline equivalence of these measures across schools and intervention orders was examined in separate 2 (school: 1, 2) by 2 (order: Intervention-first; Control-first) ANOVAs on the time 1 measures. For Recalling Sentences, there was a significant main effect of school, $F(1, 59) = 11.20$, $p = 0.05$, $n_p^2 = 0.016$, with School 1 (M = 38.95; SD = 12.42) having significantly lower scores than School 2 (M = 50.23, SD = 13.89). Neither the main effect of order, $F(1, 59) = 1.006$, $p = 0.32$, $n_p^2 = 0.017$, nor the interaction between school and order were significant, $F(1, 59) = 0.750$, $p = 0.39$, $n_p^2 = 0.013$. In all remaining analyses on the Recall Sentences data, school was retained as a variable.

For the Listening Retell, there was no significant main effect of school $F(1, 59) = 0.348$, $p = 0.55$, $n_p^2 = 0.006$, order $F(1, 59) = 0.030$, $p = 0.86$, $n_p^2 = 5.11 \times 10^{-4}$, nor the interaction effect between school by order $F(1, 59) = 0.451$, $p = 0.50$, $n_p^2 = 0.008$. In all remaining analyses of the Listening Recall data, school was not retained as a variable.

	Recalling Sentences Raw Scores								Recalling Sentences Scaled Scores							
Time Points	T1		T2		T3		T4		T1		T2		T3		T4	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Intervention-first Classes																
School 1 (n=17)	35.81	11.57	40.86	12.97	45.00	13.30	46.12	10.72	8.37	2.24	9.60	2.99	10.12	2.91	10.43	2.22
School 2 (n= 15)	50.00	13.78	51.26	16.43	57.14	17.84	58.40	17.41	11.20	2.93	11.53	3.46	12.71	3.64	12.86	3.46
Control-first Classes																
School 1 (n= 18)	42.10	13.28	45.50	10.13	48.57	13.22	47.61	16.72	9.21	2.95	10.27	1.77	10.63	2.60	10.33	3.64
School 2 (n= 13)	50.46	14.42	55.30	17.34	58.23	16.75	62.38	17.39	10.61	3.28	11.92	3.81	12.53	3.50	13.38	3.77

	Listening Retell Raw Scores								Listening Retell Scaled Scores							
Time Points	T1		T2		T3		T4		T1		T2		T3		T4	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Intervention-first Classes																
School 1 (n= 17)	8.00	4.91	12.75	4.96	13.31	4.31	17.37	4.91	5.68	2.30	8.06	2.83	8.50	2.30	11.25	3.94
School 2 (n= 15)	9.86	4.51	14.66	4.95	14.66	6.14	15.80	3.50	6.66	2.46	9.26	2.96	9.53	4.03	9.80	2.51
Control-first Classes																
School 1 (n= 18)	8.73	7.24	10.15	6.26	11.52	5.41	15.00	6.00	6.10	4.16	6.73	3.41	7.57	2.96	9.47	4.47
School 2 (n= 13)	8.61	5.62	11.76	5.21	10.61	5.51	15.84	2.99	5.92	2.66	7.69	2.81	7.07	2.78	9.76	2.04

Table 2: Raw Scores and Scaled scores mean (M) and standard deviation (SD) for Recalling Sentences and Listening Retell for each school and study phase

(ii) Oral Language Outcomes Across Study Phases

In order to assess intervention outcomes for Recalling Sentences, an ANOVA with school (1, 2) and intervention order (Intervention-first; Control-first) as between group factors and intervention phase (time 1, time 2, time 3, time 4) as a within groups factor was completed. The main effects of intervention phase, $F(3, 165) = 35.98$, $p < .001$, $n_p^2 = 0.39$, and school, $F(1, 55) = 0.16$, $p < .001$, $n_p^2 = 0.16$, were significant with the latter reflecting the lower scores of School 1 (M = 43.49, SD = 12.68) than School 2 (M = 55.75, SD = 16.44). All remaining main effects and all interactions effects were not significant: order, $F(1, 55) = 0.84$, $p = 0.36$, $n_p^2 = 0.01$; school x order, $F(1, 55) = 0.21$, $p = 0.21$, $n_p^2 = 0.01$; school x intervention phase $F(3, 165) = 1.77$, $p = 0.15$, $n_p^2 = 0.03$; order x intervention phase $F(3, 165) = 0.08$, $p = 0.96$, $n_p^2 = 0.01$; school x order x intervention phase $F(3, 165) = 1.37$, $p = 0.25$, $n_p^2 = 0.02$. For the main effect of intervention phase, pairwise post hoc comparisons revealed significant differences between all pairwise times ($p < .001$, all cases) except between times 3 and 4 ($p > .05$).

For Listening Retell, in order to assess intervention outcomes, an ANOVA with intervention order (Intervention-first; Control-first) as a between groups factor and intervention phase (time 1, time 2, time 3, time 4) as a within groups factor was completed. Only the main effect of intervention phase was significant $F(3, 177) = 31.23$, $p < .001$, $n_p^2 = 0.34$. The main effect of order, $F(1, 59) = 3.08$, $p = 0.08$, $n_p^2 = 0.05$, and the interaction effect of order by intervention phase, $F(3, 177) = 1.48$, $p = 0.22$, $n_p^2 = 0.02$, were not significant. For the main effect of intervention phase, pairwise post hoc comparisons revealed significant differences between all pairwise times ($p < 0.001$, all cases) except between times 2 and 3 ($p > .05$).

The results of the oral language measures show changes for both Recalling Sentences and Listening Recall between times 1 and 2 regardless of whether or not intervention was completed between these time points. Recalling Sentences improved between time 2 and 3 when no intervention was completed, and Listening Recall improved between time 3

and 4 regardless of intervention order. No oral language score changes were associated with intervention order.

(iii) Written Language Outcomes – Available data

Teachers were asked to allow for a 5-minute writing sample two times per week during the study, however the total number of samples collected for each classroom varied: School 1 intervention-first: group n= 217; control-first group: n = 62; School 2 intervention-first group: n= 151; control-first group: n = 114. Table 3 shows descriptive statistics for the number of written samples collected for each child in each classroom in each phase of the study. Generally, participants completed 1 to 2 written samples per intervention phase. Of the 16 possible time points that any of the 4 classes could complete written samples, there were 6 time points, 12 cases, and 11 unique children for which no written samples were submitted. The number of missing samples for each school, class and time point is as follows: School 1, Intervention-first, time 3, n=1; School 1, Control-first, time 1, n= 2, time 2, n= 3, time 3, n= 1, time 4, n=4; School 2, Control-first, time 4, n=1. Given the paucity of data for some classes, only the first 2 written samples (where available) in each phase were included in further analyses. Two samples were available for 76% of the participants. For the rest only the one written sample was considered except for those participants with no samples at given time points.

Intervention Phase	T1 (M(SD)) R	T2 (M(SD)) R	T3 (M(SD)) R	T4 (M(SD)) R
Intervention-first Classes				
School 1 (n= 17)	5.35 (1.11) 3-7	3.35 (0.86) 1-4	1.64 (0.70) 0-3	2.76 (0.56) 2-4
School 2 (n= 15)	3.80(0.41) 3-4	1.86(0.35) 1-2	1.93(0.25) 1-2	2.53(0.64) 1-3
Control-first Classes				
School 1 (n= 18)	0.88(0.32) 0-1	0.83(0.38) 0-1	0.94(0.23) 0-1	0.83(0.38) 0-1
School 2 (n= 13)	2.69(0.48) 2-3	2.69(0.48) 2-3	1.76(0.43) 1-2	1.61(0.65) 0-2

Table 3: Mean (M), standard deviation (SD) and Range (R) for the number of written samples collected from each participant in each class and school during each intervention phase

(iv) Written Language Outcomes – Preliminary data

Table 4 shows descriptive statistics for each class and phase of the study for the written language outcome measures including holistic score, number of words spelled correctly, total number of words produced , percent words spelled correctly , number of words at 5-minute mark, total letters in 5 different longest words spelled correctly, and mean length of 5 different longest words spelled correctly. In a preliminary analysis, baseline equivalence of these measures across schools and intervention order was examined in separate 2 (school: 1, 2) by 2 (intervention order: Intervention-first; Control-first) ANOVAs on the time 1 measures. For analyses of number of words spelled correctly, percent words spelled correctly, number of words at 5-minute mark, total letters in 5 different longest words spelled correctly, and mean length of 5 different longest words spelled correctly, the pattern of the results was the same: there was a main effect of school for all outcome measures, $F > 23$, $p < .05$ (all cases), but no significant main effect of intervention order, $F < 12$, $p > .05$ (all cases). As well, the interaction effect between school and order was not significant, $F < 2.5$, $p > .05$ (all cases). See Table 5 for full statistical results. For the Holistic Score and Total number of words produced there was a significant main effect of school, $F > 34$, $p < 0.001$, and a significant interaction effect of school and order, $F > 5.54$, $p < 0.05$, but no main effect of order, $F < 3$, $p > 0.05$. In all remaining analyses, School was retained as a variable.

Holistic Score								
Time Points	T1		T2		T3		T4	
	M	SD	M	SD	M	SD	M	SD
Intervention-first Class								
School 1	3.43	1.29	4.43	1.93	3.51	1.54	3.34	1.46

School 2	5.80	1.28	5.53	0.90	6.45	1.33	6.25	1.24
Control-first Class								
School 1	2.60	1.51	3.45	1.43	3.53	1.54	3.67	1.52
School 2	4.84	1.79	5.30	1.32	5.19	2.22	5.25	1.18
Number Of Words Spelled Correctly								
Intervention-first Class								
School 1	4.51	2.85	10.85	7.45	6.23	4.14	5.29	4.56
School 2	19.15	11.36	18.60	8.88	30.28	13.08	31.67	16.88
Control-first Class								
School 1	2.75	3.75	4.92	3.44	6.19	5.54	6.69	5.40
School 2	13.32	9.48	15.96	8.65	19.48	12.68	18.02	11.87
Total Number Of Words Produced								
Intervention-first Class								
School 1	11.35	4.98	23.15	13.21	14.13	7.90	12.09	7.42
School 2	32.88	18.64	27.44	12.07	42.45	15.79	44.73	21.28
Control-first Class								
School 1	5.98	4.81	9.95	8.22	11.05	9.35	12.44	9.64
School 2	18.92	11.88	19.80	11.47	23.51	14.64	21.85	13.53
% Words Spelled Correctly								
Intervention-first Class								
School 1	35.68	17.43	44.46	18.03	40.51	16.87	36.67	20.98
School 2	59.05	16.20	68.64	14.81	70.43	14.48	70.77	16.64
Control-first Class								
School 1	40.84	30.21	58.80	34.10	56.35	27.71	56.87	19.04
School 2	68.94	12.93	82.59	13.10	79.38	16.36	81.17	10.77
Number Of Words At 5-Minute Mark								

				Intervention-first Class				
School 1	11.15	4.87	16.84	10.57	14.13	7.90	12.09	7.42
School 2	26.58	14.10	26.67	12.25	37.43	16.79	40.83	20.29
				Control-first Class				
School 1	5.98	4.81	9.89	7.75	11.05	9.35	10.66	8.24
School 2	18.92	11.88	18.76	10.49	23.51	14.68	21.85	13.53
				Total Letters In 5 Different Longest Words Spelled Correctly				
				Intervention-first Class				
School 1	9.39	6.15	14.51	7.55	10.61	5.77	10.10	7.08
School 2	18.01	3.88	18.92	3.66	24.55	5.81	22.80	4.47
				Control-first Class				
School 1	5.57	7.12	8.70	5.87	12.46	8.43	10.03	6.86
School 2	19.01	7.14	22.51	4.15	21.73	7.33	22.22	5.90
				Mean Length Of 5 Different Longest Words Spelled Correctly				
				Intervention-first Class				
School 1	1.87	1.23	2.90	1.51	2.12	1.15	2.02	1.41
School 2	3.60	0.77	3.78	0.73	4.91	1.16	4.56	0.89
				Control-first Class				
School 1	1.11	1.42	1.74	1.17	2.49	1.68	2.12	1.32
School 2	3.80	1.42	4.50	0.83	4.34	1.46	4.44	1.18

Table 4: Mean (M) and standard deviation (SD) for Holistic Score, number of words spelled correctly, total number of words produced, % words spelled correctly, number of words at 5-minute mark, total letters in 5 different longest words spelled correctly, mean

Written Language Outcomes	Main effect of school F^1, p, n_p^2	Main effect of order F^1, p, n_p^2	Interaction effect of school x order F^1, p, n_p^2
Holistic Score	36.79, <.001, 0.39*	0.02, 0.876, $4.39x10^{-4}$	5.54, 0.22, 0.08
Number Of Words Spelled Correctly	42.29, <.001, 0.42*	1.09, 0.30, 0.01	3.83, 0.505, 0.06
Total Number Of Words Produced	34.98, <.001, 0.38*	2.17, 0.146, 0.03	11.01, 0.002, 0.16*
% Words Spelled Correctly	23.31^2, <.001, 0.30*	^{1}0.19, 0.659, 0.01	1.99^2, 0.164, 0.03
Number Of Words At 5-Minute Mark	33.24, <.001, 0.36*	0.25, 0.615, 0.01	6.80, 0.012, 0.10*
Total Letters In 5 Different Longest Words Spelled Correctly	47.59,<.001, 0.45*	2.26, 0.137, 0.03	0.77, 0.383, 0.01
Mean Length Of 5 Different Longest Words Spelled Correctly	47.59, <.001,0.45*	2.26, 0.137, 0.03	0.77, 0.383, 0.01

Note: 1 - degrees of freedom of 1 and 57; 2 - degrees of freedom of 1 and 54; * indicates significant result

Table 5: Statistics for main effect of school, order and interaction effect of school x order for all written language outcome measures

(v) Written Language Outcomes Across Study Phases

In order to assess intervention outcomes, an ANOVA with school (1, 2) and intervention order (Intervention-first; Control-first) as between group factors and intervention phase (time 1, time 2, time 3, time 4) as a within groups factor was completed on each of the written language measures.

For the Holistic Score, the main effect of school, $F(1,48) = 42.71$, $p < 0.001$, $n_p^2 = 0.47$, was significant due to the higher scores of School 2. The interaction effect between school, order and intervention phase, $F(3,144) = 4.99$, $p = 0.003$, $n_p^2 = 0.09$, was also significant (see figure 2). All remaining effects were not significant: intervention phase, $F(3,144) = 2.11$, $p = 0.101$, $n_p^2 = 0.04$; order, $F(1,48) = 2.98$, $p = 0.90$, $n_p^2 = 0.05$; school x order $F(1, 48) = 0.26$, $p = 0.606$, $n_p^2 = 0.01$; school x intervention phase $F(3, 144) = 2.10$, $p = 0.103$, $n_p^2 = 0.04$; order x intervention phase $F(3, 144) = 0.390$, $p = 0.760$, $n_p^2 = 0.008$. In order to unpack the interaction, pairwise comparisons were completed. Potential intervention effects were examined by considering differences between order groups across intervention phases within schools. Within schools, the intervention-first and

control-first groups did not differ significantly except for time 3 in School 2 ($p < .05$). In fact, there were no significant within group effects across time points for both schools and order ($p > .05$, all cases). Additional comparisons examined school differences based on intervention order and phase (see Figure 2). Across schools for the intervention-first classes, scores were significantly lower ($p < .05$) for School 1 than School 2 at all time points except time 2, where no significant differences were found ($p > .05$). When comparing across schools for the control-first classes, scores were significantly lower for School 1 than School 2 at time 1 (all pairs), time 2 (at time 1, 3, 4, but not time 2), and not significantly different at time 3 and time 4.

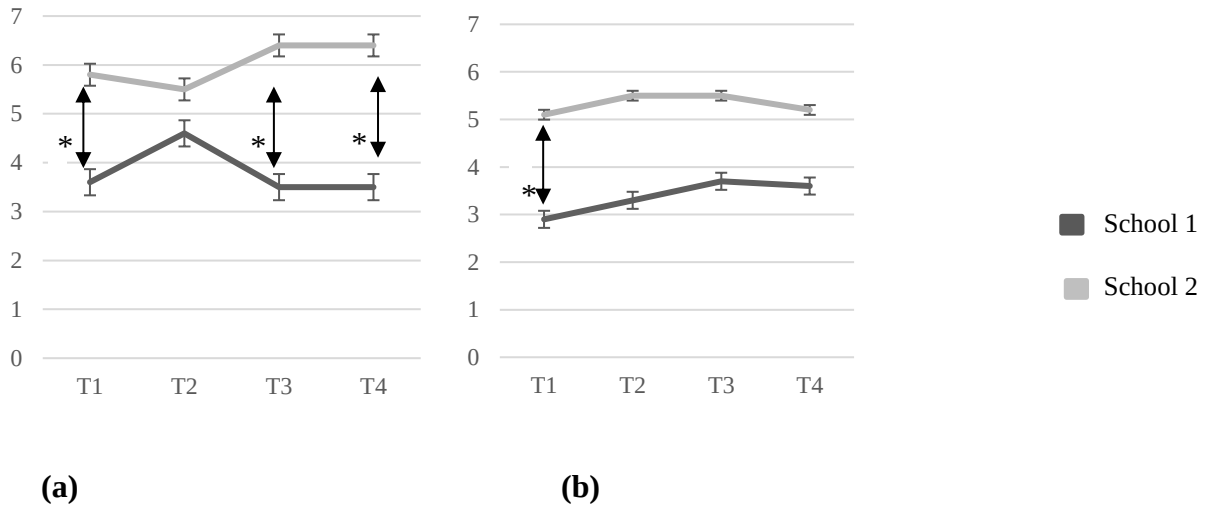


Figure 2: Significant intervention phase differences between schools for the (a) Intervention-first and (b) Control-first classes (only same-phase differences are shown)

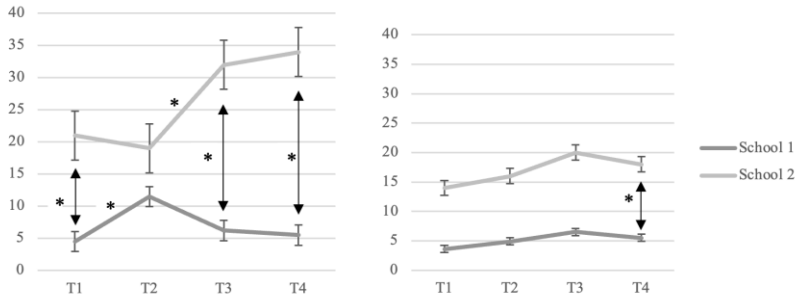
For Number of Words Spelled Correctly, all main effects were significant: intervention phase, $F(3, 144) = 13.37$, $p < 0.001$, $n_p^2 = 0.21$, for which pairwise post hoc comparisons revealed significant differences between time 1 and 2, and 2 and 3, ($p < 0.001$, all cases) but not 3 and 4 ($p > .05$), school, $F(1,48) = 56.09$, $p < 0.001$ $n_p^2 = 0.53$, due to the higher scores of School 2 and order, $F(1,48) = 6.29$, $p = 0.016$, $n_p^2 = 0.11$, due to the higher scores of those who received intervention-first. These main effects were modified by significant interactions between school and intervention phase, $F(3,144) = 13.59$, $p < 0.001$ $n_p^2 = 0.22$, and between school, order, and intervention phase, $F(3,144) = 9.55$, $p <$

0.001, n_p^2= 0.16 (see figure 3). Remaining effects were not significant: school x order, $F(1, 48) = 2.73$, $p = 0.105$, n_p^2= 0.05; order x intervention phase $F(3,144) = 1.55$, $p = 0.203$, n_p^2= 0.03. Coming to the significant interaction between school and intervention phase, pairwise comparisons revealed significant within group effects of intervention phase for School 1 between time 1 and 2 and for School 2 between time 2 and 3, and time 3 and 4 ($p < 0.05$, all cases). To unpack the significant three-way interaction of school, order and interaction phase, pairwise comparisons were examined. Investigations of order and intervention phase effects within schools revealed a significant effect of time from time 1 to 2 for the intervention-first group in School 1, and time 2 to 3 and 3 to 4 in the intervention-first group in School 2 ($p < 0.05$, all cases), the remaining pairwise comparisons between time points showed no significant effects ($p > 0.05$). Within schools, the intervention-first and control-first groups did not differ significantly except for time 4 in School 2 ($p < .05$). School differences were explored further. Potential intervention effects were examined by considering differences between order groups across intervention phases within schools and showed that for the intervention-first classes, scores were significantly lower ($p < .05$) for School 1 than School 2 at all time points except time 2 (all pairs), where no significant difference was found ($p > .05$). When comparing across schools for the control-first classes, scores were significantly lower for School 1 compared to School 2 at time 1 compared to time 2, 3, and 4 (but not at time 1), at time 2 compared to time 1, 3 and 4 (but not time 2), and at time 3 compared to time 1, 2, and 3 (but not time 4; the time 4 points also did not differ).

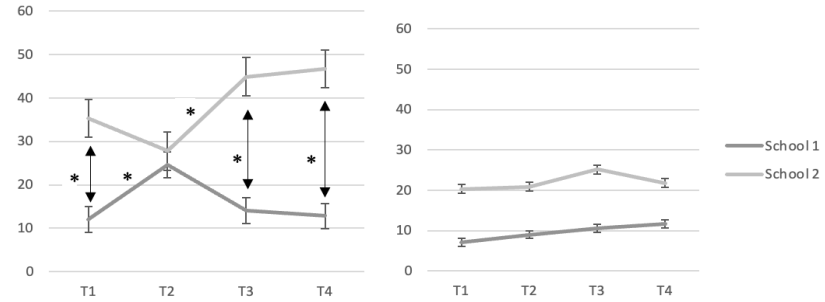
These same patterns of significant effects were observed for Total Number of Words Produced, Number Of Words At the 5-Minute Mark, Total Letters In 5 Different Longest Words Spelled Correctly, Mean Length Of 5 Different Longest Words Spelled Correctly. The same patterns in pairwise comparisons were observed for the Intervention-first classes across measures. The pairwise comparison patterns were similar for the Control-first classes with no significant School differences between measures except in the case of total letters in 5 different longest words spelled correctly, for which differences occurred at each timepoint. Figure 3 shows the interaction effect (and paired comparisons) for all mentioned measures and table 6 summarizes the statistics for these measures.

(i) Number of Words Spelled Correctly

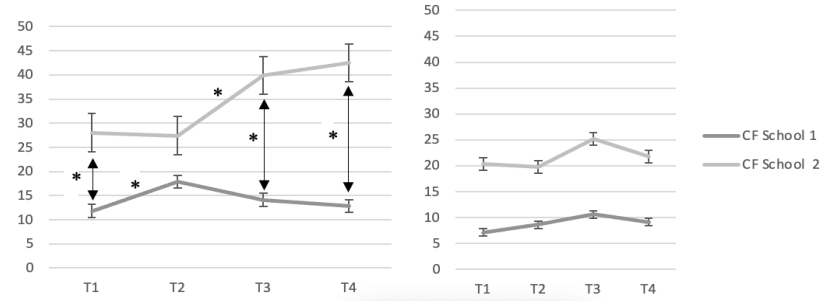
(a) Intervention-first **(b) Control-first**



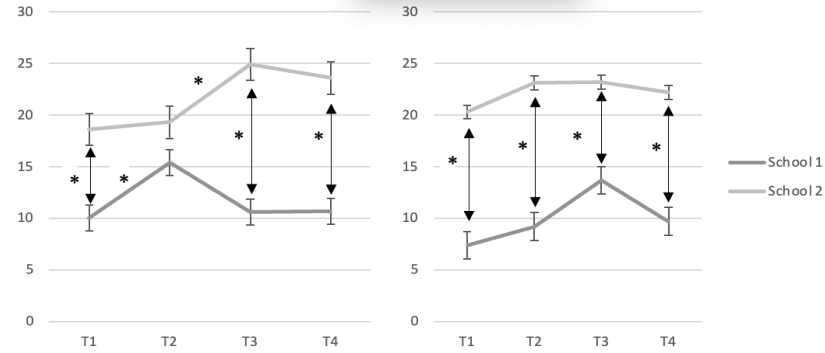
(ii) Total Number of Words Produced



(iii) Number Of Words At the 5-Minute Mark



(iv) Total Letters In 5 Different Horizontal (Category) Axis **Correctly**



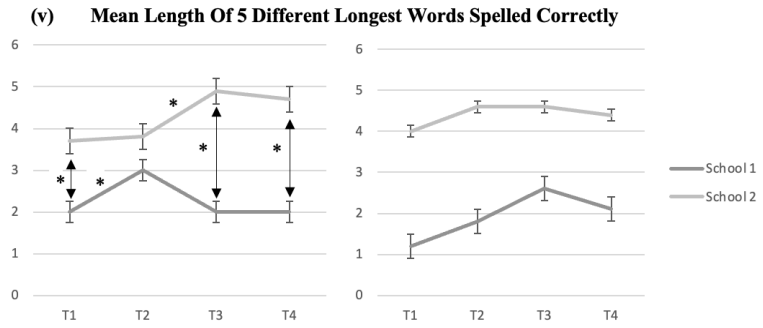


Figure 3: Interaction effects for (i) Number of Words Spelled Correctly, (ii) Total Number of Words Produced, (iii) Number Of Words At the 5-Minute Mark, (iv) Total Letters In 5 Different Longest Words Spelled Correctly, (v) Mean Length Of 5 Different Longest Words Spelled Correctly of school by order by intervention phase for (a) Intervention-first and (b) Control-first (*$p < .05$)

	school x order F^1(df), p, n_p^2	order x intervention phase F^3(df), p, n_p^2	school x intervention phase F^3(df), p, n_p^2	school x intervention phase x order F^3(df), p, n_p^2
Number Of Words Spelled Correctly	2.73, 0.105, 0.05	1.55, 0.203, 0.03	13.59, < 0.001, 0.22*	13.59, < 0.001, 0.22*
Total Number Of Words Produced	3.57, 0.065, 0.06	0.34, 0.793, 0.01	10.68, < 0.001, 0.18 *	10.53, < 0.001, 0.18*
Number Of Words At 5-Minute Mark	1.92, 0.172, 0.039	1.78, 0.153, 0.036	7.47, < 0.001, 0.135*	5.12, 0.002, 0.096*
Total Letters In 5 Different Longest Words Spelled Correctly	0.83, 0.365, 0.017	0.692, 0.558, 0.014	2.19, 0.091, 0.044	7.95, < 0.001, 0.142*
Mean Length Of 5 Different Longest Words Spelled Correctly	3.47^2, 0.312, 0.022	0.52^4, 0.664, 0.011	2.30^4, 0.079, 0.047	9.27^4, < 0.001, 0.165*

Note: 1 - degrees of freedom of 1 and 48; 2 - degrees of freedom of 1 and 47; 3 - degrees of freedom of 3 and 144; 4 - degrees of freedom of 3 and 141 * indicates significant result

Table 6: Interaction effects for the written language measures

In the case of Words Spelled Correctly, all main effects were significant: intervention phase, $F(3,141) = 7.30$, $p < 0.001$, $n_p^2 = 0.13$, due to differences between time 1 and 2 ($p < 0.05$, all cases) but not 2 and 3, or 3 and 4 ($p > .05$), school, $F(1,47) = 43.03$, $p < 0.001$, $n_p^2 = 0.47$, and order, $F(1,47) = 12.45$, $p < 0.001$, $n_p^2 = 0.20$. There were no significant interactions: school x order $F(1, 47) = 0.36$, $p = 0.546$, $n_p^2 = 0.01$; school x intervention phase $F(3, 141) = 0.96$, $p = 0.413$, $n_p^2 = 0.02$; order x intervention phase $F(3, 141) = 0.75$, $p = 0.521$, $n_p^2 = 0.01$; school x order x intervention phase $F(3, 141) = 0.465$, $p = 0.707$, $n_p^2 = 0.01$. Main effects patterns were the same as for the previous analyses for school and order.

The results of the written language measures show changes for the intervention-first class in School 1 between time 1 and time 2 and between time 2 and time 3 for School 2 for 5 out of 7 of the written outcome measures. The control-first groups who received intervention second show no changes between time 3 and 4. The results also showed consistently higher scores for School 2 overall. No written language scores changes were associated with intervention order.

Chapter 4

(iv) Discussion

The purpose of the current study was to examine the influence of oral narrative instruction on oral and written language outcomes in kindergarten children when they either did or did not receive a whole class narrative intervention, Story Champs™. For the oral language measures, significant increases in Listening Retell were observed between time 1 and time 2 and time 3 and time 4 for all children irrespective of intervention and for Recalling Sentences between time 2 and time 3 and time 3 and 4. Notably, scores were higher for School 2 than School 1 regardless of intervention phase or order for only the Recalling Sentences but not Listening Retell oral language measures. With regards to written language measures, text production quality scored holistically based on story grammar macro- and microstructure revealed higher scores for School 2 than 1 intervention classes at all time points except time 2, and control classes at times 1 and 2. No individual groups made significant changes across study phases. Across five word/transcription level production measures, the pattern of responses was the same for 4 of the 5, Numbers Of Words Spelled Correctly, Total Number Of Words Produced, Percent Words Spelled Correctly and Number Of Words At 5-Minute Mark. Overall, the intervention-first groups had higher scores, but these scores occurred in different phases across schools: In School 1, the intervention-first class showed a significant increase from time 1 to 2 whereas in School 2, the intervention-first group showed a significant increase from time 2 to 3 and 3 to 4. Importantly, intervention and control groups within Schools did not differ at any study phase except for higher scores by the intervention-first group at time 4 in School 2. School 2 scores were significantly higher with the exception of time 2 for the intervention-first classes and times 3 and 4 for control-first classes. For the mean length of the 5 different longest words spelled correctly, there were significantly higher scores for the intervention-first classes, changes between times 1 and 2, and School 2.

(i) Oral language outcomes

With regards to the first hypothesis concerning narrative intervention and oral language outcomes, this study failed to provide convincing evidence of an increase in oral language scores following narrative intervention. There were no significant increases associated with intervention order for either the Listening Recall or the Recalling Sentences measure. The findings are surprising given the number of studies that have shown this effect in the past (Gilliam et al., 2024; Kirby et al., 2021; Nelson et al., 2021; Petersen et al., 2014; Spencer & Slocum, 2010; Spencer et al., 2013). The current study used the Listening Recall measure created and used by Spencer, Peterson and colleagues across a number of studies including studies with a similar sample size (Kirby et al., 2021; Nelson et al., 2021; Peterson et al., 2021) and comparable large group intervention settings (Peterson et al., 2021; Peterson et al., 2020; Spencer et al., 2015). Despite no differences associated with the intervention order, children did improve on the Listening Recall measure over repetitions of the task with differences observed between administrations at time 1 and 2, and time 3 and 4.

A similar lack of a difference associated with intervention order was also observed in the other oral language measure included in the study, Recalling Sentences. On this measure. improved scores were observed for children between time 1 and 2, and time 2 and 3. Although an oral language measure, Recalling Sentences is somewhat different from the intervention as the focus of said intervention was on narratives, which may make the lack of a difference in results less surprising. Nevertheless, Recalling Sentences has been used occasionally (Spencer et al., 2020) to assess oral language outcomes in children following the oral narrative intervention and significant differences observed in children receiving intervention.

Thus, despite the intervention, no significant changes were associated with the order of intervention for both Recalling Sentences and Listening Retell. This finding shows that the intervention's timing (whether implemented first or second) does not significantly influence the improvement in oral language skills. All children improved on the oral language measures at different time points and there was some indication of school

difference for Recalling Sentences. The potential influence of these factors will be discussed below.

(ii) Written language outcomes

With respect to the second hypothesis concerning the impact of oral narrative language instruction on written language measures, this study again failed to provide substantial evidence of an increase in any of the written language scores following oral narrative intervention. The holistic measure was based on the Spencer and Peterson's (2016) Narrative Language Measure, which has been used in a variety of other studies (Kirby et al., 2021; Nelson et al., 2021; Peterson et al., 2020; Peterson et al., 2021; Spencer and Peterson. 2018) and is the most closely related written language measure to the intervention provided in this study. The current findings revealed no significant holistic score increases associated with the order of intervention. The only pattern of interest with regards to the intervention was the nonsignificant difference between School 1 and 2 at time 2 for the intervention-first classes where these groups differed at all other time points. The lack of significant changes related to the intervention in the holistic score differs from findings of previous studies showing a significant improvement in grade 1 children (Spencer and Peterson, 2018), grade 2 children (Peterson et al., 2020) and grade 3 children (Nelson et al., 2021) using Story Champs™ as the intervention for oral narrative language instruction. Writing samples taken from all children in the above studies showed longer sentences, inclusion of more story grammar elements, and episodic levels.

This study additionally included several written word level measures: numbers of words spelled correctly, total number of words produced, percent words spelled correctly, number of words at 5-minute mark, mean length of the 5 different longest words spelled correctly. These word measures are best aligned with the transcription component of the Not-So-Simple view of writing (Berninger and Graham, 1998) that requires both, idea generation and decoding or transcription skills. Arguably, these skills were not directly addressed in the oral narrative intervention provided in this study. Perhaps not surprisingly, then, there were no clear intervention effects. The current study is the first to examine these written word level effects in response to oral narrative intervention or oral

language intervention more generally. Previous work has focused on writing intervention with systematic reviews revealing positive outcomes with a focus on either transcription (Datchuk and Kubina, 2012; Graham et al., 2012) or idea generation (McMaster et al., 2017; Pennington et al., 2012).

Notably, for the majority of word-level written measures, significant changes were observed for the intervention-first groups, but time points differed across schools with School 1 showing a change between time 1 and 2 and School 2 between time 2 and 3 and 3 and 4. With regards to these patterns across schools, one possibility is that the baseline differences across schools masked a more consistent intervention effect. For example, School 1 had a lower baseline score and thus a greater potential for score increase, which may have resulted in an early response to the intervention. School 2, on the other hand, started with higher scores and responded later. Perhaps then, the lack of response in the School 2 control-first group at time 4 (after they had received intervention) is because their skills were not monitored for a sufficiently long enough period following intervention.

Because there were little to no improvements in written language outcomes post-intervention, we could conclude that oral language narrative instruction may not impact word-level and sentence-level structures for written narratives. This finding is not unexpected for the word-level findings given that the intervention did not specifically address these skills. The sentence-level, holistic results are more surprising given the alignment of this measure with the narrative intervention. Nevertheless, oral language narrative instruction does not explicitly instruct children to formulate complex oral and written narratives (Kirby et al., 2021; Spencer et al., 2013), which could account for the lack of change on the sentence-level written measure.

(iii) School differences

One challenge in the present study was the consistently higher scores of School 2 than School 1. Although the scaled scores of the majority of children in the study were within one standard deviation of the standardized mean (i.e., scaled score lower than 7), the children in School 1 scored lower overall and had a higher percentage of children with a

scaled score lower than 7 on one of the oral language measures (School 1: 36%; School 2: 27%). These school differences could reflect the differences in the socioeconomic regions within which each school was situated as reflected by the higher average annual income of household in the School 2 region (Government of Canada, Statistics Canada, 2022). Children from families with a higher socioeconomic status have been shown to perform better academically (Suna, 2020), in learning tasks (Lareau, 2011). They have also shown to use more complex vocabulary and sentence structures (Sirin, 2005).

There are two ways in which the differences in socioeconomic status for the two schools could have influenced the pattern of results in the current study. Specifically, as argued above, the lower baseline scores of School 1than School 2 could have allowed School 1 to have more room for change resulting in higher scores immediately following intervention. It must be noted, however, that the suggestion of an intervention response by the lower and not higher socioeconomic group in the present study would not be consistent with the well documented Matthew Effect describing the pattern of those with higher academically related skills benefiting more from interventions (Walberg & Tsai, 1983).

(iv) Repeated testing

An additional pattern across a number of the measures in the present study (Listening Retell, word-level and text-level measures) was the increase in score regardless of intervention. It is well known that repeated testing can result in increases (Karpicke & Roedigeriii, 2007; Killeen, 2005; Metcalfe & Kornell, 2003). Generally, it is recommended that the same version of a standardized test not be administered more frequently than 4 weeks (Killeen, 2005). Many standardized tests publish parallel versions for this purpose. In the present study, the same oral language measures were administered at each time point with increases observed generally between time 1 and 2. These findings could be attributed to a testing effect. It is observed that children use repeated retrieval or the ability to recall previously learnt information to perform better on repeated outcome measures (Karpicke & Roedigeriii, 2007). Considering the current study, repeatedly performing the tasks could have led to the children learning the tasks

well enough. This possible overlearning (Postman, 1962) may have helped the children augment their long-term retention of the tasks and perform better.

However, the changes between times 1 and 2 for the written language measures are not as readily interpreted as repeated measures effects given that children were given different written prompts for each sample.

(v) Limitations

Several study limitations must be considered. One limitation was the small sample size. Although there were 63 participants in the study in total, they were spread across four classrooms, and the intervention was administered to the whole class. As a result, only two instantiations of the intervention were completed. It is possible that involving more classrooms could have revealed intervention effects. Also, the current study had only measure to assess transcription quality. There were several instances where children did not complete written samples across the study, as planned. As a result, the number of written samples considered for analyses had to be limited to a maximum of two per test time. Future research could focus more on working closely with the teachers to collect more samples. Additionally, there were significant baseline differences between the schools. A larger sample size could aid in better understanding the efficacy oral language narrative intervention especially as it relates to changes in written language skills. Additionally, using the same oral language measures could have led to benefits due to repeated retrieval across the study. Future research must include different parallel test versions to avoid repeated testing.

Chapter 5

(v) Conclusion

This study examined the impact of an oral language narrative intervention on oral and written language in grade 1 children. Contrary to a considerable amount of research evidence (Gilliam et al., 2024; Kirby et al., 2021; Nelson et al., 2021; Petersen et al., 2014; Peterson et al., 2020; Peterson et al., 2021; Spencer and Peterson. 2018; Spencer & Slocum, 2010; Spencer et al., 2013), positive effects of the intervention on oral and written language were not observed. Despite high fidelity to the published oral narrative intervention, Story Champs™, intervention effects on an oral narrative measure, a recalling sentences measure, or word- or sentence-level written measures were not observed. The lack of effects could be related to the small sample size, baseline differences between the schools, and a small number of writing samples for analysis. There was some indication of different intervention responses across schools, which could suggest that factors such as specific school context and baseline student abilities influenced the outcome. By understanding and addressing the unique needs of each school, educators may be able to enhance the effectiveness of interventions and promote better language outcomes for all students. Further research examining narrative interventions in relation to individual student and classroom-level factors is warranted.

References

Bishop, D. V. M., & Edmundson, A. (1987). Language-Impaired 4-Year-Olds. *Journal of Speech and Hearing Disorders*, *52*(2), 156–173. https://doi.org/10.1044/jshd.5202.156

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 and Allied Disciplines*, *31*(7), 1027–1050. https://doi.org/10.1111/j.1469-7610.1990.tb00844.x

Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology*, *101*(2), 294–304. https://doi.org/10.1037/a0013844

Botting, N., Simkin, Z., & Conti-Ramsden, G. (2006). Associated Reading Skills in Children with a History of Specific Language Impairment (SLI). *Reading & Writing*, *19*(1), 77–98. https://doi.org/10.1007/s11145-005-4322-4

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*(1), 371-399. https://doi.org/10.1146/annurev.psych.53.100901.135233

Catts, H. W., Adolf, S. M., & Ellis Weismer, S. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. Journal of Speech, Language, and Hearing Research, 49(2), 278–293.

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating Children's Reading-Comprehension difficulties. *Psychological Science*, *21*(8), 1106–1116. https://doi.org/10.1177/0956797610375449

Conger, R. D., & Donnellan, M. B. (2007). An interactionist perspective on the socioeconomic context of human development. *Annual Review of Psychology*, *58*(1), 175–199. https://doi.org/10.1146/annurev.psych.58.110405.085551

Dickinson, D. K., & Snow, C. E. (1987). Interrelationships among prereading and oral language skills in kindergartners from two social classes. *Early Childhood Research Quarterly*, *2*(1), 1–25. https://doi.org/10.1016/0885-2006(87)90010-x

Gillam, S. L., Gillam, R. B., Magimairaj, B. M., Capin, P., Israelsen-Augenstein, M., Roberts, G., & Vaughn, S. (2024). Contextualized, multicomponent language instruction: from theory to randomized controlled trial. *Language, Speech & Hearing Services in Schools*, *55*(3), 661–682. https://doi.org/10.1044/2024_lshss-23-00171

Griffin, T. M., Hemphill, L., Camp, L., & Wolf, D. P. (2004). Oral discourse in the preschool years and later literacy skills. *First Language*, *24*(2), 123–147. https://doi.org/10.1177/0142723704042369

Hatcher, P. J., Hulme, C., Miles, J. N., Carroll, J. M., Hatcher, J., Gibbs, S., Smith, G., Bowyer-Crane, C., & Snowling, M. J. (2005). Efficacy of small group reading intervention for beginning readers with reading-delay: a randomised controlled trial. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*(8), 820–827. https://doi.org/10.1111/j.1469-7610.2005.01559.x

Hayward, D., & Schneider, P. (2000). Effectiveness of teaching story grammar knowledge to pre-school children with language impairment. An exploratory study. *Child Language Teaching and Therapy*, *16*(3), 255–284. https://doi.org/10.1177/026565900001600303

Karpicke, J., & Roediger‌iii, H. (2007). Repeated retrieval during learning is the key to long-term retention☆. *Journal of Memory and Language*, *57*(2), 151–162. https://doi.org/10.1016/j.jml.2006.09.004

Kemper, S. (1984). The development of narrative skills: explanations and entertainments. In *Springer series in cognitive development* (pp. 99–124). https://doi.org/10.1007/978-1-4613-9508-9_5

Killeen, P. R. (2005). An alternative to null-hypothesis significance tests. *Psychological science*, *16*(5), 345-353.

Kim, Y. G., Park, C., & Park, Y. (2015). Dimensions of discourse level oral language skills and their relation to reading comprehension and written composition: an exploratory study. *Reading & Writing*, *28*(5), 633–654. https://doi.org/10.1007/s11145-015-9542-7

Kirby, M. S., Spencer, T. D., & Chen, Y. I. (2021). Oral narrative instruction improves kindergarten writing. *Reading & Writing Quarterly*, *37*(6), 574–591. https://doi.org/10.1080/10573569.2021.1879696

Lareau, A. (2011). *Unequal childhoods: Class, race, and family life* (2nd ed.). University of California Press.

Manhardt, J., & Rescorla, L. (2002). Oral narrative skills of late talkers at ages 8 and 9. *Applied Psycholinguistics*, *23*(1), 1–21. doi:10.1017/S0142716402000012

Metcalfe, J., & Kornell, N. (2003). The dynamics of learning and allocation of study time to a region of proximal learning. *Journal of Experimental Psychology General*, *132*(4), 530–542. https://doi.org/10.1037/0096-3445.132.4.530

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. https://doi.org/10.1111/j.1467-9817.2004.00238.x

Nelson, B. S., Petersen, D. B., & Rai, A. (2021). The effects of a multi-tiered system of language support on oral narrative language, writing, and reading comprehension in India. *Language and Education*, *36*(1), 74–94. https://doi.org/10.1080/09500782.2021.1898633

Paris, A. H., & Paris, S. G. (2003). Assessing narrative comprehension in young children. *Reading Research Quarterly*, *38*(1), 36–76. https://doi.org/10.1598/rrq.38.1.3

Petersen, D. B., & Spencer, T. D. (2016). Using narrative intervention to accelerate canonical story grammar and complex language growth in culturally diverse preschoolers. *Topics in Language Disorders*, *36*(1), 6–19. https://doi.org/10.1097/tld.0000000000000078

Petersen, D. B., Mesquita, M. W., Spencer, T. D., & Waldron, J. (2020). Examining the effects of multitiered oral narrative language instruction on reading comprehension and writing. *Topics in Language Disorders*, *40*(4), E25–E39. https://doi.org/10.1097/tld.0000000000000227

Sedita, J. (2024, June 3). *The not so simple view of writing*. Keys to Literacy. https://keystoliteracy.com/blog/the-not-so-simple-view-of-writing/

Silva, M., & Cain, K. (2015). The relations between lower and higher level comprehension skills and their role in prediction of early reading comprehension. *Journal of Educational Psychology*, *107*(2), 321–331. https://doi.org/10.1037/a0037769

Silva, M., & Cain, K. (2017a). The use of questions to scaffold narrative coherence and cohesion. *Journal of Research in Reading*, *42*(1), 1–17. https://doi.org/10.1111/1467-9817.12129

Singh, K. (2014). Right to education and equality of educational opportunities. *Journal of International Cooperation in Education*, 16(2), 5–19.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*(3), 417-453. https://doi.org/10.3102/00346543075003417

Spencer, T. D., & Slocum, T. A. (2010). The effect of a narrative intervention on story retelling and personal story generation skills of preschoolers with risk factors and narrative language delays. *Journal of Early Intervention*, *32*(3), 178–199. https://doi.org/10.1177/1053815110379124

Spencer, T. D., Petersen, D. B., Slocum, T. A., & Allen, M. M. (2014). Large group narrative intervention in Head Start preschools: Implications for response to intervention. *Journal of Early Childhood Research*, *13*(2), 196–217. https://doi.org/10.1177/1476718x13515419

Spencer, T. D., & Petersen, D. B. (2018). Bridging Oral and Written Language: An Oral Narrative Language Intervention study with writing outcomes. *Language, Speech &*

Hearing Services in Schools, *49*(3), 569–581. https://doi.org/10.1044/2018_lshss-17-0030

Spencer, T. D., Moran, M., Thompson, M. S., Petersen, D. B., & Restrepo, M. A. (2020). Early efficacy of Multitiered Dual-Language instruction: Promoting preschoolers' Spanish and English oral language. *AERA Open*, *6*(1), 233285841989788. https://doi.org/10.1177/2332858419897886

Spencer, T. D., Kirby, M. S., Garcia, A. R., & Sturm, B. (2023). Additive Effect, Feasibility, and Cost-Effectiveness of Family Engagement Activities that Supplement Dual Language Preschool Instruction. *Early Childhood Education Journal*. https://doi.org/10.1007/s10643-023-01571-y

Suna, H. E. (2020). Akademik Başarının Yordayıcıları: Sosyoekonomik Düzey ve Okul Türü. *Journal of Economy Culture and Society*, *61*(1). https://doi.org/10.26650/jecs2020-0034

Vandewalle, E., Boets, B., Boons, T., Ghesquière, P., & Zink, I. (2012). Oral language and narrative skills in children with specific language impairment with and without literacy delay: A three-year longitudinal study. *Research in Developmental Disabilities*, *33*(6), 1857–1870. https://doi.org/10.1016/j.ridd.2012.05.004

Walberg, H. J., & Tsai, S. (1983). Matthew effects in education. *American Educational Research Journal*, *20*(3), 359–373. https://doi.org/10.3102/00028312020003359

Weddle, S. A., Spencer, T. D., Kajian, M., & Petersen, D. B. (2016). An examination of a multitiered system of language support for culturally and linguistically diverse preschoolers: Implications for early and accurate identification. *School Psychology Review*, *45*(1), 109–133. https://doi.org/10.17105/spr45-1.109-132

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child development*, *69*(3), 848–872.

Westerveld, M. F., & Gillon, G. T. (2009). Oral narrative context effects on poor readers' spoken language performance: Story retelling, story generation, and personal narratives.

International Journal of Speech-language Pathology, *12*(2), 132–141. https://doi.org/10.3109/17549500903414440

Curriculum Vitae

Name: Diya Nair

Post-secondary Education and Degrees:

Post Graduate Institute of Medicine Education and Research
Chandigarh, India
2016-2020 Bachelor's in Aud. & Sp. Lang. Path.

The University of Western Ontario
London, Ontario, Canada
2022-Present M.Sc. Speech Language Sciences

Related Work Experience

Teaching Assistant
The University of Western Ontario, Canada
2022-2024

Speech Language Pathologist
Latika Roy Foundation, India
2020-2022

Intern – SLP
Post Graduate Institute of Medicine Education and Research
2019-2020

Publications:

(i) Alam, N., Munjal, S., Kumar, R., Nair, D., & Panda, N. (2020). Application of Speech Range Profile in Hindi in Hyperfunctional Voice Disorders: A Case-controlled Observational Study. *International Journal of Phonosurgery & Laryngology*, *10*(1), 18-22.

(ii) Nair, A., Nair, D., Girdhar, M., & Gugnani, A. (2021). Optimizing developmental outcomes by setting smart goals individualized home program for children with disabilities during COVID-19. *Int. J. Physiother. Res*, *9*, 4028-4034.