Head Start Teachers' Vocabulary Instruction and Language Complexity During Storybook Reading: Predicting Vocabulary Outcomes of Students in Linguistically Diverse Classrooms

Miriam G. Lipsky

University of Miami, miriamglipsky@gmail.com

Follow this and additional works at: https://scholarlyrepository.miami.edu/oa_dissertations

Recommended Citation


This Open access is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarly Repository. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.
HEAD START TEACHERS’ VOCABULARY INSTRUCTION AND LANGUAGE COMPLEXITY DURING STORYBOOK READING: PREDICTING VOCABULARY OUTCOMES OF STUDENTS IN LINGUISTICALLY DIVERSE CLASSROOMS

By

Miriam G. Lipsky

A DISsertation

Submitted to the Faculty of the University of Miami in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Coral Gables, Florida

May 2011
HEAD START TEACHERS’ VOCABULARY INSTRUCTION AND LANGUAGE COMPLEXITY DURING STORYBOOK READING: PREDICTING VOCABULARY OUTCOMES OF STUDENTS IN LINGUISTICALLY DIVERSE CLASSROOMS

Miriam G. Lipsky

Approved:

Maria Carlo, Ph.D.  Terri A. Scandura, Ph.D.
Associate Professor of Teaching & Learning  Dean of the Graduate School

Jeanne Shay Schumm, Ph.D.  Nicholas Myers, Ph.D.
Professor of Teaching & Learning  Assistant Professor of Educational & Psychological Studies

Daryl Greenfield, Ph.D.  Rebecca Shearer, Ph.D.
Professor of Psychology & Pediatrics  Assistant Professor of Psychology
Previous research indicates that joint storybook reading between caregivers (parents or teachers) and children can have positive effects on the oral language development of young children (Whitehurst et al., 1988; Dickinson & Smith, 1994). This study aimed to add to this body of research by providing information on the relationship between teachers’ language complexity and vocabulary strategies used during storybook reading and vocabulary outcomes for monolingual and dual language learners in linguistically diverse Head Start classrooms. Videotapes of 23 Head Start teachers were coded and analyzed for vocabulary instruction strategies and language complexity during storybook reading using hierarchical regression techniques to determine how these factors related to gains in student vocabulary over the course of a year. Students’ oral language was assessed using the Peabody Picture Vocabulary Test (PPVT) and the Learning Express vocabulary sub-test (LE). Teacher’s language complexity was calculated with respect to both quantity (mean length of utterance) and quality (type/token ratio and use of uncommon words). Teacher’s use of vocabulary strategies was examined with respect to the words chosen for instruction and the strategies used to instruct those words. Overall, teachers in this study tended to choose high utility words to
instruct, but the strategies used to teach those words, and the number of words chosen for instruction, were often not optimally aligned with best practices in vocabulary instruction (Beck et al., 2002). For the PPVT outcome measure, teachers’ use of higher numbers of vocabulary instruction strategies per word was differentially related to students’ vocabulary outcomes based on the student’s prior vocabulary knowledge, such that the use of more vocabulary instruction strategies per word was negatively related to vocabulary outcomes for students who began the year with the smallest vocabularies. There was also a significant interaction between teachers’ use of uncommon words and students’ prior vocabulary knowledge, though this relationship was only statistically significant for the PPVT outcome. Teachers’ use of more complex language was differentially related to students’ outcomes. Students who began the year with the lowest vocabulary levels exhibited a positive relationship between the teacher’s use of uncommon words and vocabulary outcomes, while average or higher vocabulary students showed a negative relationship between increased exposure to these uncommon words and their vocabulary outcomes. The implications for teachers’ professional development are discussed.
Dedication

I would like to dedicate this work to my family. First and foremost, my husband Jack, who has believed in me and supported me from the time I chose to start the masters program for my initial certification in elementary education, through some low-paying and challenging teaching jobs, and throughout this doctoral program. He acknowledged that this was a journey worth taking because he saw that I believed in what I was doing, even as he sometimes asked, “What is it that you do, and what exactly do you plan to do when you’re done?” It is this unconditional support that has kept me going during those times when I felt COMPLETELY overwhelmed. Jack also served as my personal IT department, providing computer support 24/7, and making sure that there were multiple back-up copies of my dissertation.

I also dedicate this work to my children, Megan and Lauren, who provided my initial inspiration for this career change through their love of children’s literature (which I also loved reading to them). They have endured many changes to our family routines in order to accommodate my studies. With a mother who studies children’s vocabulary acquisition, they have learned that when they ask the meaning of a word, it will be dissected into its morphological components (and hopefully they have learned from this strategy). They have occasionally accompanied me to classes – always on their best behavior, and they have understood when my academic work has prevented me from helping them with homework or attending various school functions (although this was rare). I am also indebted to Lauren for reading close to thirty children’s books with me in one afternoon as part of the process of selecting the book to be used by the teachers in this study.
I am grateful for the support of my parents, without whom this endeavor would not have been possible. My mother, Ruth Glaser, chauffeured my children, fed my family, walked our dog, and was always there to listen in my times of stress. My father, Dr. Luis Glaser, was happy to act as a sounding board as I considered various dissertation topics, gave career advice and guidance when asked, and provided valuable insight into life at “the U.” He also made sure I ate well, with frequent lunches/advisory meetings at his favorite spot – The Bagel Emporium.

I would also like to thank my late father-in-law Bernard Lipsky, who told me when I began my masters degree that before I knew it, I’d be done. At the time, I thought he clearly did not understand the overwhelming nature of the task I planned to undertake, but in retrospect, he was right, and I am sorry that he is not here to celebrate this milestone with me. Special thanks also go to my mother-in-law, Terri Lipsky, for taking me for “retail therapy” to break up my long stretches of studying.

Last, but not least, I am grateful for the companionship of my dog Max, who spent countless hours lying next to my desk while I read, researched, wrote, and edited. He made the quiet times a bit less lonely. I am also thankful that Max did not eat my dissertation, as he has been known to often consume important homework assignments… this particular assignment being the most important one ever.
Acknowledgement

I would like to thank my chair, Dr. Maria Carlo, for her support, creativity, editing, and guidance in the preparation of this dissertation, and especially for her help in narrowing my choice of dissertation topics down to this one, which has turned out to be a wonderful match for my interests.

I would also like to thank Dr. Daryl Greenfield for allowing me to conduct my dissertation work in conjunction with his Early Childhood Hands on Science (ECHOS) grant. I could never have collected the volume of data that I needed for this study if I had been working on my own, and I am extremely grateful to Daryl for including me on this project and permitting me to collaborate with his lab. I am also thankful to the Institute for Education Sciences (IES), and to Daryl as the PI on the IES training grant, for providing the funding for my predoctoral fellowship.

Dr. Nick Myers has been invaluable to me as an advisor for my statistical methods. I am grateful to him for his guidance in terms of methodology, and for reviewing and making suggestions when things did not turn out as expected, which he had warned me was… to be expected!

Dr. Jeanne Schumm was the first to suggest – at my Master’s graduation luncheon – that I should return to UM for my doctorate. I thank her for planting the idea of a doctorate in my head, even though it took a few years for the idea to germinate. I am also grateful to Jeanne for her careful review of my proposal, and for her encouragement as I was writing my dissertation.

I also would like to thank Dr. Rebecca Shearer for her help with the methodological considerations I encountered when rating classroom interactions.
Rebecca made valuable suggestions, which were especially helpful in figuring out issues related to rater agreement.

I am especially grateful to Dr. Janna Fucillo, who allowed me to become a part of her study of teacher-child interactions, and was willing to modify her study design somewhat to accommodate the needs of my study. Janna’s coordination of the transcription of the videos used in this study was enormously helpful to me. I would also like to thank Lizzy Bell for her help in providing clarification and research when I had questions about student data.

I am thankful to Alain Bengochea, who provided tremendous assistance to me in coding some of the transcripts in this study for my calculation of inter-rater agreement. I would also like to thank Dr. Amanda Goodwin for her encouragement and camaraderie as we journeyed through our doctoral work together.
# TABLE OF CONTENTS

| LIST OF FIGURES | ix |
| LIST OF TABLES | x |

**Chapter**

1 INTRODUCTION AND THEORETICAL FRAMEWORK ........................................ 1  
Addressing the 30 Million Word Gap ......................................................... 2  

2 REVIEW OF THE LITERATURE ...................................................................... 4  
Interaction during Storybook Reading ....................................................... 7  
Matthew Effects in Relation to Storybook Reading ...................................... 13  
Impact of Teacher Questioning During Storybook Reading ............................ 19  
Studies of Teacher Reading Style .............................................................. 25  
Present Study ............................................................................................. 29  
Research Questions .................................................................................... 30  

3 METHODS .................................................................................................... 32  
Participants .................................................................................................. 32  
Procedures .................................................................................................... 35  
Measures ........................................................................................................ 36  

4 RESULTS ...................................................................................................... 43  
Teachers’ Vocabulary Instruction Strategies and Language Complexity .......... 43  
Instructed Words and Instructional Strategies ............................................... 47  
Impact of Teacher Strategies and Language Complexity on Student Outcomes 49  
Hierarchical Regression Analysis for PPVT Outcome .................................... 52  
Hierarchical Regression Analysis for Learning Express Outcome .................. 57  

5 DISCUSSION ................................................................................................ 60  
RQ #1:  
Teacher’s Vocabulary Instruction Strategies and Language Complexity .......... 61  
RQ #2:  
Frequency of Vocabulary Instruction and Words Chosen for Instruction ........ 63  
RQ #3:  
Relationship of Vocabulary Strategies and Language Complexity to Students’ Vocabulary Outcomes ............................................................... 65  
Limitations of the Study and Directions for Future Research ........................ 72  
Implications for Professional Development .................................................. 74
REFERENCES .................................................................................................. 76

APPENDICES

A Language Acquisition ..................................................................................... 83
B Head Start: A Historical Perspective ............................................................ 95
C Figures and Tables .......................................................................................... 97
D Computation of Intra-Class Correlation ....................................................... 111
E Pearson Correlation Matrix ........................................................................... 113
LIST OF FIGURES

FIGURE 1.................................................................................................................. 97
FIGURE 2.................................................................................................................. 106
FIGURE 3.................................................................................................................. 107
FIGURE 4.................................................................................................................. 108
FIGURE 5.................................................................................................................. 109
FIGURE 6.................................................................................................................. 110
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE 1</td>
<td>98</td>
</tr>
<tr>
<td>TABLE 2</td>
<td>99</td>
</tr>
<tr>
<td>TABLE 3</td>
<td>100</td>
</tr>
<tr>
<td>TABLE 4</td>
<td>101</td>
</tr>
<tr>
<td>TABLE 5</td>
<td>102</td>
</tr>
<tr>
<td>TABLE 6</td>
<td>103</td>
</tr>
<tr>
<td>TABLE 7</td>
<td>103</td>
</tr>
<tr>
<td>TABLE 8</td>
<td>103</td>
</tr>
<tr>
<td>TABLE 9</td>
<td>104</td>
</tr>
<tr>
<td>TABLE 10</td>
<td>104</td>
</tr>
<tr>
<td>TABLE 11</td>
<td>105</td>
</tr>
<tr>
<td>TABLE 12</td>
<td>105</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction and Theoretical Framework

In their groundbreaking study of children’s vocabulary growth during the first three years of life, Hart and Risley (1995) followed 42 families for 2 ½ years. The families ranged in socio-economic status (SES) from professionals, to working class, to those living in poverty. Hart and Risley describe in great detail how both the quantity and quality of language experienced by children differed significantly by socio-economic status (SES), with the children of professional (college educated) parents hearing an average of 2153 words per hour, while the children living in poverty heard an average of 616 words per hour (Hart & Risley, 2003). Taken cumulatively, this has been referred to as “the 30 million word gap,” indicating that by the age of 3, children from more privileged families have heard over 30 million more words than children from less privileged circumstances. In the Forward to Hart and Risley’s (1995) book, Lois Bloom eloquently quantifies this 30 million word gap:

… the finding is heartbreaking that by the time the children were 3 years old, parents in less economically favored circumstances had said fewer different words in their cumulative monthly vocabularies than had the children in the most economically advantaged families in the same period of time. (Hart & Risley, 1995, p.xiii)

How could these differences in the language environment produce such large differences in outcomes? The relationship between the language environment and children’s language outcomes can best be understood within the interactionist framework of language acquisition, which is described in Appendix A.
When viewing language acquisition through an interactionist lens, it is clear that both the language inputs and the interaction associated with those language inputs play a significant role in children’s oral language development, and vocabulary development in particular. In conditions of high need, where children live in poverty, the work of Hart and Risley (1995) and Pan et al. (2005) indicates that poverty is often not only an economic condition, but can at times be accompanied by a poverty of linguistic input and interaction as well. The next section will describe efforts by the federal government to address this linguistic poverty as part of a larger war on poverty.

Addressing the 30 Million Word Gap

The work of Hart and Risley (1995) and others (Pan, Row, Singer, & Snow, 2005; Hoff-Ginsburg, 1991) has attempted to quantify the relationship between caregivers’ child-directed speech CDS, which constitutes the child’s language environment, and children’s vocabulary development. Hart and Risley (1995) found that both the quantity and quality of parents’ CDS, and the language environment in the homes in general, differed based on the SES of the parents. As Hart and Risley (2003) continued to follow the children in their original (1995) study, they found that vocabulary use and rates of vocabulary growth at age 3 predicted measures of language skill (receptive vocabulary, listening, speaking, semantics, and syntax) at age 9-10, demonstrating the impact of early language skills on academic outcomes at 3rd grade. A longitudinal study of a larger number of children by Pan et al. (2005) investigated the predictors of vocabulary production for children from age 1 to 3, all of whom came from low income families. Pan and her colleagues found that there was a great deal of variation in children’s vocabulary growth, and that this variation was related to the diversity of mothers’ lexical
input, as well as maternal literacy and language skills. Unlike the findings of some previous studies (Hart & Risley, 1995; Hoff-Ginsburg, 1991), Pan et al. found that the quantity of maternal speech was not a factor in children’s vocabulary growth, only the quality of the input was significant.

This research indicating that children living in poverty are often exposed to a lower quality and/or quantity of lexical input provides some of the rationale for programs such as Head Start and Early Head Start, which are designed to enhance children’s development and school readiness. A historical perspective of Head Start is presented in Appendix B.

The present study will investigate a prevalent practice in Head Start classrooms - and preschool classrooms in general - shared storybook reading. There has been a trend in the literature to investigate shared storybook reading as a vehicle for exposing children to novel vocabulary and thereby increasing vocabulary growth. As the following review of the literature will demonstrate, there is fairly strong evidence to indicate that the quantity of time children spend in shared storybook reading, either with their parents or teachers, has a positive impact on oral language development. The mechanisms for that influence – the particular qualities of the storybook reading interaction which influence vocabulary development – are less clear. In the present study, teacher’s incidental vocabulary instruction and language complexity during storybook reading will be investigated in relationship to children’s vocabulary outcomes in order to identify qualities of teacher language and instruction which are most strongly related to children’s vocabulary outcomes.
Chapter 2: Review of the Literature

A recent report by the National Early Literacy Panel (NELP, 2009) stated that, “shared-reading activities are often recommended as the single most important thing adults can do to promote the emergent literacy skills of young children” (National Early Literacy Panel, p. 153). The NELP conducted a meta-analysis of randomized control trials and quasi-experimental intervention studies which focused on shared reading, and found that most studies measured the impact of the interventions on oral language skills. The meta-analysis found a moderate sized effect on children’s oral language for shared reading interventions. The report goes on to say, “given the ubiquity of both the practice of and the recommendation for shared reading in early childhood education settings, it is somewhat surprising that more studies have not investigated the impact of these practices” (National Early Literacy Panel, p. 162). The studies included in this meta-analysis included interventions which involved a shared reading intervention, conducted by either teachers or parents, which usually lasted from 1-6 months, and involved a change in either the quantity or style of shared reading activities. This meta-analysis found positive, substantial, and significant effects of these interventions on children’s oral language skills and print knowledge, with effect sizes considered to be in the moderate range of 0.66 – 0.73 for oral language.

Additionally, a recent meta-analysis by Marulis and Neuman (2010) suggests that vocabulary interventions, including those that had a storybook reading component, can improve the vocabulary development of at risk children. Sadly, the results of this meta-analysis indicated that middle-and upper-income at risk children were far more likely to
benefit from vocabulary training than those children who were at risk and poor, for whom vocabulary interventions were not powerful enough to close the vocabulary gap.

The effectiveness of shared reading in developing children’s oral language had been the focus of a pair of earlier studies (Scarborough & Dobrich, 1994a; Bus, van Ijzendoorn, & Pellegrini, 1995), which had called into question the effects of shared reading on literacy outcomes; however the majority of the studies in these reviews dealt mainly with the quantity of shared book reading, rather than the qualities of the readings. Scarborough and Dobrich (1994a) reviewed studies that dealt with frequency of storybook reading (the majority of studies), quality of storybook reading, and intervention studies. Outcomes were broken into academic skills, emergent literacy skills, and language skills. Overall, this review found weak evidence for positive effects of storybook reading, which elicited much subsequent discussion/rebuttal (Lonigan, 1994; Dunning, Mason, & Stewart, 1994; Scarborough & Dobrich, 1994b) questioning Scarborough and Dobrich’s analysis.

One year after Scarborough and Dobrich’s review, Bus, van IJzendoorn, and Pellegrini (1995) conducted a meta-analysis of storybook reading studies. Due to the small number of studies which considered the quality of book reading, and the large variety of outcome measures of those studies, research on the quality of book reading was not included in this meta-analysis. The focus of this meta-analysis was on studies of quantity/frequency of reading only. In contrast to Scarborough & Dobrich (1994a), Bus et al. found strong positive effects of storybook reading with respect to outcomes such as oral language (vocabulary), emergent literacy, and reading achievement. Bus et al. (1995) noted that storybook reading seemed to particularly influence acquisition of the
written language register. The overall effect size of \( d = .59 \), which would be considered a medium effect size, indicated that approximately 8% of the variance in the outcomes (oral language, emergent literacy, and reading achievement) could be explained by storybook reading frequency. Previous work by Purcell-Gates (1988), Dickinson, De Temple, Hirschler & Smith (1992), and Snow and Tabors (1993) also indicated that shared book reading influences children’s acquisition of the written language register.

There are many studies of shared storybook reading that are not included in these meta-analyses, either because they are more qualitative in nature or because they did not meet other criteria for inclusion. Studies that focused on the quality of the storybook reading interaction were often excluded from the previous meta-analyses because of the diversity of measures employed in those studies. The studies that focus on the quality of the storybook reading interaction are relevant to the present research however, and will be included in this review. These studies focus on shared storybook reading, as do those in the NELP meta-analysis, and involve teachers, parents, or a combination of teachers and parents reading with children. Additionally, some studies focus on teachers reading to a whole class, while others focus on teachers reading with a small group of children (a subset of the class). There are a number of studies that focus on an adult reading one-on-one with a child, and while the adult studied is often a parent, it is also sometimes a researcher or teacher.

The studies chosen for this review all had outcome measures related to oral language, and specifically vocabulary, or they focused on defining teacher reading style. The studies with vocabulary measures evaluated receptive vocabulary, expressive vocabulary, or both. Most of these studies focused on the quality or quantity of
interaction before, during, and immediately after the book reading event, with a subset of
studies looking at the quality or quantity of interaction involved in the storybook reading
event and the differential effect of this interaction on children with relatively higher or
lower levels of initial vocabulary. Another subset of studies focused on the child-directed
questions which were part of the storybook interaction, while an additional group of
studies investigated children’s vocabulary acquisition in relation to the number of
readings of a storybook or number of exposures to the target words. The studies of
teacher reading style were more qualitative, and they informed the development of the
protocol for coding video transcripts used in the present study.

To identify relevant studies, descriptors such as Storybook Reading, Story Reading,
Vocabulary, Vocabulary Development, Oral Language, Emergent Literacy, Interactive
Reading, Preschool, and Head Start were used to search ERIC, Education Full Text, and
PsycINFO. Over 130 abstracts from peer-reviewed journals were evaluated for inclusion in
this literature review. References from publications found in the previous database searches
were also examined, and some were selected for inclusion in this review.

Interaction During Storybook Reading

Just over 20 years ago, Whitehurst, Falco, Lonigan, Fischel, DeBaryshe, Valdez-
Menchaca, and Caulfield (1988) conducted a seminal study that demonstrated the power
of shared storybook reading to increase children’s expressive language ability. In this
study, Whitehurst et al. introduced the term “dialogic reading,” which was defined to
mean an interactive style of reading which parents would use when reading storybooks to
their children. When parents in the study read to their children using a dialogic style,
they were encouraged to use a variety of techniques, rather than one isolated behavior
change. These techniques included encouraging the child to speak through the use of evocative, open-ended questions; repeating, expanding, and recasting the child’s speech by providing maximally informative feedback to the child; including expansions of what the child said and/or corrective modeling (for example, child: “he goed home”, parent: “yes, he went home on the bus.”); and providing praise and corrective feedback contingent on the child’s level of speech. All of these techniques are consistent with the social interactionist perspective of language development. The parents were taught to use these techniques during two 25-30 minute training sessions over the course of the one-month intervention. The results indicated that children in the dialogic reading group increased both their expressive and receptive vocabularies, as well as mean length of utterance (MLU) when compared to the comparison group. Although the duration of the intervention in this study was relatively short (one month), the study’s strong methodological framework – a longitudinal design with random assignment to experimental and control conditions – adds to the credibility of the findings.

Subsequently, a number of studies have examined reading style in greater detail. Senechal and Cornell (1993) examined the use of 4 reading styles on the acquisition of novel vocabulary by 4 and 5 year olds. The reading styles were: use of questions, use of recasts, use of repetition to emphasize certain words, and reading the book verbatim. Interestingly, the data indicated that reading style did not affect vocabulary acquisition – children in all four styles did not differ significantly on vocabulary acquisition, according to receptive measures. While statistically significant results were found for children’s receptive vocabulary, there were no conclusive results for expressive vocabulary. Previous work by Whitehurst and DeBaryshe (1989) also indicated that the process of
acquiring receptive and expressive vocabulary may differ. Additionally, children were only exposed to the target vocabulary in a single reading of the book. The authors suggest that while a single reading of a book may be sufficient to produce results on receptive measures, they may not be sufficient to produce results on expressive measures. A single reading of the book in a particular style also may be insufficient to allow differences in vocabulary acquisition to surface as a function of reading style.

In an intervention which included training for parents and school-based and home-based parent/child activities centered around storybook reading, Jordan, Snow, and Porche (2000) reported that kindergarteners made substantial gains in language skills when participating in the program with their parents. The program, called Project EASE (Expanded Access for Student Enrollment), was designed to increase language interactions to promote developing literacy abilities. While this intervention did not focus exclusively on vocabulary development, strengthening vocabulary was one aspect of the parent training program, and vocabulary was part of one of the composite measures of language skills measured by the study. Results indicated that, in the area of vocabulary, the intervention impacted children’s recall of superordinate terms, which have been shown to be related to reading skills (Snow, 1990). The intervention did not have a significant impact on children’s vocabulary as measured by the Peabody Picture Vocabulary Test, but the researchers pointed out that this measure is designed more to capture incidental vocabulary acquisition than the effect of a particular curriculum.

One issue that has not commonly been considered in the implementation of parent training programs that aim to increase the quality and/or quantity of shared reading is how to implement these programs among language-minority or culturally-diverse
families. Recently, Jimenez, Filippini, and Gerber (2006) considered this question when they implemented a dialogic reading program where Latina/o caregivers read to children in Spanish, their primary (home) language. While this study did not measure outcomes related to vocabulary, it did find that parents in the intervention significantly increased their use of the dialogic reading strategies (which they were taught as part of the intervention) over the course of the intervention, and that children’s quantity and variety of language increased (more productive language was used). The authors note that future research should focus on the possibility of increasing academic skills through shared storybook reading in the child’s primary language. Current theories of cross-linguistic transfer of academic skills suggest that literacy-related skills learned in a child’s primary language often transfer to the second language (Cummins, 1991). Whether storybook reading in a child’s first language can impact vocabulary in a second language is clearly an area which requires further research. Regardless of the impact of first-language reading on second-language vocabulary, introducing culturally diverse children to a dialogic style of reading at home may familiarize them with this style of reading which they may encounter when entering school.

The first to implement a small-group, school-based storybook reading intervention for children who were not exposed to this practice at home were Klesius and Griffith (1996). Their descriptive study focused on developing the language and literacy skills of children, including second-language learners, who enter school with limited exposure to storybook reading. The purpose of this study was to assess whether an interactive style (a more dialogic style) of storybook reading could be implemented in a classroom environment. Given that children who acquire literacy skills early in their
schooling tend to come from homes where storybook reading is a frequent occurrence (Clark, 1984; Durkin, 1974/1975). Klesius and Griffin hoped to close the gap for at-risk children (those with limited home storybook reading experiences) by introducing them to this more intimate form of reading in a small-group setting within the school. During the three weeks in which this program was implemented, the kindergarteners who participated (in small groups of up to five children per adult reader) experienced many of the components typically found in one-on-one shared storybook reading including: questioning, scaffolding, increased participation/talking by children, “lap reading ambience” (defined as praise and closeness with the reader), and increased interest in books. These efforts were followed by a number of studies which systematically evaluated such practices.

Subsequently, several studies have found a correlation between storybook reading and vocabulary development, while others have not. In a more quantitative and long-term project than the one described by Klesius and Griffith (1996) above, Lonigan, Anthony, Bloomfield, Dyer and Samwel (1999) designed a study to determine whether typical shared reading or dialogic (interactive) shared reading had different effects on preschool (2-5 year-old) children’s oral language, listening comprehension, and phonological sensitivity over the course of a two-year intervention. Findings indicate that both interventions affected emergent literacy skills (descriptive use of language increased significantly in the dialogic reading group, while listening comprehension and alliteration detection increased significantly in the typical shared reading group), but no statistically significant findings emerged on the receptive or expressive vocabulary measures. The authors suggest that part of the difficulty in implementing dialogic
reading in groups is that some of the elements of dialogic reading, namely tailoring responses to a child’s zone of proximal development, may be difficult or impossible to implement in a group setting with children at different developmental levels, thus leading to the lack of significant effects on some areas of literacy development when implemented in a group setting. The authors further explain that part of the reason why more differences were not observed between the dialogic and typical shared reading groups may be because the books which were chosen for the intervention had many colorful illustrations which supported the text of the story, thus providing rich context in which to understand the story and any novel vocabulary which was introduced. Furthermore, the readers in the typical shared reading group not only read the book to the children, but also commented on the pictures and answered children’s questions as part of the intervention, and thus may have provided enough scaffolding for these children to have significant oral language outcomes, similar to those children in the dialogic reading group. Thus, the authors’ choice of books for this study created a situation in which there was not as much contrast between the two conditions as would have been ideal, thereby limiting the interpretation of the findings.

Another study (Silverman, 2007a) of reading styles focused on teachers’ approaches to vocabulary instruction during storybook reading, but rather than examining teachers’ natural reading styles. This investigation compared three types of instruction implemented as part of an intervention for kindergarten students. Approximately one third of the students studied were English language learners (ELLs), and one third were from low SES backgrounds, although it is unclear from the study description if these classifications were mutually exclusive. This study utilized 3 types of instruction:
contextual (based on connecting words to their use in books and to children’s personal experience), analytical (an enhancement of contextual instruction which included semantic analysis of words in contexts other than the books and children’s experience), and anchored (augmented analytical instruction with attention to the spoken and written forms of the word). Silverman found that both analytical and anchored methods of instruction were significantly more effective than the contextual method at promoting children’s learning of words targeted for instruction, as measured by a researcher-created assessment targeting these specific words. Interestingly, data in this study indicated that for children from low-SES or language-minority backgrounds, the analytical method of instruction during reading seemed to be the most effective. This is an important finding, since much of the previous research on storybook reading has not focused on ELLs, and this provides support for the theory that direct instruction is most effective for these students.

Matthew Effects in Relation to Storybook Reading

The term “Matthew effect” was first coined by Merton (1968) in the field of sociology, where a Matthew effect referred to a phenomenon where “the rich get richer and the poor get poorer.” This phrase derives from the biblical Gospel of Matthew. In the literature related to reading, a study by Stanovich (1986) proposed that there are Matthew effects in reading. Stanovich argued that there is an organism-environment correlation such that different organisms are exposed to different, and non-random, levels of environmental quality (which is similar to what is indicated by the 1995 and 2003 research of Hart and Risley, discussed earlier). Stanovich further indicated that the relationship between reading ability and cognitive processes related to reading is
bidirectional, so that differences in children’s processing abilities could lead to differential opportunities to learn, and that these differential opportunities to learn lead to differences in the level of knowledge which is acquired. In simpler terms, children who are better readers read more, and continue to improve their reading-related skills, while children who are not as proficient in reading read less, and ultimately have fewer reading skills than their well-read peers. While Stanovich’s work applied to children who were already reading on their own, a number of studies have examined Matthew effects in relation to younger children who participate in storybook reading interactions.

Senechal, Thomas, and Monker (1995) examined the effects of children’s active participation in storybook reading. They conducted experiments with 4-year-old children to determine how children with relatively larger vocabularies acquire new vocabulary from listening to storybooks being read aloud as compared to children with relatively smaller vocabularies. In this study, children were placed in one of three conditions, they either listened passively to a storybook read by an adult (a researcher), they pointed to pictures of target vocabulary during the reading, or they labeled target vocabulary during reading. The results indicated that children with larger vocabularies comprehended more novel words than children with smaller vocabularies (a Matthew effect), and children who actively participated by pointing or labeling learned more words than children who listened passively. This study assessed both expressive and receptive vocabulary, and found that children who were more actively engaged in the reading interaction (by labeling target words) were able to produce more of those words at post-test, indicating that the labeling condition had a greater impact on productive vocabulary.
Robbins and Ehri (1994) also considered the effect of children’s vocabulary size on acquisition of new vocabulary. In their study of kindergarteners, children were read the same storybook twice (by the experimenter), two days apart, and then assessed on their receptive knowledge of 22 unfamiliar words, half of which appeared in the story. Some target words appeared in the story once and some appeared twice. At post-test, children recognized more words that were in the story than not in the story, and four exposures to words seemed to be necessary, but not sufficient, for higher rates of word learning. Once again, children with larger entering vocabularies made greater gains. These differences in vocabulary acquisition among children with larger and smaller vocabularies were hypothesized by Senechal (1995) to be due to children with larger vocabularies having more efficient retrieval mechanisms, a theory which is congruent with Stanovich’s (1986) work on Matthew effects in reading.

Further evidence for the positive effects of active participation by children in storybook reading with adults comes from Ewers and Brownson (1999), who also studied children’s vocabulary size in relation to vocabulary acquisition. Their study compared expressive and receptive vocabulary acquisition of children in two storybook reading conditions: active participation by answering “what” or “where” questions regarding target words or passive participation by listening to recasts containing the target words. Children in the active participation category acquired more words, and children with higher initial vocabulary acquired more words than children with lower initial vocabulary at pre-test. This reconfirms the Matthew effect found by Senechal et al. (1995) with regard to children with larger vocabularies acquiring more novel words during storybook reading than those with smaller vocabularies, and also reaffirms the work of Whitehurst
et al. (1988) with regard to children’s active participation in reading leading to better oral language outcomes.

Coyne, McCoach, Loftus, Zipoli, and Kapp (2009) found similar results in a study that compared embedded instruction of target words (during storybook reading) with extended instruction, in which words were instructed not only during storybook reading but also outside the context of the story. The results of this study indicated that students in both the embedded and extended instructional conditions learned approximately two thirds of the target words which were introduced, but those in the extended instruction condition developed more complete knowledge of the words. Students in both conditions significantly outperformed a control condition in which students were just incidentally exposed to the target words during storybook read alouds. In this study, Coyne and his colleagues also found that children with higher initial PPVT scores tended to score higher on the target vocabulary outcome measures, providing further evidence for the Matthew effect discussed above.

The work of Collins (2005) was the only study which reported the effects of a storybook reading intervention exclusively on a group of ELLs, and also examined the effect of initial vocabulary on language outcomes. These Portuguese-speaking preschool students participated in an intervention where they heard one pair of stories read three times per week for three weeks, with rich explanations provided for the target vocabulary words. Rich explanations included pointing to illustrations, using gestures, providing brief definitions and synonyms (where applicable), and using the target word in a sentence that differed from the sentence in the book. A control group heard the same stories, but without the rich explanations. Results based on a receptive vocabulary
posttest of target words showed that rich explanation led to significant gains in new vocabulary in the intervention group, as opposed to the control group. Children with higher initial vocabulary scores in the L2 (English) learned more words than those with lower initial vocabulary scores (again a type of Matthew effect), but this finding would not be unexpected in a population of ELLs, since those with more limited vocabulary may have had difficulty in understanding the instruction or explanations which were given. This study, along with that of Silverman (2007a), highlights the significant impact that purposeful vocabulary instruction, incorporated into storybook reading, can have on children from language-minority backgrounds.

In another intervention study which focused on a reading style similar to the analytical style described by Silverman (2007a), Justice, Meier, and Walpole (2005) examined whether adult elaboration of targeted words in a storybook reading context would lead to increased acquisition of those targeted vocabulary words by at-risk kindergarteners. Children were defined as at-risk because greater than two-thirds of their school population qualified for free or reduced-price lunch. Children were pre-tested using a standardized receptive vocabulary test in order to classify them into high and low vocabulary skill groups, so that the impact of the intervention could be measured for these groups separately. The children in the treatment group completed 20 small-group storybook reading sessions with 10 books (each book was read twice). Within these ten books, 60 target words were chosen: 30 of the 60 target words were elaborated by the reader (who provided meaning followed by an example), and 30 of the 60 target words were not elaborated, children were just incidentally exposed. Gains in word learning were modest, but there were greater gains on elaborated words than non-elaborated
words. It is possible that the vocabulary measures used by Justice et al. (2005), which were only expressive vocabulary assessments, may not have been sensitive enough to capture all of the actual gains in vocabulary. The addition of a receptive vocabulary assessment to this study might have helped to better capture children’s word-learning gains. Remarkably, children with lower initial vocabulary scores made greater gains on elaborated words. This finding stands in contrast to most of the studies described so far, which found that children with larger initial vocabularies made greater gains (a Matthew effect), and it shows that purposeful instruction, embedded within storybook reading, can help to reduce the gap between children with different initial levels of vocabulary.

A second study which suggests that explicit teaching of vocabulary word meanings during storybook reading benefits children with lower initial receptive vocabulary scores was conducted by Coyne, Simmons, Kame’enui, and Stoolmiller (2004). In this classroom-based intervention for kindergarteners focusing on explicit vocabulary instruction during storybook reading, students received 108 half-hour lessons which were designed to accompany 40 classic and/or award-winning children’s books. The intervention group scored significantly higher on an experimenter-developed test of explicitly taught vocabulary words. Students with lower initial receptive vocabulary demonstrated higher gains in word learning than children with higher initial receptive vocabulary. In contrast to some of the earlier studies which reported Matthew effects, this study is significantly longer and is conducted in a classroom setting, providing encouraging evidence for the possibility that storybook reading interventions in a classroom setting which focus on explicit vocabulary instruction can provided beneficial results in terms of closing the “30 million word gap” (Hart & Risley, 2003).
Clearly, the research on Matthew effects related to storybook reading is mixed. It seems that in shorter interventions, Matthew effects manifest themselves in children with larger initial vocabularies showing more substantial growth on outcome measures related to vocabulary. However, in studies of more long-term interventions, there is evidence that direct, explicit instruction may narrow the vocabulary gap between children who score relatively higher on vocabulary measures at the beginning of the study and those who score relatively lower. The latter findings are indeed encouraging to those of us who hope to narrow the 30 million word gap.

*Impact of Teacher Questioning During Storybook Reading*

One key component of dialogic reading interventions involves questioning children during reading in an effort to increase comprehension and keep them engaged in the story. A number of studies have considered this aspect of the storybook reading interaction in an attempt to better define the qualities of the questions which impact vocabulary growth.

The types of questions which affect vocabulary acquisition were considered in a recent study by Walsh and Blewitt (2006), who examined the effects of adult questioning style on novel word acquisition during storybook reading. In their study, three-year-olds were assigned to one of three conditions: eliciting questions (example for target word pagoda: “what is this?” “pagoda”), non-eliciting questions (“what color is the pagoda?” “black”), and no questions. Children were assessed using measures of both expressive and receptive target vocabulary. Walsh and Blewitt found that children’s word-learning increased more in both the questioning conditions than in the control. Although one could argue that the cognitive demands of the non-eliciting questions are higher than
those of the eliciting questions, the findings of this study indicate that the type of question being asked is not as important as the child being actively engaged in questioning in general,

The impact of adult questioning during storybook reading was also examined by Justice (2002). In this study, she examined the impact on preschool children’s acquisition of novel vocabulary when adult readers either labeled novel words or asked questions about them. Additionally, within the questioning condition, children were asked either perceptual or conceptual questions about the words. Perceptual questions included questions about concrete features of the story, such as “what color is the house?” or “what did the girl have in her hand?” Conceptual questions required children to make predictions or judgements, or provide explanations, such as “what do you think will happen next?” or “why do you think the boy is running away?” Children’s acquisition of the vocabulary was assessed using both expressive and receptive measures. Results indicated that adults’ labeling of novel words facilitated receptive word learning more than adult questioning. No effect of perceptual vs. conceptual questions was found. This study focused on words that were very specific to a particular context, with children being asked to learn the names of various species of birds (wren, sparrow, finch, etc.) and plants (geranium, forsythia, etc.). Given that much vocabulary instruction, especially for preschool children, is more focused on learning words which can be used in a broader context, the results of this study may have been different if the words chosen were more “general purpose” words.

The effect of questioning during storybook reading was one area of focus of a study by Allison and Watson (1994), who examined parent and teacher storybook reading
style and its impact on emergent reading (but not vocabulary specifically) of kindergarten students. Emergent reading level was defined as the child’s ability to independently read the book which had previously been read aloud to the child by both the parent and the teacher. This measure is problematic because it not only tests a child’s reading ability, but also the child’s ability to memorize the story, so a child who could memorize well would be rated as having high emergent literacy, when in fact that child may just have good memorization skills. Parent reading style was measured in one-on-one reading sessions with the child, while teacher reading style reflected the teacher’s style in reading to the whole class of students. Allison and Watson considered four independent variables in their analysis: the number of questions with high cognitive demand, the total number of cognitive demands (total questions with high and low cognitive demand), the number of storybook sessions per week, and the age at which the parent began reading to the child. They found that the age at which parent began reading to child, and teachers’ use of “high cognitive demand” questions were the only significant predictors of emergent reading level, and that these two variables together accounted for 30% of the variability in emergent reading levels. Parent reading style (as defined by the number and cognitive demand of questions) was not a significant predictor of emergent reading. Because this study does not measure impacts of the variables on the development of vocabulary per se, and may have in fact measured memorization skills, the findings should not be interpreted as contrary to those of the other studies discussed here, which do show significant effects of adults’ storybook reading on children’s vocabulary development.

Other studies indicate that vocabulary instruction embedded in storybook reading can be effective, but more word learning occurs if the book reading is followed up by
targeted vocabulary instruction. This is the strategy used by Beck and McKeown (2007) in a set of studies on kindergarten and first grade students in which the efficacy of Beck and McKeown’s Text Talk program was evaluated. In the Text Talk program, teachers are encouraged to read high quality picture books to their students – books which contain more sophisticated vocabulary than young children would be able to read independently in the early primary grades. Teachers then select approximately three words from the story for additional instruction, which occurs after the story is read. In this additional instruction, the use of the word in the story context is reviewed, often re-reading the page(s) where the word is encountered, and additional explanations and practice in use of the word are provided including giving the definition of the word, asking children to repeat the word aloud to create a phonological representation of it, using examples of the word in other contexts, asking children to make judgments about the teacher’s use of the word in several sentences, asking children to create sentences where they demonstrate proper use of the word, and giving the word’s definition and asking children to provide the target word (for example, “What word means to jump very high?” target word: leap). Results of this study by Beck and McKeown indicated that children demonstrated significant gains in word learning with words which were instructed in the manner described above.

The work of Beck and McKeown (2007) indicates that additional exposure to, and discussion of, target words can improve children’s retention of those words. In an attempt to quantify the vocabulary learning which takes place with multiple exposures to a word vs. a single exposure, Senechal (1997) compared vocabulary acquisition of 3- and 4-year-old children in three storybook reading conditions: a single reading of the target
storybook, repeated reading (3 times), and repeated reading with questioning (3 times). Children were read the story individually by the experimenter. In the questioning condition, children were asked to use the novel vocabulary words to label target items which were introduced in the book. This study allowed for the comparison of results between the single reading condition and the two multiple reading conditions, as well as the comparison of the two multiple reading conditions (with and without questions). Results indicated that repeated reading led to increased acquisition of expressive and receptive vocabulary, whereas repeated reading with questioning showed greater benefits for the acquisition of expressive vocabulary than receptive vocabulary, when compared to reading without questioning.

A study which examined the effects of teacher’s storybook reading style (which included questioning and other techniques) on oral language development in classroom settings is that done by Dickinson and Smith (1994). The researchers investigated the long-term effects of preschool teachers’ storybook reading style on low-income children’s story comprehension and vocabulary. In their study, Dickinson and Smith videotaped 25 book reading sessions in 25 different classrooms, where each teacher read a storybook of their own choosing to the class. The researchers then analyzed transcripts of those book reading episodes by coding each utterance that was not part of the story text. Utterances were coded at three levels: the first level indicated the timing of the utterance (before, during, or after reading of the text); the second level distinguished between offers of information, requests for information, and responses to requests in order to determine if an utterance was responsive to a previous utterance; and the third level of coding related to the specific content of the utterance and fell into three broad
categories – cognitively challenging talk, talk with lower cognitive demands, and talk which served to manage the interaction. Based on the frequencies and groupings of these various codes, Dickinson and Smith were able to classify teachers’ book reading style according to one of three approaches. The co-constructive approach was characterized by high amounts of talk by both children and teachers during book reading, frequent talk of an analytical nature, and little talk before and after the reading. The didactic-interactional approach was characterized by a limited amount of talk as the book was read and little talk before and after the book, with interaction based on immediate-recall questions or chiming a repeated phrase or familiar text. The performance oriented approach was characterized by talk that was mostly before and after the story was read, little talk during book reading (and talk that occurred during the reading was analytical in nature), extended book introductions, encouragement of predictions and personal connections, vocabulary analysis/explanation in discussions as text was read, and extended talk after the book was read which included reconstruction of the story or drawing links with children’s personal lives/experiences. Dickinson and Smith found that the performance oriented approach predicted greater gains in vocabulary one year later, even when controlling for other classroom factors. One variable in particular, which reflected teachers’ and children’s prompted and responsive utterances related to analysis, prediction, and vocabulary, was most strongly related to children’s scores on the outcome measures. This framework of describing teachers’ reading behaviors both holistically and at a more detailed level has been particularly influential in the design of the present study. Unlike the Dickinson and Smith study, however, teachers in the present study all read the same book to their students, thereby controlling for differences in style which might be
caused by the type of book selected. For example, in Dickinson and Smith’s didactic-interactional approach, teachers read a book which was very familiar to their students, so the students could chime in throughout the story. This was one of the key features of this style, but it would clearly not be possible in the current study, since the book which was used was unfamiliar to both the teachers and the students.

Questioning children during reading is a key component of a dialogic reading style. In the majority of studies reviewed above, questioning related to vocabulary, either during or after reading, led to improved retention of that vocabulary. Part of the reason that this may be the case is that answering questions about a particular word often involves additional exposure to that word. The work of Dickinson and Smith (1994) indicates that asking questions as part of storybook reading may occur more frequently with some teachers than with others as part of their natural reading style. Although the concept of “style” is somewhat difficult to quantify, a number of studies were reviewed in an effort to better determine ways in which “style” may be defined.

*Studies of Teacher Reading Style*

Several studies of teacher reading style were reviewed to determine the ways in which the authors explained or quantified a teacher’s “style.” Although none of these studies included quantitative data on children’s outcomes, their discussions of teacher reading style or vocabulary instruction related to storybook reading were informative for the design of the coding scheme for this study.

Enhancing children’s vocabulary and comprehension abilities to help narrow the large gaps which exist between children upon entry to kindergarten is the goal of a program called Text Talk, designed and described by Beck and McKeown (2001). To
better define the types of behaviors which might be grouped into the classification of a teacher’s reading “style”, I reviewed this study and several others. These studies did not report on student outcomes but sought to better define the types of behaviors which differ among teachers in the area of storybook reading. In observing kindergarten and first grade teachers reading storybooks to their classes, Beck and McKeown found that children often responded to questions based on the pictures in the book, not on the text of the story which they heard. They also found that students often responded to questions based on their background knowledge, regardless of what information had been conveyed by the text of the story. Further, Beck and McKeown found that teachers’ interactions with the students during reading mainly fell into two categories: clarifying or explaining content or vocabulary, and attempting to involve the students in the story by asking questions about the text which had just been read. Unfortunately, these questions were often what Beck and McKeown referred to as constrained questions, or questions which only required very short (often single word) answers.

The goal of the Text Talk program is to give children experiences with decontextualized language during storybook reading by not only listening to book language, but talking about the ideas found in the book in more lengthy discussions, not just with the use of constrained questions. In terms of vocabulary development, the Text Talk program makes use of some of the sophisticated language found in children’s picture books by encouraging teachers to choose several words from the story to target for explicit instruction after the story is read. In this vocabulary instruction, teachers are encouraged to recall the use of the word in the story with the students, define the word, use the word or respond to use of the word by the students, and encourage the students to
repeat the word out loud in order to have a phonological representation of the word. These vocabulary-based elements of the Text Talk program have been included in the research protocol for this dissertation in order to relate their use by teachers to student outcomes.

In another study that sought to better define teacher reading style, Martinez and Teale (1993) observed six kindergarten teachers who read the same 4 storybooks to their classes. For analysis, data were classified at the level of topic units (TU). A TU consists of all contiguous talk contributed by the teacher and/or students directed toward the same aspect of a story feature or story-related feature. Each TU was analyzed for the three major facets of style, which were identified as: focus, type of information, and instructional strategies. Focus was defined as the focus of the teacher’s talk during the storybook reading; type of information referred to the information that the teacher chose to talk about, or would encourage the student’s to talk about while reading the storybook; and instructional strategies were those strategies used by the teacher while reading the book. Storybook reading styles were identified as being different among the six teachers in terms of the frequency and way in which they mixed these three facets of storybook reading, but there was no comparison of child outcomes to determine if certain styles were better than others.

In a qualitative study of one teacher’s storybook reading style over the course of a year in an urban kindergarten classroom, Sipe (2003) set out to define one very successful teacher’s storybook reading style, and explain how the storytelling aspects of her reading style allowed her to make connections with her culturally diverse students. He suggests that research on story-reading style indicates that teachers tend to be
consistent in their reading style from book to book, but wonders if teachers should not be encouraged to vary their styles based on make-up of their classrooms in order to better match their style with the sociocultural backgrounds of their students.

Adding to the body of work that seeks to define the characteristics of a successful read-aloud technique for storybooks, McGee and Schickendanz (2007) define the characteristics of an effective interactive read-aloud based on their work as directors of several Early Reading First grants. They define a repeated interactive read-aloud technique which is:

- a systematic approach that incorporates teachers’ modeling of higher-level thinking, asking thoughtful questions calling for analytic talk, prompting children to recall a story in some way within a reasonable time frame, reading a single book repeatedly, and reading books related by topic. It also involves a systematic approach to developing children’s understanding of vocabulary, such as inserting short definitions of words and phrases during reading. (McGee & Schickendanz, 2007, pp. 742-743).

Many of these characteristics are gleaned from the research on one-on-one and group storybook reading described above, and McGee and Schickendanz seek to consolidate this research into a format that will be helpful to practitioners. These characteristics were also used to inform the selection of teacher-related variables in this study.
Present Study

Based on the existing literature, it is clear that incidental exposure to novel vocabulary words is probably insufficient to close the gap between children who begin school with relatively depressed vocabularies and those who begin school with richer vocabularies. The promising research coming from the social interactionist philosophy of language acquisition indicates that the unique characteristics of child-directed speech assist children in acquiring new vocabulary, but which of those characteristics of child-directed speech is most facilitative has yet to be discovered. Interactionist studies of language acquisition indicate that both the quantity and quality of language inputs (in this case, teacher language) and the quantity and quality of the interaction between the adult (teacher) and child are important to children’s vocabulary growth. A conceptual model of this theoretical perspective is presented in Figure 1 in Appendix C. Although this model shows that teacher language may influence teacher/child interactions, the focus of the present study is on the individual contributions of teacher language and teacher/child interactions to children’s oral language outcomes (the darker arrows), and not on the interaction between teacher language and teacher/child interactions.

In more structured settings, purposeful instruction of words that are likely unfamiliar to children has been shown to relate to word learning and increases in receptive and/or expressive vocabulary. Many of these findings, discussed in the review of the literature above, come from intervention studies in which one or more forms of vocabulary instruction were compared to a control condition, which usually consisted of incidental exposure to the target words, or “business as usual” in the control classrooms or groups. In the current study, the goal was to determine which vocabulary instructions
strategies, if any, are employed by teachers in Head Start classrooms, and the relationship between the frequency of use of these strategies and children’s vocabulary outcomes at the end of one school year. Additionally, both the quality and quantity of teacher language were measured during the storybook reading interaction to determine the relationship of these two elements to students’ end-of-year vocabulary outcomes. Based on previous research, I predicted that teachers who employed more vocabulary instruction strategies during storybook reading would have students with larger receptive vocabularies at the end of the year, as measured by the PPVT and Learning Express, after controlling for students’ initial vocabulary (at the beginning of the school year). I also expected, based on the work of Hart and Risley (1995) and Pan et al. (2005) that teachers who use more language, and more complex language, will have students who make larger gains in vocabulary. The small body of existing literature on teaching vocabulary through storybook reading for dual-language learners (DLLs) indicates that similar results should be expected for students who are monolingual English speakers as for those from language-minority backgrounds (Silverman, 2007b), but may differ based on students’ initial level of English (Collins, 2010), so a dummy-coded variable indicating DLL status was included in this analysis in order to investigate this further.

Research Questions

1. What vocabulary instruction strategies are Head Start teachers using when reading a storybook to their students, and what is the level of language complexity found in these storybook reading interactions?
2. How frequently do Head Start teachers use vocabulary instruction strategies while reading a storybook to their students (how many words are chosen for instruction, and how many instructional strategies are used)?

3. Do teachers’ vocabulary instruction strategies and complexity of language during storybook reading relate to outcomes in student vocabulary over the course of a year within Head Start classrooms?
Chapter 3: Methods

Participants

Teachers. The participants are 23 preschool teachers and 210 students from 5 Head Start centers in the southeastern United States who are participating in an Institute for Education Sciences (IES)-funded quasi-experimental study evaluating the effects of a preschool science curriculum on school readiness outcomes. For the larger study, in the 2008-2009 school year, six Head Start centers were selected from a pool of centers (N = 36) that met the following criteria: (a) had at least two Head Start classrooms, (b) located within 20 miles of the university’s campus, and (c) used the online version of the Galileo System for the Electronic Management of Learning. Thirty classrooms across six centers consented to participate. Directors, teachers and teacher assistants from 20 of the 30 classrooms consented to participate in the science curriculum training. The directors, teachers and teacher assistants from the remaining 10 classrooms consented to participate in the study without receiving training in the science curriculum. Consent was also requested from parents of children in participating classrooms.

For the purposes of the current study, teachers were approached to allow videotaping during observations of one storybook reading session at a time which was scheduled in advance with the researcher. Twenty-three of the larger study consented to participate in this additional component of the study, 17 of whom were participating in the experimental preschool science program and six of whom were participating in the study as controls. The experimental preschool science program did not include training on storybook reading, so teachers’ status as either experimental or
control group participants was not expected to influence this study. All of the participating teachers were female, 15 (65%) were Latino (all of whom were bilingual), seven (31%) were African American, and one (4%) was Asian. Four of the teachers held Masters degrees (in education or special education), 10 held bachelor’s degrees (5 in education, 1 in special education, 1 in psychology, 1 in English, 1 in Fine Arts, and one was not specified), and 9 held associates degrees (8 in early childhood/elementary/CDA, and one was not specified). The teachers’ experience in the classroom ranged from 0 to 30 years, with the teachers who reported “0” years having been in the classroom for less than one year, \(M=10.4, SD=7.6\). Teachers were asked to report on any training they had participated in within the three years prior to the school year in which these data were collected. In addition to training in the preschool science curriculum, which was provided as part of the larger study, five teachers reported participating in training on using music in the classroom. None of the teachers reported that they had been a part of any special projects or training related to storybook reading.

**Students.** In the larger study of the science curriculum, a sample of children was selected by first stratifying each classroom by age (younger and older than 4 years of age by September 1st) and then gender to balance the demographic composition of the sample. Initially, 10 children per class were randomly sampled, with the exception of three classrooms which only had nine eligible children. This resulted in an initial stratified random sample of 297 children. The children who remained were assigned to ‘alternate’ status. From this original random sample, a total of 77 children were dropped from the study: 38 due to low English proficiency, 12 due to non-compliance, 8 due to chronic absenteeism or tardiness, 7 due to relocation, 7 due to lack of assent, and 5 due to
lack of parental consent. A total of 110 children from the ‘alternate’ group, of the same
gender and/or age group, were selected to replace children who were dropped, ensuring
that the sample was equally distributed across age and gender. Thirty of the children from
this alternate group were subsequently dropped for similar reasons. After the Fall test
wave, which included 300 children, 21 children were dropped in subsequent test waves:
15 left the program or changed classrooms, five failed a subsequent PPVT-III screening
test, and one had a developmental language delay.

The final sample for the study of the science curriculum included 279 children;
51.3% were girls. Children’s ages at the beginning of the school year ranged from 36 to
59 months (M = 48.29; SD = 6.44). Ethnicity was reported for 99% of the sample,
indicating that 63% were Black or African American, 30% Hispanic or Latino, and 7%
other ethnicities. All children met federal criteria for enrollment in Head Start,
specifically that they are from low income families (which meet the criteria for low
income established by the federal government), or from families who qualify for public
assistance (Early Childhood Learning and Knowledge Center, 2009). For example,
federal guidelines for 2009 consider a family of four earning less than $22,050 annually
as having low income.

For the present study, a total of 210 children in classrooms whose teachers
consented to participate in the storybook reading portion of the study were assessed for
school readiness across the year according to the data collection procedures described
below. This represents approximately 46% of the student population in these 24
classrooms. The students ranged in age from 36 to 59 months at the start of the school
year (M=47.8, SD=6.3), they were 51% female, 61% African American, 30.5.2%
Hispanic or Latino, and 8.5% Asian, White, or Other. Of these 210 children, 90 (42.9%) spoke a language other than English, or in addition to English, 118 (56.2%) were monolingual English speakers, and two children’s files did not contain information on home language. The children who spoke a language other than English, or in addition to English, will be considered DLLs for the purpose of coding the child-level variable on language in this analysis.

Procedures

Read Aloud

Teachers were asked to read a novel storybook to the class in the manner that they typically would read a new storybook. The book which was provided to the teachers, *Edward the Emu* (Knowles, 1988), was chosen because of the number of words embedded in the story which were likely to be unfamiliar to preschool children, and thus provided the opportunity for vocabulary instruction. Additionally, this book was not part of the holdings of the county’s public library system, nor was it familiar to any of the teachers, so it seemed likely to be novel to the students.

Each teacher was recorded on digital video reading *Edward the Emu* (Knowles, 1988) either to her whole class or to a small group of students. Teachers were videotaped in the spring of 2009 and these videos were transcribed in the summer of 2009.

Transcription

Storybook reading sessions were transcribed from the classroom videos for further analysis of teacher language using SALT software (Systematic Analysis of Language Transcripts, Miller & Iglesias, 2008) by two trained graduate students.
Teacher utterances were broken into communication units which included either an independent clause alone or an independent clause and its modifiers (Systematic Analysis of Language Transcripts, Miller & Iglesias, 2008). Each transcription was reviewed at least twice for accuracy against the videotape of the storybook reading session by members of the study team. The transcription began at the point where the teacher told the students that it was “story time” or that she had a book that she was going to read to them. Most teachers discussed the author, illustrator, and parts of the book (front cover, back cover, spine) at the beginning of the storybook reading, and this was included in the coding of the transcript. Additionally, some teachers had a song or rhyme that they sang or recited with the students just before they began to read, and this was included in the transcription and coding as well, since it was considered part of the storybook reading session. The time teachers spent in the storybook reading session ranged from 6.82 to 28.12 minutes ($M=13.81; SD=5.61$).

**Measures**

**Child Outcomes**

To assess children’s vocabulary growth, two measures of oral language were used, and each was collected at the beginning and at the end of the 2008-2009 school year. Both measures were administered by trained graduate students. The first measure was the Peabody Picture-Vocabulary Test (PPVT–version 3) (Dunn & Dunn, 1981), a standardized measure of receptive vocabulary. The PPVT-III is a highly reliable assessment instrument, with an internal consistency of Alpha = .92 to .98 (median: .95), Split-half: .86 to .97 (median: .94), Alternate-form .88 to .96 (median: .94), and Test-retest .91 to .94 (median: .92). The validity of the PPVT-III has been established using
correlations with several established measures of oral language, including the OWLS Listening Comprehension scale (correlation = 0.69) and the OWLS Oral Expression scale (correlation = 0.74), as well as with measures of verbal ability such as the WISC-III VIQ (correlation = 0.91), KAIT Crystallized IQ (correlation = 0.89), and K-BIT Vocabulary (correlation = 0.81) (Pearson, 2010). Additionally, the PPVT-III has been found to be culturally fair and appropriate for use with African-American populations (Washington & Craig, 1999). For purpose of this analysis, the PPVT W-ability scores were used. These scores are based on a 1PL IRT (or Rasch) model, so the scores were thought to be most comparable (in terms of interpretation) to the second measure of oral language, the Learning Express, which also uses ability scores as described below.

With children whose families live in poverty, the PPVT is sometimes not sensitive enough to capture vocabulary growth, so the vocabulary subscale of the Learning Express (LE) (McDermott, Angelo, Waterman, & Gross, 2005) was also used to assess vocabulary. The vocabulary subscale of the LE is similar in design to the PPVT, but is structured to be more sensitive to the vocabulary growth of children who enter preschool with more limited vocabularies. The LE is based on a 2PL IRT model, and in addition to measures of receptive vocabulary, the LE also assesses expressive vocabulary. The reliability of the LE ranges from .93-.98 across subscales. The validity of the LE vocabulary subscale has been established using correlations with the OWLS Oral and Written Language Scales (correlation = 0.61) and with the PPVT-III (correlation = 0.69).

Both the PPVT and LE assess children’s receptive vocabularies by having the examiner state a word while presenting the child with a panel of four pictures, from which the child needs to point to the picture that represents the given word. The types of
words assessed by the PPVT include nouns, verbs, and attributes. The types of words assessed by the LE include nouns, verbs, and categorical words in the receptive measure, as well as nouns and verbs in the expressive measure (McDermott, Fantuzzo, Waterman, Angelo, Warley, Gadsden, & Zhang, 2009).

Data Coding: Word-level Coding

Transcripts were coded and analyzed using a researcher-developed coding protocol (see Table 1 in Appendix C) which describes each element of the storybook reading interaction and provides details about the behaviors that are considered part of that element. Transcripts were analyzed using SALT to quantify all of the elements of each storybook reading interaction.

Vocabulary strategies. For vocabulary strategies, each teacher received a score for the number of definition strategies, contextual strategies, and orthographic/phonological strategies used, and these scores were summed to obtain a total vocabulary strategy score (VocStrat). The elements for analysis related to vocabulary instruction strategies were chosen mainly based on the work of Whitehurst et al. (1988), Dickinson and Smith (1994), Beck and McKeown (2001), and Silverman (2007a).

Code switching. The teacher’s density of second language use during the storybook reading session was captured using a code switching (CS) code, which is one of the standard measures available with the SALT package. Code switching is operationalized, for the purpose of this study, as the number of non-English words used by the teacher during the storybook reading session.
Words instructed. Additionally, each time a particular word was instructed, the utterance was coded to indicate which word was instructed so that a total frequency of instruction for each word could be obtained, along with the total number of word instruction interactions across all words, for the purpose of answering research question #2. These interactions were further analyzed to determine the number of different word types taught, as well as the total number of word instruction interactions. For example, if a teacher taught the word crept twice, the word identical once, and the word reside four times, this would represent three word types taught, but seven word instruction interactions.

Word utility. Words taught were further divided into three categories representing the utility of the word for instruction following a process similar to the system of Tiers 1, 2, and 3 outlined in Beck, McKeown and Kucan (2002). In the work of Beck et al., Tier 1 words are “‘the most basic words. …words in this category rarely require instructional attention to their meanings in school” (p.8). Tier 2 words are “words that are of high frequency for mature language users and are found across a variety of domains” (p.8), and Tier 3 words are “words whose frequency of use is quite low and often limited to specific domains” (p.8). For the purpose of this study, words were divided into three categories as well: frequent, high utility, and low utility. Table 2 in Appendix C provides information on all words instructed, and their classification according to this system. The first group was words which were considered frequent words. Any word which was instructed and was also on the list of the 3000 most common words according to the Educator’s Word Frequency Guide (Zeno, Ivens, Millard, & Duvvuri, 1995) was categorized as a frequent word. Additionally, the words teeth and zoo were added to this
list of frequent words. The word *teeth* was categorized as *frequent* because brushing teeth is part of the daily routine of the Head Start classrooms in the study, and therefore should have already been familiar to the students. The word *zoo* was also categorized as *frequent* because the discussion in classrooms across the Head Start centers in the study indicated that a field trip to the zoo was an activity common to these centers. The words that were classified as low utility included proper names (*Edward, Edwina, Edward_the_Emu*) and words whose use was limited to specific domains (*lioness, Mohawk, ostrich, and seal*). Two words that might normally be considered *low utility* words, *emu* and *zookeeper*, were classified as *high utility* for the purpose of this study because they were essential to story comprehension for the chosen book.

*Inter-rater reliability.* In order to calculate inter-rater reliability, 26% of the transcripts were coded independently by two members of the study team. This resulted in a Cohen’s Kappa of .81, which is considered to be very substantial agreement according to the benchmarks provided by Landis and Koch (1977). To control for drift, the last transcript to be coded (by the same two members of the research team) was assessed for inter-rater reliability as well. This resulted in a Cohen’s Kappa of .85, which is in the range of “almost perfect” agreement as defined by Landis and Koch (1977, p. 165).

*Data Coding: Transcript-level Coding*

Analyses of language complexity including mean length of utterance (MLU), type/token ratio, and the ratio of uncommon words to common words (ZenoRatio), with common words defined by the Educator’s Word Frequency Guide (Zeno, Ivens, Millard, & Duvvuri, 1995), were also completed for each transcript.
**Language Quantity.** MLU provides a measure of the quantity of teacher language by providing the average length of each utterance, where an utterance is defined as a communication unit. A communication unit includes either an independent clause or an independent clause and its modifiers (Systematic Analysis of Language Transcripts, Miller & Iglesias, 2008).

**Language Quality.** Quality of teacher language is operationalized using both the type-token ratio and the researcher-created Zeno ratio. The type-token ratio measures language complexity by calculating a ratio of the number of different words (types) in a sample divided by the number of total words (tokens) in that sample. For example, in the sentence “The dog chased the cat,” there are five tokens (five words), but only four types (unique words) because the word “the” is used twice. The Zeno ratio variable measures language complexity by calculating a ratio of the number of uncommon words (those not contained in the first 3000 words of the Educators Word Frequency Guide), to the total number of words in the transcript. Once the list of words which were outside the 3000 Zeno words was extracted from each transcript using SALT software, the list of “uncommon” words from each transcript was further refined. Specifically, proper names, contractions for common words (such as “gonna” for “going to”), and simple derived forms of words on the Zeno list (such as “looking”, which is a derived form of “look”) were excluded. The remaining words from each transcript were considered to be the uncommon words for the purpose of this analysis. For both the type-token ratio and the Zeno ratio, a higher number indicates more complex language.

**Reading group size.** Prior research (Klesius & Griffith, 1996; Lonigan, et al., 1999; Morrow & Smith, 1990) has suggested that reading group size may influence a
teacher’s ability to implement dialogic reading strategies, so a group size measure
(GrpSize) was included in the data coding as well. This group size variable indicated the
number of children who participated in the storybook reading activity.
Chapter 4: Results

Teachers’ Vocabulary Instruction Strategies and Language Complexity

The first research question asked what vocabulary instruction strategies Head Start teachers use when reading a storybook to their students, and what level of language complexity is found in these storybook reading interactions. The number of vocabulary instruction strategies varied greatly among the teachers, with some teachers doing no vocabulary instruction whatsoever, while others instructed many words using a variety of strategies. Teachers’ language tended to be relatively uncomplicated, and very few teachers used non-English words as part of the storybook reading interaction.

Vocabulary strategies. Table 3 in Appendix C shows the vocabulary instruction strategies used by each teacher. The teachers’ vocabulary instruction strategies were grouped into three categories: definition, contextualization, and orthographic/phonological, based on previous work describing teachers’ vocabulary instruction strategies by Whitehurst et al. (1988), Dickinson and Smith (1994), Beck and McKeown (2001), and Silverman (2007a). The total number of vocabulary strategies used by each teacher ranged from 0 to 60 incidents of strategy use ($M=12.10; SD=12.30$). Most of the strategies used were those considered to be definition strategies ($M=9.31; SD=9.24$). Far fewer contextualization strategies ($M=1.98; SD=3.17$), or orthographic/phonological strategies ($M=0.71; SD=1.78$) were used by this group of teachers. Five of the teachers used no vocabulary instruction strategies of any type when reading.
While 5 (22%) of the teachers in the sample did not provide any vocabulary instruction during storybook reading, those who did provide incidental or explicit vocabulary instruction tended to rely more heavily on definition strategies (used by 18 teachers/ 78% of the teachers) than on contextualization strategies (used by 10 teachers/43% of the teachers) or orthographic/phonological strategies (used by 5 teachers/22% of the teachers). The most popular strategy, in terms of the number of teachers using the strategy, was *pointing to a picture* of the word (used by 14 teachers/61% of the teachers), followed by *acting out the meaning* of the word (used by 13 teachers/57% of the teachers). *Asking children if they knew the meaning* of the word was done by approximately half of the teachers (12 teachers, 52%), as was *providing the definition of the word* (11 teachers, 48%). Less popular definition strategies included *providing a recast* (9 teachers, 39%) and *giving a synonym* (8 teachers, 35%).

Of the contextualization strategies used, *connecting the word to the illustration* was the most popular, and was used by 8 (35%) of the teachers, *connecting the word to students’ personal experience* was done by 7 (30%) of the teachers, and *connecting the word to its use in the book* was done by only 2 (9%) of the teachers. The orthographic/phonological strategies were used even less frequently, with less than one quarter of the teachers ever using these strategies during this book reading. Among teachers who did use orthographic/phonological strategies, asking the students to *say the target word aloud* was the most popular strategy, with 3 (13%) of the teachers doing this at least once, followed by *discussing either the written or spoken form of the word*, each of which was done by only 2 (9%) of the teachers.
Code switching. Another strategy sometimes used by teachers who are bilingual and who have DLL students in their classrooms is explaining things to children in their home language to help facilitate children’s comprehension (Sert, 2005). The number of words in the transcript that were spoken in another language by the teacher, which happened to be Spanish in all of the classrooms in this study, ranged from 0 to 17 words per transcript \((M=1.42; SD=4.22)\). These words tended to be interspersed with English, rather than whole sentences spoken in a second language. It is noteworthy that while the majority of the children identified as DLLs in this study had Spanish as their first or second language, many others did not. The data provided by Head Start, based on parent reports, indicated that of the children classified as DLLs for the purpose of this study, 69 spoke Spanish, 15 spoke a Caribbean language or dialect, four reported the child’s language as “other”, one reported speaking an Asian language, and one reported speaking a Middle Eastern language. A few of the children were reported to have a language other than English as both their first and second language.

Language complexity. This first research question also focused on teachers’ language complexity, which was measured along two dimensions, quantity and quality. Table 4 in Appendix C shows the measures of quantity and quality of language for each teacher. Teachers’ quantity of language was captured in the mean length of utterance (MLU), which varied from approximately 4 to 7 words per utterance, with a mean of around 6 words per utterance. This MLU seems to be in line with that of child-directed speech (adults talking to children), which was reported by Newport, Gleitman, and Gleitman (1977) to be around 4.24 (as opposed to that of adults talking to adults, where a MLU of 11.94 was reported). The teachers’ average MLU is likely slightly higher in this
study than that reported by Newport et al. due to the nature of the book reading interaction. Some of the utterances reflect the written text, which tends to have a longer MLU than spoken language.

Teachers’ quality of language was quantified by both the type/token ratio (a measure of lexical variety) and the Zeno ratio (a measure of the use of uncommon words). The type/token ratio, with a mean of 0.24 and a range of 0.15 to 0.33 would be considered relatively low for adult spoken language (Retherford, 2000), but the repetitive nature of the storybook text and the length of the transcripts were the likely cause of the depressed range of this measure. The Zeno ratio reflects that very few uncommon words - those outside of the 3000 most common words in the Educator’s Word Frequency Guide (Zeno, Ivens, Millard, & Duvvuri, 1995) - were spoken by the teachers in these storybook reading interactions ($M=0.09$, range $=0.03-0.11$). In most transcripts, the majority of these uncommon words came from the book (only 13 of the 45 words that were instructed, as seen in Table 2 in Appendix C, were considered common words).

*Group size.* Because of research indicating that group size can influence teachers’ ability to implement certain vocabulary instruction strategies during storybook reading (Klesius & Griffith, 1996; Lonigan, et al., 1999; Morrow & Smith, 1990), group size was examined as well. The reading groups ranged from five to 16 students ($M=10.22$, $SD=3.62$). According to Justia.com Laws & Regulations (2011), Head Start centers that serve predominately 4- and 5-year old children must maintain an average class size ranging from 17-20 children per class (with two staff members – a teacher and an aide, or two teachers). Head Start centers serving predominantly 3-year old children must maintain slightly smaller classes, with an average of 15-17 children per class. Given these class
size guidelines, teachers who read to fewer than 10 students were classified as reading to a small group for the purpose of this study, while teachers who read to 10 students or more were classified as reading to the whole class. Using this classification system, half of the teachers ($N=11$) in this study read to the whole class and half of the teachers ($N=11$) read to a small group, and one teacher had missing data for this item, and the size of the group could not accurately be determined from the video footage.

*Instructed Words and Instructional Strategies*

The second research question asked how frequently Head Start teachers use vocabulary instruction strategies while reading a storybook to their students (how many words are chosen for instruction, and how many instructional strategies are used). In general, teachers tended to make good choices in terms of words for instruction, however the number of strategies used, as well as the strategies that were chosen, were not reflective of best practices in vocabulary instruction (Beck et al., 2002; Johnson, 2001). The words chosen for instruction, along with the number of times each word was instructed, by teacher, are shown in Table 5 in Appendix C. Across all teachers, the total number of high utility words chosen for instruction (31) was noticeably larger than the total number of either frequent words (15) or low utility words (7). Additionally, of the words chosen for instruction, far more teachers taught high utility words (i.e., *bore*, which was taught by 11 teachers) than taught the most commonly instructed frequent words or low utility words (*jump* and *seal*, which were each taught by 3 teachers).

On a per-teacher basis, Table 6 in Appendix C summarizes the number of words instructed by their designated value for instruction (frequent, high utility, or low utility
On the whole, teachers chose to provide instruction on more words that would be considered high utility words ($M=3.85$, $SD=3.84$), than frequent words ($M=.97$, $SD=1.03$) or low utility words ($M=0.40$, $SD=0.75$). On average, they taught approximately four times more high utility words than frequent words, and almost ten times as many high utility words as low utility words.

In order to facilitate the analysis of teacher vocabulary instruction strategies, two ratios were created, based on words instructed, for use in the analyses of the impact of teachers’ vocabulary instruction strategies and language complexity on student outcomes, which is described below. Descriptive information on both of these ratios is provided below and in Table 7 in Appendix C.

*High utility ratio.* The ratio of high utility words to total words instructed (HU ratio) provides a measure of the density of high utility words chosen for instruction. For example, a teacher who instructed four high utility words out of five total words instructed during the course of storybook reading would have a HU ratio of .8, while a teacher who instructed four high utility words out of 10 total words instructed would have a HU ratio of .4. In this way, teachers who made better word choices for instruction, regardless of the overall number of words taught, had High Utility scores that were weighted more strongly in the analysis, since best practices in vocabulary instruction during storybook reading suggest the teaching of high utility words (Beck et al., 2002; Johnson, 2001), rather than frequent or low utility words. The HU ratio ($M=0.6$, $SD=0.36$) for the sample of teachers in this study indicated that on average 60% of the words chosen for instruction across all teachers were of high utility.
Vocabulary strategy ratio. A second ratio, the ratio of vocabulary instruction strategies to total words instructed (VocStrat ratio), was also created. This ratio provides an indication of the average number of strategies used to instruct each word. Current theory on the best practices in vocabulary instruction (Beck et al., 2002, Johnson, 2001) suggests that teachers should use multiple strategies to instruct vocabulary words to help students learn and retain the meaning of each word. The VocStrat ratio indicates the average number of instructional strategies per word used by each teacher. While each teacher used an average of just over two strategies per word ($M=2.13$, $SD=2.02$), some teachers used as many as eight strategies to teach a single word.

Another way to examine the nature of the words teachers chose for instruction is to look at the parts of speech of the words chosen for instruction. Of the 52 total words chosen for instruction across all teachers, 22 were nouns, 22 were verbs, seven were adjectives, and one was an adverb. This skew toward nouns and verbs is not unexpected, since nouns (because they are more concrete) are somewhat easier to instruct, and verbs can often be demonstrated. The descriptive information on parts of speech of instructed words, on a per-teacher basis, is provided in Table 8 in Appendix C.

Impact of Teacher Strategies and Language Complexity on Student Outcomes

The third research question concerned teachers’ vocabulary instruction strategies and complexity of language during storybook reading and how these strategies and language complexity relate to outcomes in student vocabulary over the course of a year within Head Start classrooms.
The original study design called for data to be analyzed using hierarchical linear modeling (HLM). HLM was chosen due to the hierarchical nature of the data in this study, where students are nested within classrooms, and the actions of the teachers in those classrooms were thought to affect student outcomes. The empirical evidence in this study, however, suggested otherwise. Analysis of the data indicated that between-class variation (variation at the teacher level), as measured by the intra-class correlation in hierarchical linear modeling, accounted for less than 2% of the variance in student outcomes for each of the outcome measures. Further details about the calculation of the intra-class correlation are found in Appendix D.

Given that this empirical evidence suggested very minor teacher-level variance, the data were analyzed using hierarchical multiple regression instead of hierarchical linear modeling. A correlation matrix of all the predictors and outcomes in the model is provided in Appendix E. Among the student-level variables, the highest correlations are among the fall and spring PPVT and the fall and spring LE measures, with correlations of 0.66 within measures (for both measures) and ranging from 0.67 to 0.71 across measures (\textit{p}<.01 for all correlations). Student age was also significantly correlated with all outcome measures, with correlations ranging from 0.36 to 0.44 (\textit{p}<.01 for all correlations).

Among the teacher-level variables, the highest significant correlations (at the \textit{p}<.01 level) were between teachers’ MLU and code switching (-.57), the type/token ratio and the Zeno ratio (.63), and the High Utility ratio with the Vocabulary Strategy ratio (.71). The only correlations that were significant at the \textit{p}<.01 level between teacher and student-level variables were the positive correlation between students’ DLL status and teachers’
type/token ratio (.19), and the negative correlation between students’ DLL status and teachers’ ratio of vocabulary strategies to words instructed (-.21).

**Predicting outcomes.** To further examine these relationships, the Spring PPVT and LE outcomes were modeled as outcomes in two separate analyses, while controlling for students’ fall performance on the PPVT or LE (prior vocabulary knowledge), dual language learner status (dummy coded 1 for dual language learners and 0 for English speakers based on home language data provided by Head Start), age, and gender. These student-level characteristics were included in the model to control for their possible influence on vocabulary outcomes, since previous research has indicated that these variables can influence children’s vocabulary growth (Huttenlocher et al., 1991; Collins, 2010; Robbins & Ehri, 1994). Exploration of the data revealed that students identified as DLLs began the year with higher scores on the PPVT ($M=65.83$, $SD=1.28$) than their monolingual counterparts ($M=63.04$, $SD=1.05$), yet at the end of the year, the DLLs’ PPVT scores ($M=68.33$, $SD=1.18$) were virtually identical to those of their monolingual peers ($M=68.95$, $SD=0.9$), indicating that the gains made during the school year by DLLs on the PPVT measure were lower than those of the monolinguals, although these differences were not statistically significant. The LE measure tells a different story, with students identified as DLLs starting the year with lower scores ($M=180.87$, $SD=5.90$) than their monolingual counterparts ($M=183.93$, $SD=4.23$), yet ending the year with virtually identical scores to those of their monolingual peers (for DLLs, $M=212.34$, $SD=5.19$; for monolinguals, $M=212.99$, $SD=3.60$), indicating that the gains made by DLLs on the LE measure were greater than those of the monolinguals, although not statistically significantly different. A similar examination of the data by gender showed
that for the PPVT measure, males ($M=64.54$, $SD=1.21$) started the year with scores approximately a half a point higher females ($M=64.06$, $SD=1.12$) and ended the year slightly higher as well (males: $M=69.32$, $SD=1.07$; females: $M=68.3$, $SD=0.95$). On the LE measure, the males also began and ended the year with higher scores than the females (LE fall score for males: $M=185.12$, $SD=5.58$; for females: $M=180.55$, $SD=4.33$; LE spring scores for males: $M=214.39$, $SD=4.71$; for females: $M=210.96$, $SD=3.93$). None of these differences in scores, by gender, were statistically significantly different.

Based on current best practices in vocabulary instruction (Beck et al., 2002), I predicted that students in classes with teachers who use a greater number of vocabulary strategies per word instructed, and who had a higher ratio of high utility words to total words instructed, would show the greatest change in vocabulary over the course of the school year. Additionally, teacher language complexity, in the form of MLU, type-token ratio, and Zeno ratio, was thought to be predictive of changes in children’s vocabulary outcomes. The possibility that children who began the school year with larger vocabularies would have outcomes that would be related to teacher strategies and language complexity has been reported in previous studies (Collins, 2010; Coyne et al., 2009; Robbins & Ehri, 1994; Senechal et al., 1995), so this possibility was tested by examining cross-level interactions between prior vocabulary knowledge and the various measures of teacher language and vocabulary instruction.

**Hierarchical Regression Analysis for PPVT Outcome**

In order to examine the contribution of student-level variables versus teacher-level variables, hierarchical regression was used to enter groups of variables in three blocks. The student-level variables: fall PPVT score (representing prior vocabulary
knowledge), DLL status, age, and gender were entered in the first block. These four variables accounted for 49% of the variance in the spring PPVT score. The omnibus test was statistically significant, $R^2 = .490$, $F(4, 191) = 45.93, p < .001$.

In the next block, the teacher-level variables were entered. These included MLU, type-token ratio, Zeno ratio, code switching, the ratio of vocabulary strategies to total words instructed (VocStratRatio), the ratio of high utility words to total words instructed (HUratio), and the reading group size. The addition of these teacher-level variables accounted for an additional 4.3% of the variance in spring PPVT score, after controlling for the student-level variables. The additional variance accounted for was statistically significant, $\Delta R^2 = .043$, $F(11, 184) = 19.11, p = .021$.

In the third block, cross-level interactions between prior vocabulary knowledge and the various measures of teacher language and vocabulary instruction were added to the model to test for the possibility that students would be differentially impacted by teacher strategies or teacher language based on their vocabulary level (as measured by the PPVT) at the start of the year. The addition of these cross-level interactions accounted for an additional 4.5% of the variance in spring PPVT score, after controlling for the student-level and teacher-level variables entered previously. The additional variance accounted for was statistically significant, $\Delta R^2 = .045$, $F(16, 179) = 15.37, p = .002$.

There were three student-level variables, one teacher-level variable, and two interactions that were statistically significant predictors of spring PPVT scores. Prior vocabulary knowledge, as measured by the fall PPVT score, was a statistically significant predictor of spring PPVT score, $b = 0.52$, $t = 10.38, p < .001$. For a one point increase in a student’s fall PPVT W score, the spring PPVT W score would be expected to increase
by 0.52 points, after controlling for the other student- and teacher-level variables in the model. Age was also a statistically significant predictor of spring PPVT score, $b = 0.34$, $t = 3.90$, $p < .001$. For a one month increase in age, the spring PPVT W score would be expected to increase by 0.34 points, after controlling for the other student- and teacher-level variables in the model. A student’s status as a dual language learner (DLL) was also a statistically significant predictor of spring PPVT score, $b = -3.12$, $t = -2.87$, $p = .002$. Students who were DLLs would be expected to have a spring PPVT W score that was 3.12 points lower than that of their monolingual English peers, after controlling for the other student- and teacher-level variables in the model. In the teacher block, only code switching was a significant predictor of spring PPVT scores, $b = -0.42$, $\beta = -0.19$, $t = -2.34$, $p < .020$. For a one SD increase in code switching, the spring PPVT W score would be expected to decrease by .19 SD units, after controlling for the other student- and teacher-level variables in the model. These findings are shown in Table 10 in Appendix C.

Code switching would presumably be beneficial only for DLLs (not for monolingual English students) in terms of vocabulary outcomes, and then only if the child’s language was Spanish (the only language used in code switching by the teachers in this study). When separate regressions based on students’ DLL status were run for the variables in this analysis, code switching was not a significant predictor of student vocabulary outcomes for either group alone, even though it was a significant predictor in the model for the entire sample of students. An interaction term between code switching and the DLL variable, when tested in the overall model, was also not significant. Even though this relationship did not reach statistical significance, a graph was created to
visually represent the relationship between teachers’ code switching, students’ DLL status, and student vocabulary gains over the course of the year, represented by the residuals of students’ PPVT scores. The graph shows that the relationship between teachers’ code switching and students’ gains on the PPVT is relatively neutral for DLLs, yet this relationship is negative for monolingual English-speaking students. This graph is shown in Figure 2.

The significant overall relationship of code switching to student vocabulary outcomes on the PPVT outcome measure may have been due to a mismatch between the language used for code switching and the language of the students in the classroom. Of the 23 teachers participating in this study, only two did a noticeable amount of code switching. The first of these two teachers seemed to switch to Spanish incidentally throughout the story, sometimes to ask questions, and sometimes just inserting one Spanish word in a sentence, such as “y” instead of “and” or “pero” instead of “but”. Of the six children who were assessed for school readiness from this teacher’s classroom as part of this study, only one was identified as a DLL, and that child’s first language was reported by the parents to be “Caribbean”. The second teacher who did a noticeable amount of code switching switched to Spanish on several occasions to explain the story/vocabulary to the students. For example, this teacher said, when describing the meaning of the word *emu*, “un avestruz grande” (a big ostrich). Although an emu is technically not an ostrich, they do look similar and perhaps the word ostrich is a more familiar word in Spanish, as it is in English. This teacher also used Spanish to give instructions related to classroom management (sit down, be quiet, etc.). Of the nine children who were assessed for school readiness from this teacher’s classroom as part of
this study, all nine were identified as DLLs, eight of whom had Spanish reported as their first language, and one of whom had “other” reported as a first language and Spanish reported as a second language. Aside from this particular classroom, where the teacher’s use of code switching seemed more purposeful, the results of this study indicate that for this population of students and teachers the code switching (in Spanish) done by teachers was incidental rather than purposeful, and did not benefit the students, the majority of whom were not Spanish speakers.

Two of the interactions entered in the third step were statistically significant, as shown in Table 9 in Appendix C. To determine the unique variance accounted for by each of these interactions, the regression was run again twice, once adding the Zeno Ratio x fall PPVT interaction alone as a 4th block (with all other interactions entered in the third block), and once adding the VocStrat Ratio x fall PPVT interaction as a 4th block (with all other interactions entered in the third block).

The Zeno Ratio x fall PPVT interaction accounted for 2.6% of the variance in spring PPVT score, after controlling for the student-level and teacher-level variables and other interactions. The additional variance accounted for was statistically significant, $\Delta R^2 = .026$, $F(16, 179) = 15.37, p = .001$. To better understand this interaction a graph was created that divided the fall PPVT scores into three groups: scores that were more than one standard deviation above the grand mean of fall PPVT scores, scores that were within one standard deviation (SD) of the grand mean of fall PPVT scores, and scores that were more than one standard deviation below the grand mean of fall PPVT scores. The graph of this interaction is shown in Figure 3. The graph suggests that for those students who had relatively lower fall PPVT scores (greater than one SD below the mean), the
teacher’s use of a higher ratio of uncommon words was positively related to spring PPVT outcomes. For those students whose scores were within one SD of the mean of fall PPVT scores or more than one SD above the mean of fall PPVT scores, the relationship between the teachers use of uncommon words and spring PPVT outcomes was slightly negative.

The $VocStratRatio \times fall \ PPVT$ interaction accounted for 3.1% of the variance in spring PPVT score, after controlling for the student-level and teacher-level variables and other interactions. The additional variance accounted for was statistically significant, $\Delta R^2 = .031, F(16, 179) = 15.37, p < .001$. To better understand this interaction a graph was created that divided the fall PPVT scores into the same three groups as in the previous interaction, based on distance from the mean. A graph of this interaction is shown in Figure 4. The regression was re-run for each of these three groups to determine whether the slopes of the three lines shown in Figure 4 were significant. For those students who were within one SD of the mean of fall PPVT scores or more than one SD above the mean of fall PPVT scores, the teacher’s use of more vocabulary strategies per word instructed was virtually unrelated to spring PPVT outcome. For those students whose fall PPVT scores fell more than one SD below the grand mean of fall PPVT scores, there was a significant negative relationship between an increase in the number of vocabulary strategies used per word and the student’s spring PPVT outcome, $b = -3.03, t = -2.42, p = .023$.

**Hierarchical Regression Analysis for Learning Express Outcome**

A second hierarchical regression was run for the Learning Express (LE) outcome measure. The student-level variables: fall LE score (representing prior vocabulary knowledge), dual-language learner (DLL) status, age, and gender were entered in the first
These four variables accounted for 45.8% of the variance in the spring LE score. The omnibus test was statistically significant, $R^2 = .458$, $F(4, 188) = 39.75, p < .001$. In the next block, the teacher-level variables were entered. These included MLU, type-token ratio, Zeno ratio, code switching, the ratio of vocabulary strategies to total words instructed (VocStratRatio), the ratio of high utility words to total words instructed (HUratio), and the reading group size. The addition of these teacher-level variables accounted for an additional 4.0% of the variance in spring LE score, after controlling for the student-level variables. The additional variance accounted for was statistically significant, $\Delta R^2 = .040$, $F(11, 181) = 16.36, p = .047$. In the third block, cross-level interactions between prior vocabulary knowledge and the various measures of teacher language and vocabulary instruction were added to the model to test for the possibility that students would be differentially impacted by teacher strategies or teacher language based on their vocabulary level (as measured by the LE) at the start of the year. The addition of these cross-level interactions was not significant, $\Delta R^2 = .008$, $F(16, 176) = 11.31, p = .706$, so the reported coefficients are based on the second model (without the interaction terms). These findings are shown in Table 12 in Appendix C.

There were two student-level variables and one teacher-level variable that were statistically significant predictors of spring LE scores. Prior vocabulary knowledge, as measured by the fall LE score, was a statistically significant predictor of spring LE score, $b = 0.51, t = 10.01, p < .001$. For a one point increase in a student’s fall LE score, the spring LE score would be expected to increase by 0.51 points, after controlling for the other student- and teacher-level variables in the model. Age was also a statistically significant predictor of spring LE score, $b = 1.17, t = 2.91, p = .004$. For a one month
increase in age, the spring LE score would be expected to increase by 1.17 points, after controlling for the other student- and teacher-level variables in the model. Only one teacher-level variable, the Zeno ratio, was a statistically significant predictor of spring LE score, $b = -567.82, \beta = -0.26, t = -3.09, p = .002$. For a one SD increase in the Zeno ratio, the spring LE score would be expected to decrease by .26 SD units, after controlling for the other student- and teacher-level variables in the model. These results are shown in Table 11 in Appendix C.

Because there was a significant main effect for the Zeno ratio on the LE outcome measure, and there had been an interaction between the Zeno ratio and the fall PPVT score on the PPVT outcome measure, I decided to graph the interaction between the Zeno ratio and the fall LE score in the same manner as was done for the PPVT outcome measure to test whether the same interaction effects existed for the LE measure, even though this interaction was not statistically significant. The graph of this interaction is shown in Figure 5. This graph indicates that although the overall effect of teachers’ use of more uncommon words on children’s vocabulary outcomes is negative, for those students who begin the year with the lowest level of vocabulary, the relationship appears to be positive, just as it was for the PPVT outcome measure.
Chapter 5: Discussion

The practice of storybook reading, a mainstay of preschool programs, is a potential vehicle for increasing children’s vocabulary (NELP, 2009; Bus, van Ijzendoorn, & Pellegrini, 1995). While many studies have investigated the relationship of the quantity of storybook reading to oral language outcomes, fewer studies have examined the more qualitative details of those interactions to determine the particular actions, strategies, behaviors, or techniques of the reader in facilitating children’s vocabulary growth (Bus, et al., 1995; Scarborough & Dobrich, 1994a; Whitehurst et al., 1988). The present study sought to categorize teachers’ strategies according to best practices for vocabulary instruction and to quantify the frequency of use of these strategies as well as quantifying the language complexity of teachers during storybook reading in an effort to determine which strategies or elements of teacher language were related to students’ vocabulary outcomes over the course of a year in Head Start preschool classrooms.

Overall, teachers in this study tended to make relatively good choices in terms of which words to instruct, but the strategies used to teach those words, and the number of words chosen for instruction was often not what would be considered optimal with regard to best practices in vocabulary instruction (Beck et al., 2002; Johnson, 2001). In terms of language complexity, teachers’ language quality, specifically the use of uncommon words, was differentially related to the prediction of students’ vocabulary outcomes. Teachers’ use of uncommon words had a slightly negative relationship to vocabulary outcomes for students who began the year with average or high vocabulary scores, while teacher’s use of uncommon words predicted higher end-of-year vocabulary outcomes for
students who began the year with the lowest vocabulary levels. The implications of these results for teacher preparation are discussed further below.

**RQ #1: Teachers’ Vocabulary Instruction Strategies and Language Complexity**

*Vocabulary instruction strategies.* The teachers’ storybook reading interactions were examined in the context of existing best practices in vocabulary instruction (Beck et al., 2002; Johnson, 2001) which were considered to be appropriate for use on an incidental basis in preschool settings during storybook reading. All of the strategies in the coding scheme were used by at least some of the teachers, with definition strategies being used more frequently than contextualization or orthographic/phonological strategies. The teachers’ vocabulary instruction strategies were grouped into three categories: definition, contextualization, and orthographic/phonological, based on previous work describing teachers’ vocabulary instruction strategies by Whitehurst et al. (1988), Dickinson and Smith (1994), Beck and McKeown (2001), and Silverman (2007a). While five of the teachers in the sample did not provide any vocabulary instruction during storybook reading, those who did provide incidental or explicit vocabulary instruction relied more heavily on definition strategies than on contextualization strategies or orthographic/phonological strategies.

When reviewing the strategies in terms of best practices (Beck et al., 2002; Johnson, 2001), some of the most explicit strategies, such as providing a definition or providing a synonym, were used by fewer teachers, while some of the less explicit strategies, such as pointing to a picture, were used by more teachers. Additionally, the strategy of asking children if they knew the meaning of a word was used by more than
half of the teachers, and this particular strategy is considered to be more distracting than helpful for many children, unless it is followed up with explicit discussion and instruction (M. McKeown, personal communication, May 2, 2010). The contextualization strategies, which provide more scaffolding for learning word meanings, were used far less frequently than the definition strategies, with only one third of the teachers using any of these strategies. Orthographic/phonological strategies, which are considered helpful in creating auditory and visual representations of words (Beck & McKeown, 2001), were used even less (with only two or three teachers using each of these strategies).

It seems that the majority of teachers (78%) do provide incidental vocabulary instruction while reading storybooks, thereby using this common classroom practice as a possible vehicle for increasing children’s oral language. Their choice of strategies however was not always what would be considered optimal in terms of best practices (Beck et al., 2002; Johnson, 2001), nor was the number of words chosen for instruction, or the amount of instruction provided for each word. The teachers as a whole tended to make good choices about which words to instruct, with many teachers choosing high utility words, as will be described in the discussion of the second research question.

**Code Switching.** The practice of code switching was included in this study to examine whether teachers who had DLLs in their classes might translate some of the more challenging vocabulary in the story to help their students understand these words. As discussed earlier, the DLLs in this study showed mixed results in terms of vocabulary gains over the course of the year, with DLLs having lower residuals on the PPVT measure, but higher residuals on the LE measure. The significant main effect for code switching on the PPVT outcome measure cannot be interpreted in isolation, because the
relationship of code switching to student outcomes would be expected to be different for
the population of students who were DLLs than for those who were monolingual English
speakers. Presumably, code switching would only benefit DLLs, and then only when the
child’s language was Spanish (since that was the only language used in code switching by
the teachers in this study), but when the regression analyses were run with the students
separated into groups by language status, code switching was not a significant predictor
of vocabulary outcomes for either group on either outcome measure. When the
relationship between teachers’ code switching and student vocabulary gains over the
course of the year (as measured by the residual scores on the PPVT outcome measure)
was examined graphically with regard to students’ status as DLLs or monolingual
English speakers, the graph (shown in Figure 2) indicated that the relationship between
teachers’ code switching and students’ gains on the PPVT is relatively neutral for DLLs,
while this relationship was negative for monolingual English-speaking students. With the
exception of one teacher, the code switching done by teachers in this study was more
incidental than purposeful, with occasional Spanish words interjected into utterances that
were predominantly English. These occasional Spanish words did not seem to provide
any sort of purposeful scaffolding for the Spanish-speaking DLLs. The relationship to
monolingual English-speaking students vocabulary gains appeared to be negative but this
relationship should be explored further in future studies.

**RQ #2: Frequency of Vocabulary Instruction and Words Chosen for Instruction**

The second research question examined the words teachers chose for instruction,
and the quantity of instruction that was provided for each word. While 22% of the
teachers in this study did not teach any words during the reading of this book, the rest of the teachers did provide a range of vocabulary instruction during the book reading, with from 1 word to 18 words being instructed. Prior research on vocabulary instruction suggests that it is best to teach a limited number of words (Beck et al., 2002), but teach them using multiple strategies (Beck et al., 2002; Johnson, 2001). Additionally, Beck et al. (2002) suggest teaching words which children are likely to encounter in academic language in multiple settings or words for which the child may understand the concept but lack the specificity that the target word provides (for example, if the child understands hate, he or she can be taught detest). Teaching high utility words is preferable to teaching words that are frequently used (words that are probably part of children’s existing vocabularies) or of limited utility (those which are specific to a particular domain of knowledge or subject area).

As a whole, the teachers in this study tended to choose good words to instruct (defined here a high utility words, rather than frequent or low utility words). On average, teachers chose to teach high utility words about 60% of the time they provided vocabulary instruction. The number of words instructed in a single storybook reading, however, was not always optimal.

Although the research on best practices (Beck et al., 2002) suggests teaching approximately three words as part of a storybook interaction, the number of words taught by teachers in this study ranged from zero to 18. Approximately half of the teachers taught more than five words during the storybook reading, while approximately 40% of the teachers taught two or fewer words during the storybook reading. In fact, the mean number of words instructed was almost twice the number recommended in the best
practices referenced above ($M=5.22, SD=4.67$). There were only two teachers who instructed three words during their book reading, and for those two teachers, only two of the three words they chose were high utility words (consistent with the overall rate of choosing high utility words across all teachers, which was about 60%). This relationship is depicted graphically in Figure 6. The large number of words chosen for instruction may have influenced the relationship between some of the teacher-level variables and student outcomes. For example, teachers who taught a greater number of uncommon words did not necessarily use more instruction strategies per word, and this may have influenced students’ acquisition of those words, as will be discussed further below.

*RQ #3: Relationship of Vocabulary Strategies and Language Complexity to Students’ Vocabulary Outcomes*

Results indicated that student age and prior vocabulary knowledge were positively and significantly related to vocabulary outcomes. Teachers’ use of uncommon words had a different relationship to student outcomes based on students’ prior vocabulary knowledge for both outcome measures, but only statistically significantly for the PPVT outcome measure (for the LE outcome measure, only the negative main effect for the use of uncommon words was significant). Additionally, for the PPVT outcome measure, a student’s status as a dual-language learner, the teacher’s ratio of vocabulary instruction strategies to words instructed, and the amount of code switching by the teacher were also negatively and significantly related to students’ vocabulary outcomes. Each of these findings is discussed in turn below.
Age. The significance of age as a predictor of vocabulary outcomes was not unexpected. Children’s vocabularies increase with age, and the number of words children learn per day has been estimated to be between 3 (Joos, 1964) and 20 (Miller, 1978), with a figure of 7 words per day being most commonly cited (Beck, et al., 2002), although these are all averages, which are influenced by other variables. The work of Hart and Risley (1995) has shown that children as young as age 3 can have large differences in vocabulary ability as compared to their same-age peers, and that these differences are highly correlated with socio-economic status (SES), with children living in poverty knowing fewer words than their more affluent peers. Given that all of the children were of relatively similar SES in this sample (all qualified for the Head Start program), one would expect that, in general, older children would have higher vocabulary ability than younger children, and the data reflect this.

Dual language learner status. A child’s status as a DLL was a significant predictor of their spring PPVT outcome, with children who were DLLs in this sample having a spring PPVT W score that was approximately one quarter of a standard deviation lower than their monolingual English-speaking peers. This was also not unexpected, since previous research has demonstrated that DLLs lag behind their monolingual peers in the acquisition of English vocabulary and specifically academic vocabulary (Snow, Burns & Griffin, 1998), some of which is assessed by the PPVT. Interestingly, this same finding was not present for the LE outcome measure. It is possible that the design of the LE, which has the goal of being able to measure vocabulary growth of children with lower vocabularies, enabled this measure to capture
more modest incremental growth among the DLLs in this sample than the PPVT measure was able to detect.

**Prior Vocabulary/Interaction Effects.** For the LE outcome measure, there was a positive, significant main effect for fall vocabulary score, and for the PPVT outcome measure, there was a significant interaction between fall PPVT score and both the Zeno ratio and the ratio of vocabulary strategies to total words instructed (VocStrat ratio). The positive main effect for prior vocabulary knowledge, found for the LE measure, was not unexpected. Previous research has shown that prior vocabulary knowledge is a good predictor of subsequent knowledge (Collins, 2005; Coyne, et al., 2009; Coyne, et al., 2004; Ewers and Brownson, 1999; Justice, et al., 2005; Robbins & Ehri, 1994; Senechal, et al., 1995). The significant interactions for the PPVT measure however, provide more interesting information about the relationship between teacher’s language and vocabulary instruction strategies and the vocabulary outcomes of children with diverse levels of vocabulary knowledge.

In this sample, the teachers’ ratio of vocabulary strategies used per total words instructed (VocStrat ratio) had a different relationship to students’ spring PPVT outcomes based on students’ fall PPVT scores. In classrooms where teachers used a higher number of strategies per words instructed, this practice was not significantly related to spring vocabulary outcomes (spring PPVT scores) for those students whose fall PPVT score was within one SD of the grand mean of fall PPVT scores or greater than one SD above the grand mean of PPVT scores. For those students whose fall PPVT scores were greater than one SD below the grand mean however, the teacher’s use of more vocabulary strategies per word was significantly negatively related to spring vocabulary outcomes.
This finding is contrary to what would be predicted by the best practices in vocabulary instruction, which suggest that using multiple strategies to explicitly teach vocabulary words is a positive instructional strategy (Beck et al., 2002; Johnson, 2001). Beck et al. (2002) also suggest that attempting to teach more than a few words in one lesson tends to have less positive results in terms of word learning. Further examination of the data revealed that the number of vocabulary strategies used per word instructed was significantly correlated with the total number of word tokens instructed (Pearson correlation = .285, p<.01), so those teachers who used a greater number of strategies per word also had more word instruction interactions overall. It is possible that this higher number of word instruction interactions, and not the number of strategies per word, is what was negatively related to the PPVT vocabulary outcome of those students who knew fewer words to begin with, and this relationship merits further investigation in future studies.

For this sample of preschool children and their teachers there was also a significant interaction between the students’ fall PPVT scores and the teacher’s use of more uncommon words (Zeno ratio). The interaction was such that for students in the high and middle groups, the Zeno ratio appeared to be slightly negatively related to vocabulary outcomes, while for student in the low group the Zeno ratio appeared to be slightly positively related to vocabulary outcomes. For the LE outcome measure, the Zeno ratio was negatively related to students’ vocabulary outcomes, with the significant main effect indicating that for an increase of one SD in the teacher’s Zeno ratio, the student’s spring LE score would be expected to decrease by approximately one quarter of a SD. Although the overall relationship between the teacher’s Zeno ratio and students’
vocabulary outcomes on the LE was negative, the data were examined to determine if the same interaction effect between prior vocabulary knowledge and teachers’ use of uncommon words was present for the LE outcome measure as in the PPVT outcome measure, and a similar pattern (though not statistically significant) was found. This interaction between teachers’ Zeno ratio and students’ vocabulary outcomes, though only statistically significant for the PPVT outcome, suggests that the teachers’ use of more complex language is differentially related to students’ vocabulary outcomes in this sample based on students’ initial vocabulary status. Students who begin the year with the lowest vocabulary levels exhibited a positive relationship between the teacher’s use of uncommon words and vocabulary outcomes, while average or higher vocabulary students had a more negative relationship between increased exposure to these uncommon words and their vocabulary outcomes.

Several possible reasons for this finding should be considered. First, there is no significant correlation between the teachers’ use of uncommon words and the number of vocabulary instruction strategies used per word (VocStrat ratio), suggesting that those teachers who used more uncommon words were not necessarily providing a high level of vocabulary instruction or scaffolding to assist children in learning those words. Additionally, there is a significant negative correlation between the teachers’ use of uncommon words and the ratio of high utility words chosen for instruction (HU ratio), indicating that those teachers who used more uncommon words also chose to teach relatively fewer high utility words. Mere exposure to novel, more uncommon vocabulary is often not enough for children to acquire that vocabulary (Senechal et al., 1995), and this may be what is indicated by these results, with the exception of those children who
began the year with the lowest vocabulary levels, for whom increased exposure to uncommon words was positively related to vocabulary outcomes. It is also possible that the uncommon words used by the teachers in this study were simply not uncommon enough to have a positive relationship to students’ vocabulary outcomes among the students who began the year with higher levels of vocabulary. This finding should be explored systematically in future studies to better determine why and how students outcomes differ based on teachers’ language complexity with respect to uncommon words.

_Code switching_. The practice of code switching was included in this study to examine whether teachers who had DLLs in their classes might translate some of the more challenging vocabulary in the story to help their students understand these words. With the exception of one teacher who did seem to use code switching in this way, there was no other evidence of the use of code switching for vocabulary instruction among the other 22 teachers.

Code switching by the teacher was, nevertheless, a significant predictor of students’ vocabulary outcomes on the PPVT measure, with a higher level of code switching associated with poorer vocabulary outcomes. This may be because the code switching measure was not capturing code switching that was meaningful in terms of providing scaffolding for DLLs to acquire English vocabulary, but rather was simply a reflection of the occasional Spanish words inserted somewhat randomly into the teacher’s discourse. This random type of code switching appeared not to affect the vocabulary residual scores of the DLLs during the course of the school year, but was negatively related to vocabulary residual scores of the monolingual English speakers over the course
of the school year, although the difference in this relationship of code switching to residual scores between the two groups (DLLs and monolingual English speakers) was not statistically significant.

The results of this study as a whole indicate that for this sample of Head Start teachers and their students, older children in general knew more words than younger ones across both outcome measures. The results also suggest that DLLs may lag behind their peers in terms of vocabulary acquisition, although this finding was only present for the PPVT outcome measure. Students’ PPVT outcomes had differing relationships to teachers’ vocabulary instruction strategies per word instructed based on students’ fall vocabulary level, with students who scored more than one SD below the mean on the fall PPVT being more negatively affected in terms of their spring PPVT outcome by teachers who used more instruction strategies per word. This is particularly disappointing in light of the fact that best practices (Beck & McKeown, 2002) suggest that more strategies per word should be helpful in terms of vocabulary acquisition. Teachers’ use of more complex language, as measured by the Zeno ratio, was differentially related to students’ vocabulary outcomes, with a negative relationship between the teachers’ use of more uncommon words and students’ vocabulary outcomes, except for those students who began the year with the lowest levels of vocabulary (although this relationship was only significant for the PPVT outcome measure). Prior research (Hart & Risley, 1995) suggests that exposure to more complex language is important in increasing the size of children’s vocabularies. Because teachers used relatively few uncommon words, it seems that the use of these words was only beneficial to those students with the most depressed vocabulary levels. Future studies should examine the variables and outcome measures
from this study with a relatively more economically advantaged population to eliminate
the range restriction on SES imposed by studying Head Start classrooms, especially in
light of the findings by Early et al. (2010) which suggest that children spend less time in
high level interactions such as storybook reading in preschools that serve low-income
students than they do in preschools serving more advantaged populations. This would
help researchers obtain a more comprehensive picture of teacher language complexity
and vocabulary instruction practices and their relationship to student vocabulary
outcomes.

Limitations of the Study and Directions for Future Research

This study examined correlations between teacher practices and student outcomes.
The findings should not be interpreted as implying causality. Experimental studies that
systematically manipulate variables such as strategy selection and number of strategies
should be done to clarify the exact nature of the relationship of these variables.

In terms of a single storybook reading providing an accurate “snapshot” of the
way in which these Head Start teachers read new storybooks to their students, this study
had several challenges. We asked teachers to read the book to their students “the way
you normally read a new book to your class.” Several teachers chose to read the book to
a small group of students while the other members of the class were engaged in a
different activity with the assistant teacher. Other teachers read the book to the whole
class, which made group size inconsistent. This may have presented some behavior
management challenges to teachers who chose to read to the whole class, which in turn
might have impacted their reading style. Additionally, some teachers reported that they
did not customarily read a whole book with the class in one sitting, but rather spread it
out over several days, beginning with a “picture walk” through the book on the first day, then asking predictive questions based on the pictures on the second day, and finally reading part, or all, of the book on a third day. In future studies, it may be beneficial to interview teachers about their reading practices before observing them so that these differences in style can be accommodated through multiple videotaped observation sessions if necessary.

Another limitation was the relatively small sample size, both in terms of teachers and students, which may have made it more difficult to detect significant interactions between student- and teacher-level variables. Additionally, sample size became an issue when conducting analyses of sub-groups (by DLL status, prior vocabulary knowledge, etc.), due to the loss of power. The sample size was also relatively modest for using hierarchical linear modeling, which would have been the preferred technique for analyzing this nested data, as discussed earlier.

The lack of detailed knowledge about the DLL population in this study also provided challenges. While all of the children in the sample had enough proficiency with English to be tested in English, field notes indicated that some children were far more proficient than others, with some students reported to have almost no English language skills, while others reportedly spoke English very well. Those students who spoke very little English may have had difficulty with the vocabulary assessment, especially in the fall. Additionally, it would have been helpful to know the degree to which students were exposed to other languages. For some students, their time in the Head Start classroom may have been their only exposure to English, but for other students their primary language may have been English, and they may have only been incidentally exposed to a
second language (for example through interaction with a grandparent who spoke another language). The degree to which these students were exposed to English over the course of the school year may have been vastly different, and this study was unable to account for these types of differences with the data provided. Future studies should attempt to obtain more detailed information about students’ home language environments to better account for their influence on students’ vocabulary outcomes.

Finally, it would have been beneficial to follow these students over a longer period of time in order to examine patterns in vocabulary growth. Prior work (Dickinson & Smith, 1994) has indicated that teachers’ storybook reading style in kindergarten can impact children’s vocabulary growth in first grade, and it would have been wonderful to have this type of longitudinal data in this study in order to test for these types of effects.

*Implications for Professional Development*

Children from high-poverty backgrounds attend programs such as Head Start as a way to improve school readiness, and one of the elements of school readiness is oral language. Storybook reading has been shown to be one vehicle for improving children’s oral language (NELP, 2009), especially when accompanied by purposeful vocabulary instruction (Beck et al., 2001). This purposeful vocabulary instruction, embedded in the ubiquitous practice of storybook reading, may help these students catch up to their more economically advantaged peers in terms of oral language proficiency. Future studies should investigate the possibility of providing professional development programs for Head Start teachers that focus on storybook reading. One emphasis of this professional development should be to provide training on the purposeful selection of high utility
vocabulary words for instruction, the most effective strategies for instructing those 
words, and what constitutes a reasonable number of words for instruction during the daily 
classroom routine of storybook reading.

In addition to vocabulary instruction, teachers’ language complexity may also 
contribute to students’ vocabulary outcomes. The findings of this study with respect to 
teachers’ use of uncommon words suggest that teachers use relatively few uncommon 
words during storybook reading, and the majority of those words come from the text of 
the book. The findings further suggest that teachers’ use of uncommon words is often not 
accompanied by instructional scaffolding for those words. In this study, only students 
who began the year with the lowest levels of vocabulary had outcomes that were 
positively related to their exposure to teachers’ uncommon words. It is possible that the 
“uncommon” words used by the teacher were already familiar to the students who began 
the year with higher levels of vocabulary – these words were not uncommon *enough* to 
make a difference in vocabulary outcomes for these children. Additional professional 
development should focus on the complexity of teacher language, helping teachers to 
provide a rich oral language environment, with appropriate scaffolding, to help students 
gain the breadth and depth of vocabulary they will need for academic success.
References


Appendix A: Language Acquisition

Language Acquisition

Interactionist theorists contend that children learn language through interacting with more experienced role models, a process that falls in line with Vygotsky’s (1986) theories of learning. These role models are parents and caregivers in the first few years of life, and as the child grows older, the role models to whom they are exposed expand to include teachers, older or more experienced peers and community members, and others. To understand the interactionist approach to language acquisition, it is helpful to distinguish it from the other major theories of language acquisition which preceded it, since the interactionist theory includes some elements of these other theories.

According to Bohannon and Bonvillian (2005) there are three main theories of language acquisition. At one end of the spectrum are behavioral theories, at the other end are linguistic theories, and in the middle are theories which are interactionist, including Piaget’s cognitive approach, the information processing approach, and the social interaction approach. These three families of theories vary according to three dimensions: structuralism/functionalism, competence/performance, and nativism/empiricism.

Structuralism versus functionalism. Along this dimension, structural theories attempt to discover the mechanisms or processes that can explain observed language data, while functional theories attempt to explain the relationship between particular situations and/or environmental factors and the production of language. The rules of grammar that a linguist might use or the stimulus-response interaction which a behaviorist might explain are both examples of structural descriptions of language. On the other hand, a
functional description of language might include a description of the setting in which an utterance occurred in order to explain that utterance. For example, if a child asks her mother for a cookie, and this behavior is usually followed by the mother giving the child a cookie, the functionalists are not concerned with the form or structure of the utterance (it could be “want cookie!”, “may I please have a cookie?”, or “give me a cookie NOW!”), but rather with the pragmatic, social use of language.

*Competence versus performance.* This distinction is rather similar to the distinction between receptive and productive vocabulary, where receptive vocabulary is the ability to understand a word which is presented (sometimes using context), while productive vocabulary is the ability to produce that word and use it correctly in either written or spoken communication. Similarly, an individual’s competence refers to his or her understanding of the rules governing language use, his or her abstract linguistic knowledge. For example, an individual cannot possibly know every grammatically correct combination of words in the English language which could be assembled to make a sentence, yet that individual’s competence can be assessed by evaluating the grammaticality of his or her utterances, which reflect knowledge of the rules of grammar. Performance, on the other hand, refers to an individual’s actual production and use of language, not on the grammaticality of those utterances.

*Nativism versus empiricism.* This third dimension of the explanation of language acquisition refers to the extent to which emphasis is placed on the individual (nativism) or the environment (empiricism). Nativists claim that the complexities of language and the rate at which children acquire language indicate that there must be some innate language system present in all individuals. Empiricists liken language to other behaviors
and postulate that language is learned much in the same way as other behaviors. While most researchers do not adhere strictly to one end or the other of the nativism/empiricism continuum, some theories clearly reflect a greater influence of one dimension over the other. The behavioral, linguistic, and interactionist explanations of language acquisition all vary along these three continua.

**Behavioral theories**

In behavioral theories of language acquisition, performance is emphasized over competence, functionalism over structuralism, and empiricism over nativism. The behavioral theory states that operant conditioning can account for relationships of words and word meanings and children’s learning of these relationships. For example in the case of receptive language, if a mother says, “soft” prior to the child touching a puppy, the child will learn the meaning of the word *soft*. Productive language is a bit more complex. Behaviorists contend that operant conditioning accounts for development of productive language because parents reinforce speech that is more “adult-like” and ignore speech that is inappropriate or incorrect. Language is thus learned by a process of training through shaping and imitation. The concept of reinforcement plays a major role in language acquisition in this theory.

There is not a great deal of evidence to support behavioral theories outside the laboratory setting, where experiments with adults who are trained in artificial languages have provided evidence that their patterns of language acquisition (and associated errors) paralleled those of children learning their native language (Palermo & Eberhart, 1968). Criticisms of the behavioral approach have included the fact that results from laboratory settings are not often replicable in real life. For example, a number of studies (Brown &
Hanlon, 1970; Hirsh-Pasek, Treiman & Schneiderman, 1984; Penner, 1987) indicated that parents did not reinforce language consistently in the home setting. Additionally, Markov sentence models, in which behaviorists propose that the first word of the sentence acts as a stimulus for the second word, and the second word for the third word (and so on, as the result of parents training children in simple word combinations through shaping and imitation), cannot account for the production of sentences with complex embedded clauses (such as the one you are reading now) (McNeill, 1970). Because the acquisition of language is, at its heart, a learning process, the theories of behaviorists should clearly be part of the explanation of this complex phenomenon, but other theories provide explanations for parts of the process which cannot be accounted for by behavioral theories alone.

**Linguistic theories**

Unlike behavioral approaches, linguistic approaches to language acquisition assume that children are never told which sentences are correct or incorrect. In linguistic theories of language acquisition, competence is emphasized over performance, structuralism over functionalism, and nativism over empiricism. Linguistic approaches to language acquisition assume that a language has a set of grammars, a finite set of rules which are understood by speakers of that language, which allow those speakers to form an infinite number of comprehensible sentences. This accounts for the ability of children (and adults) to produce and understand sentences they have never heard before, unlike behavioral approaches. Noam Chomsky is the primary theorist behind the linguistic approaches to language acquisition, and although his theories have undergone multiple revisions and refinements over the past half-century, his basic contention is still the same.
That contention is that people are born with a part of the brain which Chomsky refers to as the language faculty. In its initial state, as proposed by Chomsky’s theory of universal grammar, the language faculty can be thought of as a network of connections and switches inside the brain. Based on the linguistic input experienced by a child, these connections and switches become “set,” and this is what leads a child to become a speaker of English, Hebrew, Chinese, or whatever language they are exposed to (Chomsky, 1997). Chomsky’s systems of language all have subsystems to account for semantic, syntactic, and phonological aspects of language, yet they are controversial in that they propose that a child must constantly form hypotheses about language based on the structure of the inputs they hear in order to learn language, even though some of those inputs may be ambiguous and/or indecipherable to someone just starting to learn language. This has been referred to as the learnability problem (Bohannon & Bonvillian, 2005).

How does Chomsky navigate this learnability problem? With the language acquisition device, or LAD. This “device” is thought to be a specialized language processor in the brain which provides the child with enough innate knowledge of the universal aspects of language to be able to speak (Chomsky, 1997; Pinker, 1994). It is speculated that young children categorize words by their patterns of usage in adult language and seem to be sensitive to nouns being things, while verbs represent actions or relations. This allows the child to learn new words’ meanings based on their order in a sentence (the syntax of the child’s native tongue).

Support for linguistic theories of language acquisition comes from a broad variety of studies. The work of Bock (1982) indicates that comprehending sentences involves
processing the hierarchical sentence structure into syntactic units, then deriving the meaning of those units. Studies of children’s over regularization errors, such as that done by Brown and Bellugi (1964), indicate that children’s inappropriate application of grammatical rules, with sentences such as “I goed to school”, provide evidence that children know the grammatical rule (for past tense, in this case) because they could not have heard this sentence construction in adult speech. Evidence for the universal nature of the LAD comes from the work of Slobin (1982), which indicates that very young children use subject-object word order regardless of the correct order in their native language.

Interestingly, studies involving deaf populations have provided both supporting and contrary evidence for linguistic theories of language acquisition. In a study of deaf children of hearing parents, where the parents chose not to teach the children to sign, the children developed rather sophisticated sign language of their own, which came to approximate many of the movements and signs present in established sign languages (Singleton, Morford, & Goldin-Meadow, 1993). Yet another study, this time involving hearing children of deaf parents (where the parents also chose not to teach the children to sign) found that the children, who were mainly exposed to spoken language through television, had little productive speech and no syntax upon entering preschool at age four. These children, with exposure to live (in-person) English speakers and some speech therapy, were able to improve to within the normal range rather quickly (Sachs, Bard, & Johnson, 1981). This seems to indicate that simple exposure to language through television was insufficient to produce a normal level of language learning, contrary to a more nativist perspective. These findings are in line with the work of Jerome Bruner
(1985), who argued that, while there may very well be a LAD, the LAD cannot function without the support of a Language Acquisition Support System, or LASS. Bruner’s LASS refers to the family and social network within which the child functions. Bruner, who may be considered one of the early interactionist theorists, argued that the social routines of the family, which include a language component, provide extensive support for the language development of the child, and he felt that this critical component was missing from Chomsky’s theory. Macnamara (1972) offered another alternative to the LAD. He postulated that children have a unique ability to read meaning into, and understand, social situations, and it is this ability that helps them to acquire language.

Linguistic approaches to language acquisition have further been called into question by Pinker’s (1984) “poverty of imagination” postulate, which states that just because you can’t show how a behavior is learned does not mean the behavior not learned (that it is innate). While some research had shown a negative correlation between parental use of recasts and children’s later use of the correct forms which had been the subject of the recasts (Morgan, Bonamo, & Travis, 1995), additional and subsequent studies (Farrar, 1992; Link & Bohannon, 2003) indicated that children who receive a recasted correction were more likely to repeat that correction, thus providing evidence for language learning, contrary to the linguistic perspective.

Interactionist approaches

In between the extreme approaches to language learning postulated by the behavioral and linguistic theories is a group of interactionist approaches to language acquisition. These include Piaget’s cognitive approach, the information processing approach, and a social interaction approach. According to Vygotsky (1962), cognitive
and social factors modify language, and language acquisition modifies cognitive and social skills – the causal relationships are reciprocal. It is these reciprocal relationships which the interactionist approaches strive to explain.

*Piaget’s cognitive approach.* The cognitive theory put forth by Jean Piaget (1926) is, in many ways, similar to the linguistic approach to language acquisition in that it emphasizes competence over performance, and structuralism over functionalism, but Piaget’s theory strikes more of a balance between nativism and empiricism. Where Piaget’s theory differs from the linguistic approach is that Piaget feels that language development is not separate from the more general cognitive development of the child. He argues that the complex structures of language are neither totally learned nor totally innate, but result from the child’s interaction with his/her environment at each successive level of cognitive functioning. Piaget’s interactive approach has come to be known as constructivism. To illustrate this link of language to cognitive functioning, Sinclair-de Zwart (1969) explained that a child who is in Piaget’s sensorimotor stage of development lacks the concept of object permanence, and therefore has no need for language or symbols to represent objects which are not present, because the child cannot conceive of objects which are not present in their immediate environment. As the child develops the concept of object permanence, he/she both needs and can make use of symbols or language, and this demonstrates how object permanence is a necessary precursor of language development. Additional work by Sinclair-de Zwart (1973) suggested that early sentences are better described by the cognitive-semantic categories of agent, action, patient, rather than the syntactic forms of subject, verb, object, and Pinker (1984) argues that this is perhaps the strongest evidence for this constructivist approach.
Criticism of the constructivist approach to language acquisition includes the contention that this cognitive-interactionist approach often explains correlations as if they were causal relationships. There is evidence from both deaf children (Bonvillian, Orlansky, Novack, & Folven, 1983) who learn to sign before obtaining full object permanence, and from precocious speakers (Ingram, 1981) whose spoken language develops more quickly than their normally-progressing cognitive skills, that some cognitive skills (such as object permanence) may be necessary for language development at a particular stage, while full completion of that Piagetian stage may not be necessary for the associated language skills to develop.

The information processing approach. The information processing approach uses a computer science paradigm to explain language learning. It is a structural, rather than a functional approach. Unlike the previous approaches, information processing research is often focused on adults, with children viewed as novice “users” who are becoming more skilled. The competition model (Bates & MacWhinney, 1989) is one of these types of models. Information processing models emphasize performance over competence, but indicate that performance may account for competence. Information processing models also clearly emphasize empiricism over nativism. These models propose that when a child wants to communicate something, for example the fact that they see several mice, the linguistic representation of “mouse” may be processed through the mental representations in the brain and the child may say “mouses.” When the child realizes that “mouses” is not similar to the adult speech exemplars, the mental representation which produced “mouses” is decreased in strength, while a representation that reflects the plural of mouse to be mice (perhaps through a parental recast) may be strengthened. Through
these incremental increases and decreases in the strength of various mental representations, children eventually produce more adult-like language.

Support for the information processing approach comes largely from computer-based research with parallel data processing (PDP) models. Because these models cannot yet simulate the conversational or social context in which language occurs, they have been criticized as being significantly underspecified (Bohannon & Bonvillian, 2005).

The social interaction approach. Social interactionists combine aspects of both behavioral and linguistic theories of language acquisition. They include the linguists’ propositions that language has a structure and rules, but also include the behavioral emphasis on the role of the environment in influencing the structure of language. The functions of language in the social environment are considered essential throughout the child’s development, so while social interactionists agree with linguists about the general acquisition of grammatical skills, they focus far more on the functional aspect of language. On the competence-performance continuum, social interactionists consider both aspects. While they may investigate children’s competence with language in their research studies, this competence cannot be measured outside the social interactions in which it occurs (the performance domain). Like the linguists and supporters of Piaget’s cognitive approach, social interactionists recognize that children cannot acquire language prior to certain stages of cognitive development. Yet like the behaviorists, social interactionists feel that the environment, and social interactions in particular, are critical to language development. One especially important social interaction which supports language development is that between a mother and child. The specialized speech which mothers direct toward their children is referred to as motherese or child-directed speech.
(CDS), and is thought to provide the basis for children’s eventual ability to segment the sound stream of speech into individual words, as well as learn conversational conventions such as turn-taking (Stern, Beebe, Jaffe, & Bennett, 1977).

The importance of gestures is another aspect of CDS which facilitates language acquisition. Mothers, or other caregivers, tend to draw children’s attention to words that are nouns by offering them the referent (such as holding up a teddy bear and saying “bear”) or by providing demonstrations of words which are verbs. Pairing words and their referents in this way assists children in learning the association between the object and the verbal code which refers to that object (Zukow-Goldring, 1996). As children grow older, the role of the language environment continues to be critical to development. Caregivers tend to use language which is appropriate to the developmental level of the child, adjusting this level as children grow older (Bohannon & Marquis, 1977). As part of the language teaching environment, a parent may use recasts to model more mature or correct forms of speech to a child. Nelson (1977) has argued that these parental recasts are particularly influential in the child’s language development in that they come at a time when the child’s attention is focused on the environment and the role their speech plays in obtaining what they want. For example, a child may say, “Wanna cookie,” and the mother replies, “Oh, you would like a cookie! Then you need to say, ‘May I please have a cookie?’” It is thought that the close proximity of the more mature form of the language to the original utterance by the child facilitates the child’s acquisition of the more mature form. Although imitation is not as important in social interactionist theory as it is in behavioral theories of language acquisition, social interactionists do believe that
imitation can play a role in children learning new speech patterns, specifically in helping children to test out hypotheses about various linguistic forms (Snow, 1978).

Unique support for the social interactionist theory of language acquisition comes from studies involving child directed speech. In research conducted on neglected children, Culp, Watkins, Lawrence, Letts, Kelly, & Rice (1991) found that while neglected children did not differ from their adequately-treated peers on cognitive development, they were significantly delayed on both measures of receptive and expressive language skills. Taken in combination with the many studies which report the positive effects of CDS (Bohannon & Bonvillian, 2005), the work of Culp et al. (1991) demonstrates the importance of social interaction in the development of language. Contrary evidence for the social interactionist theory is somewhat scant, due to the relatively recent advancement of this theory. Critics of the role of CDS point to the correlational nature of the studies conducted to date (Baker & Nelson, 1984). Additionally, much research on CDS has focused on the global construct of CDS as opposed to the specific features of CDS which might be critical for language development. Bohannon and Bonvillian (2005) suggest that the social interactionist approach to language acquisition is a promising theory in that it incorporates aspects of behavioral theories (by recognizing the importance of environmental inputs) and linguistic theories (considering children to be “language processors” who acquire the language code over time). It is this social interactionist approach which informs the research in the present study.
Appendix B. Head Start: A Historical Perspective

Head Start and Early Head Start are both federally funded programs which are administered by local agencies. The Head Start program was launched in 1965 as part of President Johnson’s War on Poverty, with the purpose of providing a comprehensive child development program designed to enhance children’s social competence by providing education, health, nutrition, and social services to children living in poverty, as well as providing avenues for parental involvement. President Johnson, who had been a teacher in a one-room schoolhouse in Texas, strongly believed in the power of education to break the cycle of poverty. Although Head Start was initially launched as a summer program, it quickly morphed into a nine-month half-day program, and eventually into a full-year program. In the program’s early years, the emphasis on “community involvement” included hiring many parents of Head Start children to work in the program, resulting in educational services being provided by people who were not professional educators and experiences for the participating children that were not necessarily enriching as compared to their home environments. Subsequently, the Child Development Associate (CDA) training program was launched in 1972 to raise the quality of Head Start teachers and their aides.

Under the Bush administration, and No Child Left Behind, Head Start programs were required to meet even higher standards in terms of the qualification of teachers and the performance of children related to school readiness skills (Snow, 2004; Pennsylvania Head Start Association, 2010). According to the US Department of Health and Human Services, under which the Head Start program falls, the teacher credentialing
requirements are set to increase in 2011, and again in 2013. By 2011, a CDA will no longer be a sufficient credential for Head Start teachers. Instead, they will be required to hold at least an associate’s degree in early childhood education. By 2013, at least 50% of the Head Start teachers nationwide will be required to hold a minimum of a baccalaureate degree in early childhood education (or in another subject, with the equivalent of baccalaureate level work in early childhood education) (Early Childhood Learning and Knowledge Center, 2010). It is hoped that these efforts to increase teacher credentials will have a positive impact on the language and literacy skills of the children enrolled in Head Start.

The current guidelines for Head Start indicate that the program’s purpose is to support children’s cognitive and language development, including emergent literacy skills, which are defined as being the precursors of reading, such as learning the letters of the alphabet (Early Childhood Learning and Knowledge Center, 2010). Head Start centers have broad flexibility in how they choose to meet these goals for developing the emergent literacy skills of their students. Given that the Head Start program is targeted to children between the ages of 3 and 5 living in poverty, and knowing that by age 3 these children are already far behind their more affluent peers in terms of language development, the Head Start program has been charged with a herculean task. If ever there were a need to implement proven, research-based practices for improving children’s language development, it would be in Head Start classrooms.
Appendix C. Figures and Tables

Figure 1. The influence of teacher language and interaction on vocabulary development.
Table 1. Elements for Analysis.

<table>
<thead>
<tr>
<th>Element</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary Strategies</td>
<td>Definitional Strategies Used by Teacher:</td>
</tr>
<tr>
<td></td>
<td>• Defines word (child-friendly definition)</td>
</tr>
<tr>
<td></td>
<td>• Acts out meaning of word</td>
</tr>
<tr>
<td></td>
<td>• Gives synonym</td>
</tr>
<tr>
<td></td>
<td>• Points to picture of word</td>
</tr>
<tr>
<td></td>
<td>• Use of recasts</td>
</tr>
<tr>
<td></td>
<td>• Asks students to define word (“What does basked mean?”)</td>
</tr>
<tr>
<td>Contextualization Strategies Used by Teacher:</td>
<td>Connect word to its use in the book– give meaning and/or example</td>
</tr>
<tr>
<td></td>
<td>• Connect word to children’s personal experience (“Do you reside in a house or an apartment? If you crept around the classroom, what would that look like? What is something you detest?”)</td>
</tr>
<tr>
<td></td>
<td>• Connect word to illustration (more than pointing, actually explaining elements of illustration)</td>
</tr>
<tr>
<td>Orthographic/Phonological Strategies Used by Teacher:</td>
<td>Discuss spoken form of the word (“What sound does the word ___ begin with?”)</td>
</tr>
<tr>
<td></td>
<td>• Discuss written form(s) of the word or show written form of the word (“What letter does the word ___ begin with?”)</td>
</tr>
<tr>
<td></td>
<td>• Ask students (s) to say the target word aloud</td>
</tr>
<tr>
<td>Code Switching</td>
<td>• Number of non-English words used by the teacher (code switching) during the storybook reading session.</td>
</tr>
<tr>
<td>Word Utility</td>
<td>• Frequent Words, High Utility Words, or Low Utility Words instructed by teachers</td>
</tr>
<tr>
<td>Language Complexity</td>
<td>Quantity:</td>
</tr>
<tr>
<td></td>
<td>• Teacher’s Mean Length of Utterance</td>
</tr>
<tr>
<td></td>
<td>Quality:</td>
</tr>
<tr>
<td></td>
<td>• Teacher’s Type/Token ratio</td>
</tr>
<tr>
<td></td>
<td>• Teacher’s Ratio of uncommon to more common words (ZenoRatio)</td>
</tr>
<tr>
<td>Group Size</td>
<td>• Number of children participating in the storybook reading session</td>
</tr>
</tbody>
</table>
Table 2. Classification of Words Instructed.

<table>
<thead>
<tr>
<th>Frequent words</th>
<th>High Utility words</th>
<th>Low Utility words</th>
</tr>
</thead>
<tbody>
<tr>
<td>balanced*</td>
<td>amusing</td>
<td>lioness</td>
</tr>
<tr>
<td>ball*</td>
<td>bask</td>
<td>Mohawk</td>
</tr>
<tr>
<td>compare*</td>
<td>beak</td>
<td>ostrich</td>
</tr>
<tr>
<td>fun*</td>
<td>bore</td>
<td>seal</td>
</tr>
<tr>
<td>jump*</td>
<td>cage</td>
<td>Edward_the-Emu</td>
</tr>
<tr>
<td>laughed*</td>
<td>casually</td>
<td>Edward</td>
</tr>
<tr>
<td>sick*</td>
<td>clamber</td>
<td>Edwina</td>
</tr>
<tr>
<td>slipped*</td>
<td>considered</td>
<td></td>
</tr>
<tr>
<td>smile*</td>
<td>curled</td>
<td></td>
</tr>
<tr>
<td>snake*</td>
<td>crept</td>
<td></td>
</tr>
<tr>
<td>surprised*</td>
<td>detest</td>
<td></td>
</tr>
<tr>
<td>swimming*</td>
<td>dived</td>
<td></td>
</tr>
<tr>
<td>teeth</td>
<td>emu</td>
<td></td>
</tr>
<tr>
<td>wings*</td>
<td>face_to_face</td>
<td></td>
</tr>
<tr>
<td>zoo</td>
<td>gentleman</td>
<td></td>
</tr>
<tr>
<td></td>
<td>growl</td>
<td></td>
</tr>
<tr>
<td></td>
<td>gumption</td>
<td></td>
</tr>
<tr>
<td></td>
<td>hissing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>identical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>impressive</td>
<td></td>
</tr>
<tr>
<td></td>
<td>overheard</td>
<td></td>
</tr>
<tr>
<td></td>
<td>peck</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pen (synonym for cage)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>reside</td>
<td></td>
</tr>
<tr>
<td></td>
<td>roar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>screaming</td>
<td></td>
</tr>
<tr>
<td></td>
<td>shy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>snarl</td>
<td></td>
</tr>
<tr>
<td></td>
<td>visitor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>whiskers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>zookeeper</td>
<td></td>
</tr>
</tbody>
</table>

* - included in the 3000 most common words in the Educator’s Word Frequency Guide (Zeno, Ivens, Millard, & Duvvuri, 1995)
Table 3. Vocabulary instruction strategies.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Definition Strategies</td>
<td>0</td>
<td>41</td>
<td>9.31</td>
<td>9.24</td>
</tr>
<tr>
<td>define word</td>
<td>0</td>
<td>14</td>
<td>1.98</td>
<td>3.17</td>
</tr>
<tr>
<td>act-out word</td>
<td>0</td>
<td>8</td>
<td>2.33</td>
<td>2.68</td>
</tr>
<tr>
<td>give synonym</td>
<td>0</td>
<td>5</td>
<td>.70</td>
<td>1.17</td>
</tr>
<tr>
<td>point to picture of word</td>
<td>0</td>
<td>9</td>
<td>1.50</td>
<td>1.91</td>
</tr>
<tr>
<td>use recast</td>
<td>0</td>
<td>5</td>
<td>.91</td>
<td>1.45</td>
</tr>
<tr>
<td>ask students for word meaning</td>
<td>0</td>
<td>11</td>
<td>1.80</td>
<td>2.76</td>
</tr>
<tr>
<td>Total Contextualization Strategies</td>
<td>0</td>
<td>14</td>
<td>2.08</td>
<td>3.20</td>
</tr>
<tr>
<td>connect word to use in book</td>
<td>0</td>
<td>4</td>
<td>.29</td>
<td>0.87</td>
</tr>
<tr>
<td>connect word to personal experience</td>
<td>0</td>
<td>7</td>
<td>1.26</td>
<td>2.03</td>
</tr>
<tr>
<td>connect word to illustration</td>
<td>0</td>
<td>3</td>
<td>.53</td>
<td>0.85</td>
</tr>
<tr>
<td>Total Orthographic/Phonological Strategies</td>
<td>0</td>
<td>7</td>
<td>.71</td>
<td>1.78</td>
</tr>
<tr>
<td>discuss spoken form of word</td>
<td>0</td>
<td>5</td>
<td>.38</td>
<td>1.22</td>
</tr>
<tr>
<td>discuss written form of word</td>
<td>0</td>
<td>1</td>
<td>.08</td>
<td>0.27</td>
</tr>
<tr>
<td>ask students to say word aloud</td>
<td>0</td>
<td>5</td>
<td>.25</td>
<td>0.93</td>
</tr>
<tr>
<td>Total Vocabulary Strategies</td>
<td>0</td>
<td>60</td>
<td>12.10</td>
<td>12.30</td>
</tr>
</tbody>
</table>
Table 4. Teachers’ Language Complexity

<table>
<thead>
<tr>
<th>Language Complexity</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLU</td>
<td>4.34</td>
<td>7.28</td>
<td>5.96</td>
<td>.72</td>
</tr>
<tr>
<td>Type/Token ratio</td>
<td>.15</td>
<td>.33</td>
<td>.24</td>
<td>.04</td>
</tr>
<tr>
<td>Zeno ratio</td>
<td>.03</td>
<td>.11</td>
<td>.09</td>
<td>.02</td>
</tr>
<tr>
<td>Frequent Words</td>
<td># of teachers instructing</td>
<td>High Utility Words</td>
<td># of teachers instructing</td>
<td>Low Utility words</td>
</tr>
<tr>
<td>----------------</td>
<td>--------------------------</td>
<td>--------------------</td>
<td>---------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>jump</td>
<td>3</td>
<td>bore</td>
<td>11</td>
<td>seal</td>
</tr>
<tr>
<td>ball, laughed, swimming, zoo</td>
<td>2</td>
<td>emu</td>
<td>10</td>
<td>ostrich</td>
</tr>
<tr>
<td>balanced, compare, fun, sick, slipped, smile, snake, surprised, teeth, wings</td>
<td>1</td>
<td>roar</td>
<td>9</td>
<td>Edward, Edward_the_Emu, Edwina, lioness, Mohawk</td>
</tr>
<tr>
<td>hissing</td>
<td>7</td>
<td>beak, reside</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>detest</td>
<td>4</td>
<td>growl, visitor</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>crept, clamber, bask, zookeeper</td>
<td>2</td>
<td>amusing, cage, casually, considered, curled, dived, face-to-face, gentleman, gumption, identical, impressive, heard, peck, pen(cage), screaming, shy, snarl, whiskers</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
### Table 6. Word Utility

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Words Instructed (Types)</td>
<td>0</td>
<td>18</td>
<td>5.22</td>
<td>4.67</td>
</tr>
<tr>
<td>Frequent Words</td>
<td>0</td>
<td>3</td>
<td>.97</td>
<td>1.03</td>
</tr>
<tr>
<td>High Utility Words</td>
<td>0</td>
<td>16</td>
<td>3.85</td>
<td>3.84</td>
</tr>
<tr>
<td>Low Utility Words</td>
<td>0</td>
<td>3</td>
<td>.40</td>
<td>.75</td>
</tr>
</tbody>
</table>

### Table 7. Word Utility and Vocabulary Strategy Ratios.

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>High utility words per total words instructed (HU ratio)</td>
<td>0</td>
<td>1</td>
<td>.60</td>
<td>.36</td>
</tr>
<tr>
<td>Vocabulary strategies per word (VocStrat ratio)</td>
<td>0</td>
<td>8</td>
<td>2.13</td>
<td>2.02</td>
</tr>
</tbody>
</table>

### Table 8. Parts of speech of instructed words.

<table>
<thead>
<tr>
<th>Part of Speech</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>0</td>
<td>8</td>
<td>1.84</td>
<td>1.92</td>
</tr>
<tr>
<td>Verb</td>
<td>0</td>
<td>9</td>
<td>2.42</td>
<td>2.44</td>
</tr>
<tr>
<td>Adjective</td>
<td>0</td>
<td>4</td>
<td>.87</td>
<td>.95</td>
</tr>
<tr>
<td>Adverb</td>
<td>0</td>
<td>1</td>
<td>.05</td>
<td>.21</td>
</tr>
</tbody>
</table>
Table 9. Student-level, teacher-level, and interaction terms as predictors of spring PPVT outcome.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Unstandardized coefficients</th>
<th>Standardized coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE</td>
</tr>
<tr>
<td>Constant</td>
<td>63.174</td>
<td>26.274</td>
</tr>
<tr>
<td>fall PPVT</td>
<td>.519</td>
<td>.050</td>
</tr>
<tr>
<td>DLL (Yes=1, No=0)</td>
<td>-3.121</td>
<td>1.086</td>
</tr>
<tr>
<td>Gender (M=1, F=0)</td>
<td>.052</td>
<td>.999</td>
</tr>
<tr>
<td>Age (in months)</td>
<td>.340</td>
<td>.087</td>
</tr>
<tr>
<td>code switching</td>
<td>-.424</td>
<td>.181</td>
</tr>
<tr>
<td>MLU</td>
<td>-2.004</td>
<td>1.124</td>
</tr>
<tr>
<td>type/token ratio</td>
<td>-127.515</td>
<td>102.753</td>
</tr>
<tr>
<td>Zeno ratio</td>
<td>20.852</td>
<td>40.381</td>
</tr>
<tr>
<td>reading group size</td>
<td>-.120</td>
<td>.171</td>
</tr>
<tr>
<td>VocStrat ratio</td>
<td>-.423</td>
<td>.384</td>
</tr>
<tr>
<td>HU ratio</td>
<td>.769</td>
<td>2.542</td>
</tr>
<tr>
<td>MLU x fall PPVT</td>
<td>.084</td>
<td>.077</td>
</tr>
<tr>
<td>Zeno Ratio x fall PPVT</td>
<td>-10.995</td>
<td>3.329</td>
</tr>
<tr>
<td>Type/Token x fall PPVT</td>
<td>2.355</td>
<td>1.657</td>
</tr>
<tr>
<td>VocStrat ratio x fall PPVT</td>
<td>.130</td>
<td>.036</td>
</tr>
<tr>
<td>HU ratio x fall PPVT</td>
<td>-.436</td>
<td>.233</td>
</tr>
</tbody>
</table>

Table 10. Student- and Teacher-level predictors of vocabulary outcomes as measured by PPVT.

<table>
<thead>
<tr>
<th>Model</th>
<th>r</th>
<th>r²</th>
<th>Adjusted r²</th>
<th>Standard error</th>
<th>Change Statistics</th>
<th>Significant F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>of the estimate</td>
<td>r² change</td>
<td>F change</td>
<td>df1</td>
</tr>
<tr>
<td>1</td>
<td>.700a</td>
<td>.490</td>
<td>.480</td>
<td>7.190</td>
<td>.490</td>
<td>45.932</td>
</tr>
<tr>
<td>2</td>
<td>.730b</td>
<td>.533</td>
<td>.505</td>
<td>7.010</td>
<td>.043</td>
<td>2.422</td>
</tr>
<tr>
<td>3</td>
<td>.761c</td>
<td>.579</td>
<td>.541</td>
<td>6.752</td>
<td>.045</td>
<td>3.865</td>
</tr>
</tbody>
</table>

a. Predictors for Step 1: constant, age, gender, DLL, fall PPVT. b. Predictors for Step 2: constant, age, gender, DLL, fall PPVT, reading group size, MLU, code switching, type/token ratio, Zeno ratio, VocStrat ratio, HU ratio. c. Predictors: constant, age, gender, DLL, fall PPVT, reading group size, MLU, code switching, type/token ratio, Zeno ratio, VocStrat ratio, HU ratio, MLU x fall PPVT, Zeno Ratio x fall PPVT, Type/Token x fall PPVT, VocStrat ratio x fall PPVT, HU ratio x fall PPVT.
Table 11. Student-level, teacher-level, and interaction terms as predictors of spring LE outcome.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE</td>
</tr>
<tr>
<td>(Constant)</td>
<td>64.624</td>
<td>31.425</td>
</tr>
<tr>
<td>fall LE Vocabulary Score</td>
<td>.513</td>
<td>.051</td>
</tr>
<tr>
<td>DLL (Yes=1, No=0)</td>
<td>-6.478</td>
<td>4.957</td>
</tr>
<tr>
<td>Gender (M=1, F=0)</td>
<td>-.390</td>
<td>4.613</td>
</tr>
<tr>
<td>Age (in months)</td>
<td>1.173</td>
<td>.402</td>
</tr>
<tr>
<td>code switching</td>
<td>.914</td>
<td>.861</td>
</tr>
<tr>
<td>MLU</td>
<td>4.085</td>
<td>5.132</td>
</tr>
<tr>
<td>type/token ratio</td>
<td>153.986</td>
<td>84.844</td>
</tr>
<tr>
<td>Zeno ratio</td>
<td>-567.818</td>
<td>183.930</td>
</tr>
<tr>
<td>reading group size</td>
<td>-1.373</td>
<td>.771</td>
</tr>
<tr>
<td>VocStrat ratio</td>
<td>-.029</td>
<td>1.780</td>
</tr>
<tr>
<td>HU ratio</td>
<td>1.149</td>
<td>11.451</td>
</tr>
</tbody>
</table>

Table 12. Student- and Teacher-level predictors of vocabulary outcomes as measured by LE.

<table>
<thead>
<tr>
<th>Model</th>
<th>$r$</th>
<th>$r^2$</th>
<th>Adjusted $r^2$</th>
<th>$r^2$ change</th>
<th>Standard error</th>
<th>$F$ change</th>
<th>df1</th>
<th>df2</th>
<th>Significant F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.677$^a$</td>
<td>.458</td>
<td>.447</td>
<td>.458</td>
<td>32.12883</td>
<td>39.749</td>
<td>4</td>
<td>188</td>
<td>.000</td>
</tr>
<tr>
<td>2</td>
<td>.706$^b$</td>
<td>.499</td>
<td>.468</td>
<td>.040</td>
<td>31.49990</td>
<td>2.083</td>
<td>7</td>
<td>181</td>
<td>.047</td>
</tr>
<tr>
<td>3</td>
<td>.712$^c$</td>
<td>.507</td>
<td>.462</td>
<td>.008</td>
<td>31.67876</td>
<td>.592</td>
<td>5</td>
<td>176</td>
<td>.706</td>
</tr>
</tbody>
</table>

a. Predictors for Step 1: constant, age, gender, DLL, fall LE.  
b. Predictors for Step 2: constant, age, gender, DLL, fall LE, reading group size, MLU, code switching, type/token ratio, Zeno ratio, VocStrat ratio, HU ratio.  
c. Predictors: constant, age, gender, DLL, fall LE, reading group size, MLU, code switching, type/token ratio, Zeno ratio, VocStrat ratio, HU ratio, MLU x fall LE, Zeno Ratio x fall LE, Type/Token x fall LE, VocStrat ratio x fall LE, HU ratio x fall LE.
Figure 2. Code Switching x DLL Interaction
Figure 3. Zeno Ratio x fall PPVT Interaction
Figure 4. VocStrat Ratio x fall PPVT Interaction
Figure 5. Zeno Ratio x fall LE Interaction

2=1 SD above mean, 
1=within 1 SD of mean, 
0=below one SD
Figure 6. Number of words instructed by word utility.
Appendix D. Computation of intra-class correlation.

The first step in Hierarchical Linear Modeling (HLM) is to compute the intra-class correlation to determine the proportion of variance in the outcome (Spring vocabulary score) attributable to between group differences. The equations used for the student level (L1) and the teacher level (L2) are shown below.

$L1\ Model:\ \text{VOC}_{ij} = \beta_{0j} + \eta_{ij}$

$L2\ Model:\ \beta_{0j} = \gamma_{00} + u_{0j}$

$\text{VOC}_{ij}$ is the score on either the PPVT or LE at time 2 for student $i$ in classroom $j$, $\beta_{0j}$ is the average VOC score at time 2 in classroom $j$, and $\gamma_{00}$ is the grand mean of VOC scores across all classrooms. The terms $\eta$ and $u$ designate random error terms and reflect the extent to which an individual student differs from his/her classroom mean ($\eta$) and the extent to which an individual teacher’s classroom differs from the grand mean ($u$).

The Intra Class Correlation (ICC) is computed according to the following formula to determine the proportion of the variance in the outcome (Spring Vocabulary score) attributable to between group differences:

$$ICC = \frac{\tau_{00}}{\tau_{00} + \sigma^2} = \frac{Var\ (u_{0j})}{Var\ (u_{0j}) + Var\ (\eta_{ij})}$$

The values used in the ICC calculation for this study are found in Table C below. Because the focus of this study was to look at the change in student vocabulary from time 1 (fall) to time 2 (spring), the model was run a second time using the residual values for the vocabulary scores of both outcome measures (the PPVT or LE) to see if the use of
change scores showed an appreciable difference in the proportion of variance at each level, and in the resulting ratio. The residuals were calculated by subtracting each student’s fall vocabulary score from their spring score, and the Intra Class Correlation was calculated for these values as well. The results indicate that with change scores as the outcome measure, there was more variance at the student level for the PPVT, but far less variance at the student level for the LE. At the teacher level, there was less variance when using change scores than there was when using spring vocabulary outcome scores across both measures. The ICCs for the change score outcome measure indicated that very little difference was attributable to between group differences, as was the result with the ICCs for the spring vocabulary outcome measures, although the ICC increased for the PPVT and decreased for the LE when the change scores were used. These results indicated that the use of hierarchical linear modeling was not supported by this data set, so hierarchical linear regression was used instead.

Table C. Values for computation of Intra Class Correlation.

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>$\tau_{00}$</th>
<th>$\sigma^2$</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring PPVT</td>
<td>1.56</td>
<td>101.22</td>
<td>0.015</td>
</tr>
<tr>
<td>Spring LE</td>
<td>34.20</td>
<td>1822.16</td>
<td>0.018</td>
</tr>
<tr>
<td>PPVT residual</td>
<td>3.36</td>
<td>75.13</td>
<td>0.04</td>
</tr>
<tr>
<td>LE residual</td>
<td>.61</td>
<td>1500.04</td>
<td>0.0004</td>
</tr>
</tbody>
</table>
## Appendix E. Pearson Correlation Matrix

<table>
<thead>
<tr>
<th>Variable</th>
<th>PPVTwFall</th>
<th>PPVTwSpring</th>
<th>Learning Express Vocabulary Score Fall</th>
<th>Learning Express Vocabulary Score Spring</th>
<th>Age</th>
<th>Gender M=1, F=0</th>
<th>DLL (1=DLL, 0=monolingual English)</th>
<th>code switching</th>
<th>MLU</th>
<th>type/token ratio</th>
<th>Zeno ratio</th>
<th>VocStrat ratio</th>
<th>reading group size</th>
<th>HU ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPVTwFall</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPVTwSpring</td>
<td>.661**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LE Vocabulary Score Fall</td>
<td>.684**</td>
<td>.694**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LE Vocabulary Score Spring</td>
<td>.676**</td>
<td>.713**</td>
<td>.656**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.402**</td>
<td>.436**</td>
<td>.357**</td>
<td>.368**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender M=1, F=0</td>
<td>.049</td>
<td>.088</td>
<td>.109</td>
<td>.072</td>
<td>.139'</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DLL (1=DLL, 0=English only)</td>
<td>.141'</td>
<td>-.017</td>
<td>-.025</td>
<td>-.020</td>
<td>.106</td>
<td>.020</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>code switching</td>
<td>-.050</td>
<td>-.133</td>
<td>-.170'</td>
<td>-.066</td>
<td>.097</td>
<td>.017</td>
<td>.177'</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLU</td>
<td>.030</td>
<td>.015</td>
<td>.147'</td>
<td>.022</td>
<td>.015</td>
<td>-.005</td>
<td>-.108</td>
<td>-.568**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>type/token ratio</td>
<td>-.018</td>
<td>-.059</td>
<td>-.010</td>
<td>.010</td>
<td>.171'</td>
<td>-.030</td>
<td>.194**</td>
<td>.155'</td>
<td>.243''</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zeno ratio</td>
<td>-.137'</td>
<td>-.075</td>
<td>-.011</td>
<td>-.144'</td>
<td>.047</td>
<td>-.006</td>
<td>-.042</td>
<td>-.056</td>
<td>.524**</td>
<td>.633**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VocStrat ratio</td>
<td>-.066</td>
<td>-.114</td>
<td>-.055</td>
<td>-.081</td>
<td>-.111</td>
<td>.010</td>
<td>-.208**</td>
<td>-.142'</td>
<td>.191**</td>
<td>-.262**</td>
<td>.088 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>reading group size</td>
<td>.041</td>
<td>-.095</td>
<td>-.094</td>
<td>-.134</td>
<td>.082</td>
<td>-.067</td>
<td>.000</td>
<td>.431''</td>
<td>-.127</td>
<td>-.040</td>
<td>-.039</td>
<td>-.051 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HU ratio</td>
<td>-.081</td>
<td>-.097</td>
<td>-.072</td>
<td>-.101</td>
<td>-.087</td>
<td>-.005</td>
<td>-.122</td>
<td>-.236''</td>
<td>.104</td>
<td>-.469''</td>
<td>-.148'</td>
<td>.713''</td>
<td>.101 1</td>
<td></td>
</tr>
</tbody>
</table>

**p<.01  *p<.05