EFFECTS OF A SPECIALIZED EARLY INTERVENTION FOR CHILDREN

WITH SEVERE LANGUAGE IMPAIRMENT

APPROVED BY SUPERVISORY COMMITTEE

Carroll W. Hughes, Ph.D.

Cheryl Silver, Ph.D.

Beth D. Kennard, Psy.D.

Peter L. Stavinoha, Ph.D.

Maryann O. Hetrick, Ph.D.

Joyce S. Pickering, Hum.D.

DEDICATION

This dissertation is dedicated to

Mom, Dad, Alex, Andrew, & Juliet

You are my favorite people

TSFY

-and-

Papo & Cateche

I will always carry you with me

EFFECTS OF A SPECIALIZED EARLY INTERVENTION FOR CHILDREN WITH SEVERE LANGUAGE IMPAIRMENT

by

VANESSA RENEE SALAZAR

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August, 2008

ACKNOWLEDGEMENTS

The process of creating this dissertation has taught me a wealth of information about language problems in children and just how profound an impact these difficulties have on their learning. I have also seen firsthand how a supportive environment that fosters their language skill development allows them to meet their full potential. I would like to thank Dr. Joyce Pickering and the Shelton School staff for allowing me to work with their wonderful students.

To my dissertation committee, thank you for your feedback and expertise in the areas of childhood learning disorders and research design. I would especially like to thank Dr. Maryann Hetrick for your constant patience, reassurance, and constructive feedback. And to Dr. Hughes, I certainly could not have crossed the finish line without you in my corner, and I am eternally grateful for your guidance and mentorship, not only with the logistics of this project, but also in my professional development.

To my clinical supervisors this year, thank you for your support and for lending an ear when I needed it. Dr. Monty Evans, your clinical expertise and patience with complex cases have given me a new understanding of the therapeutic process. Gayle Marshall, thank you for helping me get through the stress of the dissertation while teaching me to be a better clinician. Your empathic demeanor, excellent supervision, and model of what it means to serve the best interests of the patient have been a true inspiration.

To my classmates Julie, Renee, Sally, Charlotte, Michelle, Kim, and Todd, you have no idea how important your friendships have been to me throughout graduate school and the dissertation process in particular. You have been like a second family during my time in Dallas, and I am so grateful to have known you. To my dear friends from earlier days, Catherine, Jessica, Claudia, and Linda, thank you for being a constant source of support and encouraging me to stay strong. I love and appreciate you.

Saving the best for last, to my family, "you are the *best*!" I marvel at the things we have been through together, the unsurpassed humor and laughter, the losses we have suffered, and how we only seem to become stronger and closer. Thank you for always listening, caring, loving, helping, and laughing. You have supported me in all my efforts, been proud of me whether I succeeded or failed, and taught me to always find the humor in life and remember what is truly important. Newmistaca and I love you.

Colorín, colorado, este cuento se ha acabado.

EFFECTS OF A SPECIALIZED EARLY INTERVENTION FOR CHILDREN WITH SEVERE LANGUAGE IMPAIRMENT

VANESSA RENEE SALAZAR, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2008

CARROLL W. HUGHES, Ph.D.

Children with language impairment experience difficulties in grammar, vocabulary, and phonological skills, and they are susceptible to developing learning disorders without intervention (Scarborough, 1990; Tallal, Ross, & Curtiss, 1989; Van der Lely & Stollwerk, 1996). Intervention is imperative to prevent further delays in language and potential emotional and social problems stemming from poor communication skills (Bruce & Hansson, 2008). Speech-language therapy is effective for these children (Law, Garrett, & Nye, 2003), and various interventions have been investigated with mixed results. Certain factors have been found to be associated with language outcome, including expressive language difficulties (Law, Garrett, & Nye, 2004), nonverbal cognitive ability (Bishop & Edmundson, 1987; Oliver, Dale, & Plomin, 2004), age (Schery, 1985), and initial type of impairment (Boyle, McCartney, Forbes, & O'Hare, 2007; Law et al., 2004). No empirical investigations have been published on the effects of the Montessori Method Applied to Children At-Risk for learning disabilities

(Pickering, 1988) or the DuBard Association Method (DuBard & Martin, 2000), two central components of a specialized language intervention program at the Shelton School in Dallas, Texas. This program evaluation examines change in the language skills of 20 children ages 3 to 9 with language impairment during participation in this three-year intervention, the Shelton Early Intervention Language Learning Program. Variables associated with language outcomes are also examined. Using one-way repeated measures analyses of variance, significant improvement was found on measures of expressive language, expressive vocabulary, and articulation; significant decline was found on measures of receptive language and receptive vocabulary. No interaction effects were found between baseline nonverbal intelligence or age and language outcomes. Reliable change indices showed that a minimal proportion of participants improved, with the exception of the articulation measure, on which the majority of participants improved. A two-way contingency table analysis revealed that a relationship existed between baseline language impairment type and receptive language outcome, in which children who did not respond to intervention had a higher likelihood of having more pervasive language impairment at baseline than children who declined. Further research on the apparent differential response to expressive and articulation measures versus receptive measures is warranted.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER ONE
Introduction

Severe developmental language impairment negatively affects children's ability to communicate effectively and achieve academic success. Children with language deficits who do not receive remediation are predisposed to developing other learning disorders (Scarborough, 1990; Tallal et al., 1989; Van der Lely & Stollwerk, 1996), and they have a higher incidence of concurrent attention-deficit/hyperactivity disorder (e.g., Cohen et al., 2000) and motor impairment (e.g., Trauner, Wulfeck, Tallal, & Hesselink, 2000). Previous research has revealed associations between the efficacy of early interventions for children with learning disabilities and specific factors that predict improvement. Traditional speech and language therapy has been shown to be effective for children with primary oral language disorder (Law et al., 2003), and alternative educational programs and treatments have been investigated with mixed results. A limited number of studies have identified several predictors of language outcome in children with language impairment, including narrative retelling ability (Bishop & Edmundson, 1987; Botting, Faragher, Simkin, Knox, & Conti-Ramsden, 2001), expressive syntax (Botting, Faragher, Simkin, Knox, & Conti-Ramsden, 2001), phonological and vocabulary difficulties (Law et al., 2004), pervasiveness of impairment across expressive and receptive domains (Boyle et al., 2007; Law et al., 2004), and nonverbal cognitive ability (Bishop & Edmundson, 1987; Oliver et al., 2004). Furthermore, several studies have suggested that intervention be implemented at an early age to prevent further lag in language development and potential social and behavioral problems (Gillon, 2000, 2002; Hautus, Setchell, Waldie, & Kirk, 2003; Rvachew, Ohberg, Grawburg, & Heyding, 2003). However, no studies to date have empirically examined the impact of the current specialized intervention. The fact that much of the severely limited research on the intervention components used in this study is methodologically flawed or inconclusive underscores the need for more research with appropriate procedures and statistical analyses. In addition, theoretical underpinnings and years of knowledge gathered by education professionals and speech-language pathologists support the intervention components chosen to remediate language impairment. This study is important to expand our knowledge base on the effects of the current program components and contribute to the literature regarding evidence-based practice.

The current study examined the effects of a specialized, intensive early intervention at the Shelton School in Dallas, Texas, on the language skills of children with severe language impairment, as well as the factors that are associated with improvement in these children. A group of children diagnosed with primary oral language or phonological disorder were enrolled in the Early Intervention Language Learning (EI) program, which combined the Montessori teaching method ("Montessori," 2005; Montessori, 1988) adapted for children at-risk for a learning disability (Pickering, 1988, 1992, 2004b), the DuBard Association Method (DuBard & Martin, 2000; McGinnis, 1939), and occupational therapy with incorporated sensory-integration activities. This longitudinal study aims to identify the skill areas that improve over three years of intervention, with a focus on language, articulation, and oromotor skills. A selection of baseline and demographic factors that past research has shown are associated with improvement, namely expressive vs. mixed receptive-expressive language impairment, nonverbal IQ, and age, will also be examined in order to identify which variables are involved in language outcome.

CHAPTER TWO
Review of the Literature

LANGUAGE IMPAIRMENT IN CHILDREN

Language impairment (LI) involves a delay in the progression of language skills that is not explained by neurological, cognitive, or hearing impairment (Botting, 1998; Cantwell & Baker, 1987). Children with LI are at risk for dyslexia (Catts, 1993; Catts, Fey, Zhang, & Tomblin, 1999, 2001; Larrivee & Catts, 1999; Snowling, Bishop, & Stothard, 2000; Snowling & Hayiou-Thomas, 2006), and their early language delays may increase the gap between their level of learning and that of their peers as they continue through school (Stanovich, 1986). Difficulties with academics may lead to emotional and behavioral problems, and communication barriers may impact their ability to connect socially with others (Bashir & Scavuzzo, 1992; Cohen et al., 2000; Rice, Sell, & Hadley, 1991; Rutter & Mawhood, 1991). Children with receptive language impairment have difficulty understanding spoken language, while children with expressive language impairment have difficulty expressing themselves verbally (American Psychiatric Association, 2000; Leonard, 1990). Children with LI may be more susceptible to other learning disorders such as dyslexia (Bishop & Snowling, 2004; Carroll & Snowling, 2004), and psychiatric disorders, including attention-deficit/hyperactivity disorder (ADHD; Baker & Cantwell, 1992; Beitchman, Hood, & Inglis, 1990) and social phobia (Voci, Beitchman, Brownlie, & Wilson, 2006). Children with LI may experience motor impairment, sensory-integration deficits, and perceptual difficulties (American Psychiatric Association, 2000). Early language intervention tailored to their needs is regarded as imperative for these children in order to help remediate their deficits and offer them a supportive learning environment. While some children with LI make substantial improvement with intervention, others continue to struggle considerably. The population of children with language disorders is highly heterogeneous and includes various combinations of expressive, mixed receptive-expressive, articulation, speech apraxia, auditory processing, visual perception, phonological awareness, and reading problems, along with varying levels of impairment severity in these domains. This heterogeneity makes difficult work of conducting research using consistent inclusion and exclusion criteria and measures of improvement with intervention. Due to the variable methodological quality of studies in this area, the field is plagued by lack of universal guidelines for which interventions are most effective for which children. Interventions may focus on remediating different components of language, i.e., morphology, syntax, semantics, or phonological awareness; different service delivery models, i.e., speech-language pathologist vs. trained parents; and various lengths of intervention, making meta-analyses difficult to conduct. A limited number of rigorously controlled studies exists in this area, but generally research has shown that those children with mixed receptive-expressive LI do not respond as well to language intervention as those with expressive difficulties alone (Boyle et al., 2007; Law et al., 2004), an intervention period of eight weeks or longer is preferable to short-term interventions (Law et al., 2004), and phonological awareness intervention improves transition to reading and writing skills in children with LI (Gillon, 2006; Gillon & Moriarty, 2007). Several studies also suggest that the earlier an intervention is implemented in learning-disabled children, the better their prognosis. In addition, nonverbal cognitive ability accounts for some of the variance in outcome with intervention (Bishop & Edmundson, 1987). The current study examines level of improvement in expressive and receptive language, articulation, and oromotor (speech praxis) skills, as well as whether nonverbal cognitive ability, age, or pervasiveness of language impairment is associated with response to intervention.

Developmental Trajectory

Children with LI show a different developmental trajectory than children with typical language acquisition. On one end of the spectrum, these children may experience articulation problems or language delays that improve with remediation; on the other end, they may have severe deficits in grammar, syntax, and semantics that lead to the diagnosis of a language disorder. According to Leonard (1990), children with LI may have five different presentations of deficits. First, the difficulties may simply be a delay in language or some isolated areas of language that will ultimately reach the level of the child's peers. Second, acquisition of language skills may occur in the same sequence as in other children, but the delay may not be resolved, and the child with LI will reach a learning plateau. Third, the most common pattern observed is that the child with LI is developing more slowly than other children, but different language skills are developing at different rates, resulting in a mixed picture of severity of language deficits. Fourth, the child with LI may use isolated verbalizations that normal children also use, but with greater frequency (e.g., for all words beginning with 's', the child may say the 's' at the end of the syllable). Finally, the child with LI may use some verbalizations that are not observed in normally developing children (e.g., inhaling for certain consonants rather than verbalizing them). These different patterns may require different intensity and duration of intervention to achieve the best results.

The American Speech-Language-Hearing Association (ASHA; 2007) outlines specific risk factors that parents can look for to determine whether a child with language deficits in the 18- to 30-month-old range will continue to have difficulty. Children who show normal receptive language skills at this age tend to show a better prognosis in language learning while those with receptive language deficits show more prolonged language problems. Children who make greater use of physical gestures to communicate when they cannot verbally express an idea are also more likely to develop typical language skills later on. The older a child is when the diagnosis of a language disorder is made, the worse the prognosis. The 24- to 30-month-old range, when language skills typically develop very quickly, may increase the gap in learning even more between the older child with LI and his or her normally developing peers. Parents should also observe their child's language production and expect to see gains each month in characteristics such as the length of a child's utterances and the variety of purposes a word serves (e.g., "bottle" may indicate "that is my bottle" and later change to mean "I want my bottle").

Areas of Language Affected

Children with LI may have impairments in receptive language skills, expressive language skills, or both. Receptive language involves comprehension of input while expressive language involves quality of output. Receptive difficulties may be less apparent than expressive deficits, and the child may not follow instructions appropriately or may give responses to questions that are completely off topic, simply because they may not grasp the meaning of what is being said (American Psychiatric Association, 2000). A purely receptive language disorder that is developmental in nature (i.e., not due to acquired brain damage) is virtually never seen because these skills are required for expressive language to develop (American Psychiatric Association, 2000). Expressive language difficulties may be evidenced by articulation and phonological errors, repetition of syllables, and word-finding difficulties (American Psychiatric Association, 2000).

Language is made up of specific subsystems that govern how we communicate, including the *semantic* or *lexical* system, the *phonological* system, the *syntactic* or *grammatical* system, and the *pragmatic* system (Cantwell & Baker, 1991). Children with LI may experience difficulty with these various aspects of language and may show greater deficits in one area versus another (Leonard, 1990). They may demonstrate slow or limited vocabulary development and word-finding difficulties, which are both evidence of problems with semantics (Leonard, 1990). More specifically, examples of semantic or lexical difficulties include calling items by the wrong name (e.g., "finger" for "thumb"), using vague vocabulary (e.g., frequently saying "thing", "stuff" or "you know" instead of the precise word), difficulty retaining new words, trouble with abstract concepts such as time and space, and problems understanding metaphors, puns, or idioms (Cantwell & Baker, 1991). They may also have difficulty with organization or correct pronunciation of phonemes, or speech sounds in words, (e.g., saying "tee" for *see* or "sip" for *chip*) (Leonard, 1990). Phonological difficulties may be due to speech or articulation problems, problems with organization of phonemes, or oromotor problems, i.e., coordinating physical jaw, tongue, and facial muscle movements to create the correct sounds (Leonard, 1990). Syntax may be affected, which may be evidenced by shortened utterances, avoidance of using auxiliary verbs to express tense (e.g., “would have had”), difficulty sequencing words in a sentence, and trouble differentiating between similar-sounding words (Cantwell & Baker, 1991; Leonard, 1990). Difficulty using morphology, or the smallest meaningful parts of words, is seen frequently in children with LI, (e.g., difficulty with suffixes) (Leonard, 1990). Cohen (2001) explains that children with LI not only have structural language (words, sounds, and sentences) difficulties, but they also have problems with pragmatics, or using language that is appropriate to the context. Their utterances may be syntactically correct and intelligible, but they may seem tangential or socially inappropriate for the situation (Cantwell & Baker, 1991; Leonard, 1990).

Prevalence and Etiology

Boyle, Gillham, and Smith (1996) estimate that a developmental language disorder or delay occurs in approximately 7% of all children. Tomblin and colleagues give an estimate of 7.4% in kindergarten-age children (1997). Other prevalence estimates range from 1 to 15% in the child population, depending on the criteria used to define a language disorder (Law, Boyle, Harris, Harkness, & Nye, 2000). While speech and language disorders generally tend to affect more males than females, with ratios ranging from 1.2:1 to 2.3:1, two studies have found the reverse effect, with ratios as low as .46:1 for concurrent speech and language impairments (Law et al., 2000). One study found a higher prevalence of LI in African-Americans, but this finding was not replicated in other studies (Law et al., 2000). It should be noted that LI can change in severity over the course of development, so this dynamic process may affect prevalence estimates in other cohorts.

Language impairment has been given various monikers by professionals investigating different etiologies of the disorder, including neurological, developmental, psychological, linguistic, and perceptual origins (Cantwell & Baker, 1987, 1991). In order to illustrate various theories of etiology, it will be useful to examine terminology used to identify the disorder. Language impairment has been called *developmental childhood aphasia* (Cantwell & Baker, 1987) and *developmental dysphasia* (Cohen, 2001), which imply that neurological factors play a role in the disorder. Bishop (1994) notes that these terms can be confusing because they lead the reader to believe that some type of brain damage has been acquired. While some research has demonstrated that no significant structural differences exist between the brains of children with LI and those without impairment (Gauger, Lombardino, & Leonard, 1997; Preis, Lancke, Schittler, Huang, & Steinmetz, 1998), recent imaging studies have identified developmental brain abnormalities in children with LI. Plante and colleagues found perisylvian asymmetries not only in boys with LI (Plante, Swisher, Vance, & Rapcsak, 1991), but also in their parents and siblings (Plante, 1991). Several other studies have found atypical symmetry of the planum temporale, or atypical degrees of asymmetry of the planum temporale in children with dyslexia and/or LI (Foster, Hynd, Morgan, & Hugdahl, 2002; Hynd, Semrud-Clikeman, Lorys, Novey, & Eliopulos, 1990; Larsen, Høien, Lundberg, & Odegaard, 1990).

From a developmental perspective, the term *language delay* was initially used to describe children with a slower rate of language development than their peers. This term conveyed the expectation that these children would ultimately reach a normal level of language development for their age (Cohen, 2001). The term *language disorder*, on the other hand, was used to distinguish children who were not expected to approach a typical level of language development and would experience more severe difficulties. It has proven complicated, however, to differentiate a language delay from a disorder because while most problems associated with delay may resolve, residual academic difficulties may remain, especially in reading skills (Snowling, Bishop, & Stothard, 2000; Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998). In the mid-1900s, psychologists adopted terms related to auditory-processing such as *congenital auditory imperception* and *developmental word deafness* to describe language impairment (Cantwell & Baker, 1987). Linguists used terms describing the actual impaired behavior, including *linguistic delay* and *deviant language* (Cantwell & Baker, 1987).

Perceptual psychologists have posited that LI does not involve difficulty with linguistic concepts or rules of grammar, but rather it stems from problems with perception, discrimination, memory, or association (Cantwell & Baker, 1987). Current research on LI has investigated the role that higher-order cognitive processes play in the manifestation of LI, which is often accompanied by motor skills deficits and perceptual difficulties. While LI has been considered for some time to be a developmental anomaly that did not involve any type of neurological abnormalities, progressive neuroimaging studies have identified brain abnormalities that may be considered markers for language deficits (Preis, Engelbrecht, Huang, & Steinmetz, 1998; Preis, Lancke et al., 1998; Preis, Steinmetz, Knorr, & Jancke, 2000; Trauner et al., 2000; Watkins et al., 2002). Overall, the research on LI has not proposed a universally accepted etiology due to the complexity of the deficits involved and variety of skill areas affected. However, it has been suggested that LI and related reading disorders exist on a continuum of impairments characterized by the same etiology (Bishop & Snowling, 2004; Snowling & Hayiou-Thomas, 2006; Stackhouse & Wells, 1997). Other research has posited that an underlying etiology is responsible for LI and concomitant motor, sensory-integration, and perceptual skills deficits (Hill, 2001; Hill, Bishop, & Nimmo-Smith, 1998; Jancke, Siegenthaler, Preis, & Steinmetz, 2007; Trauner et al., 2000). For example, Beitchman, Wilson, Brownlie, Walters, and Lancee (1996) studied a cohort of children with and without LI by administering cognitive, language, and academic tests at age 5 and again at age 12. They investigated the stability of speech and language skills over time in four groups: children with high overall scores, poor articulation, poor comprehension, and low overall scores. They found that the high overall cluster performed best, followed by the poor articulation group, the poor comprehension group, and the poor overall group. This rank was consistent over time on cognitive, language, academic, and visual-motor integration tests, indicating that several areas of functioning are affected concurrently and continue over time without intervention. Deficits were also shown to occur with similar degrees of severity at ages 5 and 12, suggesting that underlying neurodevelopmental immaturity may explain the etiology of LI and concurrent deficits.

Cohen (2001) explains that *specific language impairment* (SLI) and the more general term *language impairment* (LI) are most frequently used in the current literature due to their neutrality with respect to etiology. While a consensus has not been reached concerning specific etiology, it is currently believed that both biological and environmental factors play a role in the development of LI (Cohen, 2001). Chronic otitis media, socioeconomic status, problems during pregnancy, and oral-motor difficulties have been associated with LI (Tomblin, Smith, & Zhang, 1997; Whitehurst et al., 1991). Children born to families with a history of LI have a greater risk of developing language problems than control children, and children of mothers who delayed or never received prenatal care are also at greater risk (Prathanee, Thinkhamrop, & Dechongkit, 2007). Genetic research has identified specific chromosomal abnormalities linked to LI, reading disorders, and problems with phonological awareness (Grigorenko, 2001; Grigorenko et al., 2003; Grigorenko et al., 1997; Spitz, Tallal, Flax, & Benasich, 1997; Tallal, Townsend, Curtiss, & Wulfeck, 1991; Wood & Grigorenko, 2001). Some investigators have found evidence for premature birth as a risk factor (Siegel, 1982; Weindrich, Jennen-Steinmetz, Laucht, Esser, & Schmidt, 1998), but others have found no significant association (Stanton-Chapman, Chapman, Bainbridge, & Scott, 2002). Other biological risk factors include male gender (Law et al., 2000), very low birthweight concurrent with broncopulmonary dysplasia (Rvachew, Creighton, Feldman, & Suave, 2005), temperament (Slomkowski, Nelson, Dunn, & Plomin, 1992), and infant otitis media with effusion (Polka & Rvachew, 2005). Environmental risk factors include low parental education level (Tomblin, Smith et al., 1997), bilingual home (Horwitz et al., 2003), low socioeconomic status (Horwitz et al., 2003), and maternal age below 18 years (Stanton-Chapman et al., 2002).

Classification

The literature is plagued by a lack of consistent criteria for classifying LI (Lahey, 1990; McCauley & Demetras, 1990; Wickstrom, Goldstein, & Johnson, 1985), primarily because these children constitute such a heterogeneous group (Aram, Morris, & Hall, 1993; Bishop, 1997). While the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR; American Psychiatric Association, 2000) defines five different subtypes of LI under the category of Communication Disorders, experts in the field have researched more clinically useful classification systems that encompass the nuances of discrete types of LI.

DSM-IV-TR Classification

The Diagnostic and Statistical Manual of Mental Disorders (APA, 2000) category of Communication Disorders includes Expressive Language Disorder, Mixed Receptive-Expressive Language Disorder, Phonological Disorder, Stuttering, and Communication Disorder Not Otherwise Specified. The defining features of Expressive Language Disorder include standardized test scores of expressive language that are substantially below standardized scores of nonverbal cognitive ability and receptive language. Clinical symptoms include an underdeveloped vocabulary, frequent errors in tense, and difficulty recalling words or producing utterances that are developmentally appropriate. Expressive Language Disorder is estimated to occur in approximately 10% to 15% of children under 3 years old and 3% to 7% of school-age children (APA, 2000). The criteria for Mixed Receptive-Expressive Language Disorder include additional impairment in receptive language based on test scores that are substantially lower than nonverbal cognitive ability scores. Children with added receptive language difficulties may have difficulty understanding words, sentences, or particular types of words. They may also have auditory processing deficits that lead to poor sound discrimination, difficulty associating word sounds with their symbols, and problems storing, recalling, and sequencing auditory information. Prevalence estimates include 5% of preschool-age children and 3% of school-age children (APA, 2000). Learning disorders in reading and writing, delayed or impaired motor skills, social withdrawal, and psychiatric disorders such as Attention-Deficit/Hyperactivity Disorder are commonly associated with language disorders.

The DSM-IV-TR defines Phonological Disorder (formerly Articulation Disorder) as difficulty with the accurate production of speech sounds, which may include substitutions of one sound for another, omissions of sounds, and lisping. When the disorder has been characterized as *developmental* due to unknown origin, prevalence in preschool children is 3%. Certain types of Phonological Disorder are commonly referred to in the literature as "developmental dyspraxia of speech" (APA, 2000, p. 65). Stuttering involves a disruption in the fluency and temporal accuracy of speech, including frequent repetitions or prolongations of words or syllables, and it is estimated to occur in 1% of preadolescent children.

Clinically Useful Classification Systems

Due to the various presentations of LI, researchers have deemed it necessary to investigate a clinically useful classification system that encompasses the complexity of the disorder. Cohen (2001) argues that the DSM-IV-TR classification has limited clinical utility and does not encompass the range of presentations seen in children with LI. A frequently cited classification system categorizes LI into three subtypes: 1) mixed receptive/expressive disorders, 2) expressive disorders, and 3) higher-order processing disorders (Rapin, 1996; Rapin & Allen, 1983). Children with mixed receptive/expressive disorders show impairment in phonology, syntax, and semantics. Those with little to no language comprehension may be nonverbal, and those with less severely impaired comprehension may exhibit limited speech that is nonfluent, poorly intelligible, and lacks appropriate grammatical structure. Children with expressive disorders and sufficient comprehension skills primarily show deficits in speech phonology. They may also exhibit varying degrees of verbal dyspraxia, the most severe of which renders these children nonverbal with intact comprehension. Children with higher order processing disorders show deficits in semantics, pragmatics, and discourse. Preschool-age autistic children usually show deficits in semantics and pragmatics. Van Weerdenburg, Verhoeven, and van Balkom (2006) have devised a statistically-based typology. They administered a comprehensive battery of standardized language tests to 147 six-year-old and 136 eight-year-old children with specific language impairment (SLI) in the

Netherlands and conducted a factor analysis to define four different typologies. The analysis revealed four distinct subtypes of impairment in both groups of children: 1) lexical-semantic impairment, 2) auditory conceptualization, 3) verbal sequential memory, and 4) speech production (Van Weerdenburg et al., 2006). Lexical-semantic skill deficits included difficulty with knowledge of word meanings and ability to understand words in text. Auditory conceptualization involved using colored blocks to represent the number, similarity, and order of phonemes that were verbally presented to the child. Verbal sequential memory involved tasks measuring ability to recall and correctly sequence phonemes. Speech production included measures of word repetition, repetition of pseudo-words, and articulation. These distinct factors demonstrate the possibility that children with SLI may have deficits in more than a single area of language, and some areas may be more impaired than others (Van Weerdenburg et al., 2006). This classification system serves as evidence that language interventions should be flexible enough to accommodate the needs of children with deficits in various aspects of language (Van Weerdenburg et al., 2006).

Due to the small sample size used in the current study (N=20), it may prove difficult to divide this sample into the four groups proposed by Van Weerdenburg and colleagues (2006). In addition, the four group scheme is based on additional measures of auditory conceptualization and verbal sequential memory that were not used in the current study. Based on previous research showing that predominantly expressive versus receptive language deficits predict improvement with language intervention, the most useful classification scheme may be a dichotomous one borrowing from Rapin and Allen's

(1996) three-group scheme and excluding the group with pragmatic language disorder due, again, to the fact that specific measures of pragmatic deficits were not used in this study. Children with mixed receptive-expressive difficulties and those with predominantly expressive difficulties will be identified in the current study to examine associations between these types of language deficits and level of improvement with intervention.

Assessment and Diagnosis

Assessment procedures for diagnosing a language disorder include a standardized measure of nonverbal cognitive ability with which to compare various tests of expressive and receptive language. While the DSM-IV-TR states that standardized expressive and/or receptive language test scores must be "substantially below" (APA, 2000, p. 64) those of standardized measures of nonverbal cognitive ability for a formal diagnosis of a language disorder, the size of the discrepancy between language and nonverbal ability scores is open to interpretation. The American Academy of Child and Adolescent Psychiatry (AACAP; 1998) outlines practice parameters for assessment of language disorders in order to clarify the discrepancy necessary for a child to receive additional educational services under the Individuals with Disabilities Education Act (IDEA). These guidelines explain that while each state is able to interpret criteria for diagnosis of a language disorder or learning disability and develop its own cutoff scores, general practice should deem any child with language skills below an age-appropriate level as in need of assessment and possible intervention. A language disorder may be diagnosed when language tests yield scores -2 to -1.25 standard deviations below the mean, which translates to standard scores of 70 to 81, respectively. This criterion has been used for several prevalence studies of language disorders (Law et al., 2000). Generally, language scores are significantly below nonverbal cognitive ability scores when diagnosing a language disorder; however, the AACAP warns against denying additional educational services to children who do not necessarily meet criteria for a specific language or learning disability because they show general low-achievement that is commensurate with a borderline IQ. Several studies also present evidence against the practice of using cognitive referencing to diagnose language disorders (e.g., Cole, Coggins, & Vanderstoep, 1999; Cole, Dale, & Mills, 1990; Dethorne & Watkins, 2006). Research has shown that a nonverbal IQ cutoff score of ≥ 85 is arbitrary (Swisher & Plante, 1993; Swisher, Plante, & Lowell, 1994), children above and below the cutoff do not vary in their pattern of language deficits (Tomblin & Zhang, 1999), and children with LI show deficits in nonverbal ability as well (Leonard, 1998).

Nonverbal cognitive assessments include tests such as the Leiter International Performance Scale – Revised (Leiter-R; Roid & Miller, 1997) and the newly developed Wechsler Nonverbal Scale of Ability (WNV; Wechsler & Naglieri, 2006), in which instructions can be pantomimed by the examiner or are pictorially represented. Nonverbal IQ tests are used in lieu of verbal measures to virtually eliminate the need for verbal communication, which is impaired in children with language disorders. While nonverbal IQ is generally seen as the more accurate representation of the language-impaired child's intellectual ability, a general factor responsible for verbal and nonverbal abilities is implicated in several models of intelligence. For instance, Carroll's three-stratum theory of intelligence (1993), based on factor analysis of a host of cognitive measures, consists of Stratum III, a general factor that plays a role in all cognitive activity, Stratum II, including broad intellectual abilities such as fluid and crystallized intelligence, and Stratum I, 70 narrower cognitive abilities that contribute to Stratum II. Tests of receptive language may include identification tasks that require the examinee to choose a picture or object in response to a question; these tasks test the child's understanding of language input. Examples of receptive tasks include the Peabody Picture Vocabulary Test – 3rd Edition (Dunn & Dunn, 1997), and the Receptive One-Word Picture Vocabulary Test (Brownell, 2000b), in which the child is given a vocabulary word and must point to one out of several picture choices that depicts the word. Tests of expressive language may include tasks requiring the child to name an object or action depicted in a picture, as in the Expressive One-Word Picture Vocabulary Test (Brownell, 2000a). These tasks test the child's ability to produce language output. The Clinical Evaluation of Language Fundamentals-Preschool (CELF:PS-2; Wiig, Secord, & Semel, 2004) and the CELF-4 (Semel, Wiig, & Secord, 2003) consist of several language tasks that yield receptive and expressive indices of language.

Connection Between Language and Reading Skills

Language impairment and reading skills are clearly related. In a longitudinal study conducted by Catts and colleagues, 25% of a sample of 527 kindergarten children diagnosed with SLI met criteria for a diagnosis of dyslexia in second, fourth, and eighth grades (Catts, Adlof, Hogan, & Ellis Weismer, 2005). Just below 20% of dyslexic children in the later grades had been diagnosed with SLI in kindergarten (Catts et al., 2005). Flax and colleagues investigated co-occurrence of SLI with reading impairments in children and their families (2003). They found that family members of children with SLI were more likely to have the combination of oral language impairment (LI) and reading impairment (RI) than either one alone. Sixty-eight percent of the children with SLI also met diagnostic criteria for dyslexia. The LI rate for family members was 25%, and the RI rate was 23%; comorbidity of the two disorders in family members was 46%. The families also had more male children than female children, and more male children were affected by both LI and RI, which is consistent with the trend in prevalence rates.

Phonological awareness, which is related to decoding of words by breaking them down into sound phonemes, is the main skill deficit that links LI to dyslexia (e.g., Snowling & Hayiou-Thomas, 2006; Stackhouse & Wells, 1997). A deficit in phonological processing is closely associated with dyslexia, but it is not associated with SLI that occurs in the absence of dyslexia (Catts et al., 2005), demonstrating its pivotal role in the overlap of these two disorders. Snowling and Hayiou-Thomas (2006) explain that several hypotheses have been proposed to describe the relationship between dyslexia and SLI, including the *severity hypothesis*, which states that SLI is a more severe form of dyslexia (Snowling et al., 2000); the *critical age hypothesis*, which speculates that if the child with SLI has language delays that continue to school age, they will develop reading delays (Bishop & Adams, 1990); and the hypothesis that they are coexisting disorders (Catts et al., 2005). Bishop and Snowling (2004) have conceptualized the two disorders as existing on a continuum of impairment, such that a two-dimensional model of phonological/nonphonological skills and impaired/not impaired levels can better account for the heterogeneity of skill deficits observed within each disorder. Their model postulates that children with intact phonological skills (PS) and nonphonological skills (NPS) have no impairment, those with intact PS and impaired NPS are poor comprehenders of language, those with impaired PS and intact NPS have dyslexia, and those with impaired PS and impaired NPS have SLI.

Botting, Simkin, and Conti-Ramsden (2006) assert that children with LI are at very high risk for developing reading difficulties as they near high school age: 80% of the 7-year-old language impaired participants in their study showed reading comprehension problems at age 11. The researchers stress that children who have impaired reading skills in addition to impaired language skills are at a distinct disadvantage in their education, which is clearly dependent upon the ability to read. In order to identify which factors differentiated the children who had no literacy impairment at age 11, they found that while approximately half of their sample showed a decrease in IQ (from > 85 at age 7 to < 85 at age 11), only one out of 33 children with average literacy skills at age 11 showed the same type of decline. Therefore, it appears that IQ is also related to reading skill development.

Concurrent Impairment in Motor Skills and Sensory-Integration

Several studies have suggested an association between developmental LI and deficits in other domains, including motor skills, sensory integration, and perceptual skills (Estil, Whiting, Sigmundsson, & Ingvaldsen, 2003; Gillberg, 1998; Hill et al., 1998; Mandelbaum et al., 2006; Trauner et al., 2000). Reviews of the literature by Hill (2001) and Webster and Shevell (2004) describe a substantial body of evidence that children with LI experience deficits that are not specific to language. Hill (2001) offers a theoretical explanation involving slower information processing in children with LI, which has been shown to occur not only in the language domain, but also on auditory processing and fine motor tasks. A decreased capacity for information processing may also explain why children with LI show deficits on both linguistic and nonverbal tasks (Hill, 2001). Immature brain development has been posited as an explanation for language and motor deficits that occur in other developmental disorders such as attention-deficit/hyperactivity disorder and dyslexia (Hill, 2001). Webster and Shevell present evidence that specific language impairment is often "not 'specific'" (2004, p. 479), given research that has shown associated impairments in nonverbal cognitive ability, motor skills, attention, slowed auditory processing, and impaired short-term memory and auditory discrimination.

Mandelbaum and colleagues (2006) compared sensory and motor performance of children with developmental language disorder (DLD) to that of autistic children with low nonverbal IQ, autistic children with high nonverbal IQ, and non-autistic children with low nonverbal IQ. Gross motor skills tasks included stressed gaits (i.e., walking on the heels or toes or hopping on one foot); balance (i.e., standing with feet together and eyes closed); and persistence (i.e., maintenance of posture with eyes closed). Fine motor skills were measured by asking the child to tap the foot or fingers, write his or her name, and complete a pegboard task. Oromotor skills were measured by asking the child to repeat a series of syllables and demonstrate facial movements such as placing the tongue in the cheek. Sensory skills included finger localization and discrimination between a penny and dime with eyes closed. The researchers found that the children with DLD and the High IQ autism group scored better than the Low IQ autism and Low IQ groups. The High IQ autism group also scored better than the DLD group in sensory and motor skills, oromotor skills, and praxis skills.

Another study of neurological examinations of children with developmental LI showed that 70% of these children showed motor abnormalities, as compared to 22% of control children (Trauner et al., 2000). These abnormalities included obligatory synkinesis (involuntarily moving one muscle while intentionally moving another), fine motor impairment, and hyperreflexia (overresponsive reflexes). The severity of these abnormalities was significantly correlated with the severity of language deficits as measured by standardized testing. Thirty-four percent of children with LI also showed abnormalities on MRI, including right ventricular enlargement, central volume loss, and multiple areas of white matter hyperintensity; none of the control subjects had abnormal scans. Jancke, Siegenthaler, Preis, and Steinmetz (2007) confirmed the relationship between language and motor impairments and neuroanatomical atypicalities. They used voxel-based morphometry techniques in MRI scans to show reduced white matter volume in the left hemisphere of children with developmental language disorder as compared to controls. Specific areas affected included the motor cortex, dorsal premotor cortex, ventral premotor cortex, and planum polare on the superior temporal gyrus. Children with language deficits also showed impairment in motor skills on standardized motor assessments, suggesting that motor and language impairments are linked in the neuroanatomy of affected children and therefore impaired to the same degree.

Despite the apparent association between LI and motor skills, the underlying general processing deficit hypothesis does not fully account for the specificity of language deficits seen in some children with LI, and nearly any deficit could be explained by problems with underlying processing (Ullman & Pierpont, 2005). The population of children with LI is decidedly heterogeneous in that not *all* children with LI have motor weaknesses. Ullman and Pierpont posit their own Procedural Deficit Hypothesis to account for deficits in procedural memory caused by abnormal development of certain brain structures to account for the association between language and motor problems. Given the debatable explanations for language and motor deficits, the current study does not assess underlying etiology of LI and motor impairments and seeks to explain the effects of language intervention on language skills only.

CHAPTER THREE
Language Interventions

A multitude of interventions has been devised for children with LI, ranging from traditional speech and language therapy that targets expressive syntax and phonological skills to supplemental nonverbal communication gestures to computerized training modules. The age ranges of participants, severity of language impairment involved, intervention methods, and instruments used to measure improvement in language intervention studies vary a great deal within the massive amount of literature available, making meta-analyses of studies with comparable methodology difficult to conduct. Several interventions are highly specialized and often require years of training or a master's degree in speech-language pathology to administer. Some more recently developed interventions also require individualized technology such as computer programs and voice output communication aids (VOCAs) to administer, making these interventions difficult to replicate without adequate resources. Cirrin and Gillam (2008) emphasize the need for evidence-based practice (EBP) of speech and language interventions to justify their use. They note that no published guidelines currently exist to govern the use of speech-language interventions with school-age children. Current empirically-investigated oral language interventions will be explored, followed by detailed descriptions, theoretical bases, and empirical studies relevant to the intervention components of the current study: Montessori Applied to Children At-Risk for a learning disability, the DuBard Association Method, and Occupational Therapy with incorporated sensory integration.

Reviews & Meta-analyses

Several reviews of the literature and meta-analyses have been conducted in an attempt to consolidate empirical research on the efficacy of speech-language interventions for children. These comprehensive articles usually establish inclusion criteria for accepting studies of high methodological quality into the review. For example, Cirrin and Gillam (2008) only accepted peer-reviewed articles published since 1985 that investigated school-aged children with a primary *spoken* language (as opposed to *reading*) disorder, used valid and reliable measures of treatment outcome, and consisted of either a randomized clinical trial (RCT), meta-analysis of RCT's, systematic review of RCT's, nonrandomized comparison study, or multiple-baseline single-subject design. They also assigned an appraisal-point rating to the publications included, based on whether the studies included random assignment to treatment conditions, a comparison control group, descriptive characteristics of participants, initial degree of similarity between groups, blinding to treatment conditions, quality of outcome measures, statistical significance levels reported, and practical significance (Cirrin & Gillam, 2008).

Other systematic reviews and meta-analyses of language interventions include those by Gallagher (1998); Leonard (1998); Yoder and McDuffie (2002); and Law, Garrett, and Nye (2003). While these reviews generally provide support for the efficacy of speech and language therapy, several concerns remain about varying methodological quality of studies included and the possibility of exaggerated effect sizes due to the statistical aggregation procedures used (Cirrin & Gillam, 2008; Law et al., 2004). Law, Garrett, and Nye's (2004) meta-analysis of 13 studies reported a positive effect of intervention when children had expressive vocabulary or phonological difficulties. Interventions for expressive syntax problems showed mixed results, and evidence for receptive language interventions were inconclusive due to the limited number of studies analyzed. No significant differences were found for clinician-led interventions versus those led by trained parents. Interventions more than eight weeks in duration led to better outcomes than shorter interventions. These conclusions, however, should be interpreted with caution because the authors noted that results of their meta-analysis were highly heterogeneous, warranting further investigation into sources of variation. Cirrin and Gillam (2008) further note that all but two of the 36 articles reviewed by Law, Garrett, and Nye (2004) studied children under the age of five, limiting generalizability of the findings to school-aged children, and their analysis of effect sizes collapsed intervention methods across targeted treatment areas, which does not allow clinicians to determine the specific intervention procedures that are most effective.

Empirically Investigated Language Interventions

Recently investigated speech-language interventions that utilize relatively sound methodological study design (repeated measures design or controlled trials) and were indicated for use with children ages 3 to 12 (the age range of the current study from beginning to end of intervention period) will be consolidated in this literature review. The following studies were cited in a meta-analysis of language interventions for preschool-aged children by Law, Garrett, & Nye (2004), a systematic review of interventions for school-aged children by Cirrin and Gillam (2008), and a comprehensive text of various language interventions edited by McCauley and Fey (2006). Due to the highly specialized nature of several of these studies, a sampling of individual studies was chosen based on those that would be appropriate for the population in the current study, i.e., they apply to the same age range and focus on remediating oral language deficits that involve articulation, syntax, vocabulary, and processing as opposed to pragmatics. The studies will be organized according to the specific deficit to be remediated: syntax/morphology, vocabulary/semantics, phonological awareness, and language processing. Augmentative and alternative communication, which involves nonverbal hand gestures and signs to supplement language; purely pragmatic language interventions, which focus on improving the child's ability to use language appropriate to specific contexts; and interventions with the primary goal of improving reading and written language will not be included in the current review due to being outside the scope of the current intervention goals of improving verbal language components. See Table 1 for an overview.

Syntax and Morphology

Several language intervention studies focus on development of syntax (grammar) and morphology skills. Bishop, Adams, and Rosen (2006) investigated the efficacy of a computer-based intervention for 36 children ages eight to 13 with receptive language difficulties. Computer training games were designed to improve sentence and story comprehension by teaching children to move objects on a computer screen to match spoken sentences of increasing grammatical complexity. Children in the two experimental conditions heard either slowed speech with a 1.2 second delay between sentence phrases, or modified speech based on FastForWord – Language software (FFW-L; Scientific Learning Corporation, 1998), which slows speech down and alters length and amplification of certain sounds depending on progress of the individual child. The experimental groups received computer-based training in 15-minute sessions over the course of 20 days, and the control group received no training. While the trained children's responses showed an increase in speed with a moderate effect size (.401), no significant differences were found between trained groups and the control group on language or auditory processing measures, including the Test for Reception of Grammar – 2 (Bishop, 1989) and mean length of utterance (MLU). While the researchers did not assess effect size from pre-test to post-test, Cirrin and Gillam (2008) calculated these values based on means and standard deviations, finding moderate improvements for experimental and control groups on the sentence comprehension measure (.37 to .77) and an overall decline on the MLU measures (-.03 to -.84). The computer-based intervention, therefore, did not appear to produce better results than language intervention methods already in place at the children's school. Cole and Dale (1986) compared the effects of direct language instruction, which involves imitation of language components presented by the teacher and specific positive reinforcement, to interactive language instruction, which involves modeling of language rules with the goal of the child's applying the rules to his/her own utterances in a natural context and uses natural responses to the child's language as opposed to positive reinforcement. Forty-four children ages 3 to 5 participated in the intervention groups, which were monitored for procedural validity to ensure that the types of intervention were reliably different on dimensions such as amount of teacher-directed instructional activities and elicited responding. Despite significant differences between the treatment types, no significant differences were found between the two treatment groups on language tests such as the Peabody Picture Vocabulary Test – Revised (PPVT-R; Dunn & Dunn, 1981) and the Preschool Language Scale – Revised (PLS-R; Zimmerman, Steiner, & Pond, 1979). Significant differences were found from pretest to posttest on every language measure except the Developmental Sentence Scoring analysis (Lee, 1974), which included a smaller number of participants because only 27 children produced enough sentences during a 30-minute interactive session with the examiner to compute the analysis from pre- to posttest. Since no control group was used, the authors used developmental quotients, which are not influenced by chronological age, for the PLS-R and the PPVT-R to rule out influence of the confounding variable of maturation on changes observed with intervention. While Language Quotients on the PLS-R increased significantly from pretest to posttest, PPVT-R scores increased non-significantly. Therefore it remains inconclusive whether results were due to maturation or the intervention itself.

Weismer and Murray-Branch (1989) used a multiple-baseline single-subject design to study children's production of grammatical targets after two types of intervention: a modeling condition in which an adult modeled the targets but did not require a response from the child, and a modeling plus evoked production condition, in which an adult both modeled the targets and allowed the child to respond, then provided the child with feedback on his/her accuracy. Both intervention types worked equally well and showed moderate to large effect sizes with the percentage of non-overlapping data statistic (PND between .64 and 1.00) for the three children who displayed expressive language deficits, but for the child with mixed receptive-expressive deficits, neither intervention was effective (PND = .20). Connell and Stone (1992) also compared a modeling condition to an imitation condition using novel, invented morphemes. The morphemes consisted of one of four suffixes attached to concrete nouns represented by pictures (e.g., "TVum" for a broken TV). Modeling consisted of an adult providing a novel morpheme in a meaningful way and instructing the child to give a command to a cartoon character on a computer; the imitation condition involved modeling the novel morpheme and having the cartoon character then instruct the child to imitate that specific morpheme. Children were also probed for understanding of the morphemes. Control children performed the tasks with relative ease in both conditions, but children with LI showed minimal use of the morphemes with modeling alone and significantly more use following the imitation instruction. Cirrin and Gillam (2008) computed effect sizes by dividing the amount of learning by the standard deviation since novel morphemes were heard only during treatment. For children with SLI in the modeling condition, a moderate effect (d = .78) was shown for production of morphemes, and a large effect (d = 1.32) was shown for comprehension. For children with SLI in the imitation condition, a moderate effect (d = .5) was shown for production, and a large effect (d = 1.3) was shown for comprehension. Both modeling and imitation interventions were equally effective for improving the language-impaired children's comprehension of morphemes. Dixon, Joffe, and Bench (2001) investigated the effectiveness of the Visualizing and Verbalizing (V & V) program, which was developed by Bell (1987) for use with children with predominantly receptive language difficulties. Bell (1991) proposed that children with language comprehension problems have trouble forming a gestalt of presented information due to weak imagery skills. Through the V & V intervention, children are trained to use "structure words" to describe the distinguishing features of a pictured object, e.g., shape, size, perspective, etc. They then move on to the more abstract task of visualizing familiar and fantasy objects and describing them using the structure words. The clinician then verbally presents sentences containing the previously imaged nouns, and the child is required to visualize these sentences. Next, the clinician presents sentences and asks questions to call on the child's ability to visualize and verbalize features of his/her sentence imagery. The child is then asked inferential and main idea questions about the overall gestalt of the paragraph presented, with the ultimate goal of improving listening and comprehension skills. Dixon, Joffe, and Bench (2001) used a counterbalanced treatment design to compare the effects of V & V to traditional speech-language therapy on eight children ages 9 to 15:1. Two of these children received the V & V intervention for ten weeks, two received traditional speech-language therapy for 10 weeks, and four children received traditional therapy for five weeks followed by five weeks of the V & V intervention. Using the Analytic Reading Inventory (Woods & Moe, 1995), the authors found significant differences of both interventions from pre- to posttest, but no additive treatment effects of V & V were found over traditional therapy. The authors postulated that V & V may have proven more beneficial over a longer period of time; they also noted that statistical power was limited due to such a small sample size.

Fey, Cleave, Long, and Hughes (1993) compared the effectiveness of a clinician- vs. parent-administered approach to remediating grammatical impairment in 30 children ages

3 years, 8 months to 5 years, 10 months. Children were randomly assigned to either the clinician- or parent-administered intervention group that lasted 4.5 months, or the delayed treatment condition that postponed treatment for 4.5 months. Clinicians met with children individually for one hour per week and in two one-hour group sessions per week with other children in a clinical setting. Parents conducted intervention with their children individually in the home setting, and the clinician visited to give feedback and demonstration of correct intervention techniques. The intervention consisted of a focused-stimulation approach, in which the clinician or parent modeled specific grammatical targets and provided opportunities for the child to produce these targets. The clinician or parent also “recast” the child’s attempts to produce a target by asking questions to correct the target (e.g., Child: “Him like that”; Adult: “Does he like that?”). Children were also read a story each week using these language remediation techniques. Further study on focused stimulation and conversational recast intervention is reviewed in McCauley and Fey’s (2006) comprehensive text. In addition to focused stimulation and recast, a cyclical goal-attack strategy was used, in which individualized grammatical goals were created for each child (e.g., accurate use of *will, is, the, I, she*), and one goal was stressed each week. Using the DSS (Lee, 1974) as the primary dependent measure of production of appropriate grammatical forms, results for the two service delivery models indicated that test scores for the clinician- and parent-administered approaches were significantly better at posttest than the delayed treatment group, demonstrating large effect sizes (.81 and .96, respectively), but no significant differences were found between these service delivery models. Some evidence was found to suggest that the clinician-administered treatment produced more consistent gains over time than the parent-led

intervention. However, specific components of the interventions (focused stimulation procedures & cyclical goal-attack strategies) were not treated as experimental variables, so it cannot be determined whether these specific elements or other facets of the intervention, e.g., clinical vs. home setting and individual vs. mixed individual and group components, are necessary to improve grammar. A second phase of this study conducted by Fey, Cleave, and Long (1997) investigated whether gains would be maintained without intervention or with five additional months of clinician or parent intervention. Results suggested that both treatment groups' DSS scores were reliably higher after Phase 2 of intervention, but the size of these gains were not statistically different from the size of Phase 1 gains for each group. The clinician group's size of Phase 2 gain in the mean main verb score of the DSS was significantly smaller than the size of Phase 1 gain. The dismissal group showed no reliable change from post-Phase 1 to post-Phase 2.

Overall, syntax and morphology interventions lack consistency in their methods, measures, and the age groups studied, making it difficult to draw broad conclusions about their efficacy. Some evidence exists for the effectiveness of syntax interventions involving imitation, modeling, modeling with evoked production, and focused stimulation with cyclical goal-attack strategies, but replication of these studies is needed to build a stronger evidence base. Computerized grammar training for children with receptive LI does not appear to produce better results than school-based language therapy. Speech-language pathologists with the goal of remediating syntax have little quality evidence on which to base a specific intervention approach.

Semantics and Vocabulary

Several studies focus on remediation of semantics and vocabulary. Crowe (2003) compared the effects of Communicative Reading Strategies (CRS; Norris, 1988) to traditional reading decoding strategies on the expressive and receptive vocabulary skills of children with oral and written language-learning disabilities (LLD). Traditional reading decoding intervention involves the interventionist's providing cues to sound out, reread, or divide a word, giving phonemic cues, and ultimately providing the word itself. CRS is an approach in which the teacher or clinician assists the child with understanding the author's message and constructing overall meaning of what is read rather than simply providing feedback on how to decode or articulate single words. The child reads passages orally so that the teacher may identify problems with comprehension and clarify meaning of vocabulary or complex sentences. Twelve children with LLD who were identified by their school district as poor readers were assigned to either a reading decoding treatment group, a CRS treatment group, or a control group. All groups continued to receive regular reading instruction and any special language instruction services as delivered by school staff. Scores on the Comprehensive Receptive and Expressive Vocabulary Test (CREVT; Wallace & Hammill, 1994) showed that children in the CRS group showed more improvement on the posttest of expressive vocabulary than the other two groups, with a moderate effect size of .5. All groups showed decreased performance on the receptive vocabulary measure, suggesting that the alternate form of the CREVT used at posttest may have been more difficult than the pretest form. The small sample used in this study limits statistical power of the results. Throneburg, Calvert, Sturm, Paramboukas, & Paul (2000) tested the effects of different service delivery models of vocabulary intervention with children in kindergarten through third grade. A vocabulary intervention focusing on 60 target vocabulary words was administered in three different ways. In the collaborative approach, a speech-language pathologist (SLP) and classroom teacher planned and administered the intervention together. In the classroom-based approach, the SLP and teacher planned and administered their lessons separately in the classroom setting. The traditional pull-out approach involved the SLP's pulling the child from class for a 50-minute session once a week. Students in the collaborative condition scored highest on a 20-item vocabulary test after the intervention, but all three groups showed large effect sizes from pretest to posttest. However, this study did not include a no-treatment control group, leaving the possibility that gains were due to outside variables, such as word-learning achieved outside of the vocabulary lessons. Weismer and Hesketh (1993) investigated the effects of three different types of verbal presentation of novel words on both word production and comprehension skills in children with SLI ages 5 years, 1 month to 6 years, 7 months. One condition involved variation of the rate of verbal presentation of the word, using a slow, normal, or fast rate; the second condition involved the presence or absence of emphasis on certain syllables of the word; and the third condition involved the presence or absence of iconic gestures to accompany the word. Effect sizes were computed by Cirrin and Gillam (2008) by dividing the amount of learning by the standard deviation since the invented words were heard only during treatment. Large effect sizes of 1.1 were calculated for children with LI on measures of both comprehension and production when words were presented at a slow rate. The presence of emphatic stress had a moderate effect on production of novel words, but no effect for comprehension. A moderate effect was shown for comprehension of words when presented with accompanying gestures, but only a small effect was noted for production of words in this condition. Due to the relatively small sample size used in this study (N=16), the analyses have limited statistical power.

Barratt, Littlejohns, and Thompson (1992) investigated the effects of two service delivery models of traditional speech-language therapy on expressive and receptive vocabulary outcomes of 39 children ages 3 years, 1 month to 3 years, 7 months. Speech-language pathologists administered therapy either on a weekly basis (40 minutes per session over the course of six months) or an intensive schedule (40-minute sessions four days a week for three weeks, in each three months of a six month period). Using the expressive and receptive vocabulary composites of the Reynell Developmental Language Scales, the authors found that the weekly group improved significantly on the comprehension composite ($p=.02$) while the intensive group's improvement nearly reached significance ($p=.07$). The intensive group improved significantly on the expression scores ($p<.01$), but the weekly group did not ($p=.18$). Effect sizes were not calculated for this study, and no control group was used; therefore, conclusions cannot be drawn about the magnitude of improvement and whether confounding variables played a role in outcome. Wing (1990) designed an intervention to improve word-finding difficulties in children with severe LI. Ten children ages 5 years, 11 months to 7 years, 1 month were assigned to either a semantic intervention, in which children assigned vocabulary words to categories or described their functions or attributes, or a phonological intervention, in which children counted syllables and phonemes in words and found rhyming words. Children received treatment for 2.5 months, and the phonological treatment group improved significantly, with a moderate effect size of .7 for the treatment. The semantic treatment group did not improve significantly, but Cirrin and Gillam (2008) calculated a moderate effect size (d = .6) for this treatment as well. No control group was used in this study, and the small sample size limits statistical power of the results.

Overall, semantic and vocabulary interventions, while executed using a variety of strategies, produce moderate to large effect sizes. Reading strategies focusing on comprehension of the gestalt produce moderate effects on expressive vocabulary, and vocabulary intervention administered by a speech-language pathologist and classroom teacher produces the largest effect sizes when administered collaboratively. It seems that a slower presentation rate of novel words improves both production and comprehension in children with LI, with large effect sizes. Intensive speech-language therapy seems to be preferable to weekly sessions, and moderate effect sizes result from phonological and semantic approaches to vocabulary intervention. These results should be interpreted with caution due to varying methodological quality of the studies reviewed.

Phonological Awareness

Phonological awareness (PA) training has been presented by Gillon (2006) as a framework for prevention of delays in reading, writing, and spelling problems in preschool-aged children with oral language deficits. PA, which focuses on the sound structure of speech, is the most implicated skill deficit noted in children with reading disorders, or dyslexia, and it is an integral link between oral and written LI (Catts, Fey, Zhang, & Tomblin, 1999; Larrivee & Catts, 1999; Stackhouse & Wells, 1997). Gillon proposed that 3- and 4-year-old children with speech-language difficulties who do not have accompanying sensory, neurological, physical, and intellectual problems or behavioral/emotional disorders would be the appropriate population to receive PA intervention due to their increased risk for reading difficulties (Gillon, 2006). PA is approached at three levels: 1) awareness that words consist of syllables; 2) onset-rime awareness (syllables have a beginning part, or onset, and a rime unit), which is necessary for recognizing and producing rhyming words; and 3) awareness that words are formed from individual speech sounds, or phonemes (Gillon, 2006). Specific tasks used in PA training include choosing the word that does not rhyme from a set of three pictures, matching words that rhyme, choosing the word from three pictures that starts/ends with a different sound, choosing the word that starts with the same sound as the target word, blending syllables to form words, and deleting a syllable from a word. Exercises include play activities that encourage phoneme awareness [e.g., "Let's find all the toys that start with the /k/ sound" (Gillon, 2005, p. 315) and "My friend 'munching monkey' is going to eat the pictures that start with an /m/ sound. Let's help him find the pictures." (Gillon, 2005, p. 324)]. Standardized assessments such as the Preschool and Primary Inventory of Phonological Awareness (PIPA; Dodd, Crosbie, MacIntosh, Teitzel, & Ozanne, 2000) and Phonological Abilities Test (PAT; Muter, Hulme, & Snowling, 1997) can be used to monitor the progress of PA skills in young children.

While the positive effects of PA training on reading and writing skills have been well-documented for children with reading delays (see Ehri et al., 2001 for a review), much of the evidence base has excluded children diagnosed with speech-language impairment. However, some early evidence suggests that phonological awareness training is beneficial for children with oral language impairments. Van Kleeck, Gillam, and McFadden (1998) studied the effects of rhyming and phoneme awareness training on 24 preschool-aged children with speech and/or language disorder. They received group instruction twice a week for two semesters; the fall semester focused on the recognition, identification, judgment, and generation of rhyming words, and the spring semester focused on teaching the participants to identify and match initial sounds or words, produce words that start with target sounds, blend sounds, and use the sounds learned to make new words. The children in the experimental group were compared to a group of older control children who had attended the same preschool classes the previous year. The children who received treatment obtained scores 1.5 standard deviations above the control children on standardized measures of phoneme awareness, constituting a large effect size of treatment. Using a sample of 91 New Zealand children ages 5 years, 6 months to 7 years, 6 months, Gillon (2000) compared PA intervention to traditional intervention and minimal consultation in a group of children with LI and a group with normal language development; an 11-month follow-up study was conducted to measure maintenance of improvements (2002). The PA intervention consisted of weekly one-hour sessions, totaling 20 hours of intervention focusing on phonological identification, phoneme manipulation and segmentation, grapheme-phoneme correspondence, and phoneme production. The traditional intervention involved 20 weekly one-hour lessons on phoneme production in isolation, syllables, words, and phrases. The minimal consultation approach involved monthly meetings about phoneme production between

SLPs, parents, and teachers. Based on standardized speech production measures, PA measures, and reading measures, Gillon found that children in the experimental group made significantly more improvement than the other intervention groups on PA tests, including the Lindamood Auditory Conceptualization Test (LACT; Lindamood & Lindamood, 1979) and the Queensland University Inventory of Literacy (Dodd, Holm, Oerlemans, & McCormick, 1996). There was no significant difference between the experimental group's test scores and the typically developing children's test scores on the LACT post-intervention. The experimental group also made significantly more improvement than the other groups on standardized reading measures. As compared to traditional intervention, PA intervention produced a large effect size on a measure of phoneme awareness (2.58) and a moderate effect size on a reading comprehension measure (.67) (Cirrin & Gillam, 2008). As compared to consultation only, the phonological awareness intervention produced a large effect size on the phoneme awareness measure (1.77) and a moderate effect size on rhyming (.67) (Cirrin & Gillam, 2008). An 11-month follow-up study on a group of 20 of the original language-impaired children receiving phonological awareness, 20 children who received traditional or consultation intervention, and 20 typically developing children showed that improvements were maintained for children who received PA intervention in the areas of phoneme awareness and word-recognition (Gillon, 2002). Children who were in the experimental group also exhibited reading ability at or above the level expected for their age on the Burt Word Reading Test (Gilmore, Croft, & Reid, 1981), as well as maintenance of improved non-word spelling ability, which shows evidence of making phoneme-grapheme connections (sound-units to written units of language). A large effect size (2.42) was calculated by Cirrin and Gillam (2008) for PA intervention on word-recognition from pre-test in the original Gillon study to 11-month follow-up. Children who received the other interventions also made large gains from pre-test to follow-up, showing an effect size of 1.52 (Cirrin & Gillam, 2008).

Children as young as three years old with speech impairment who do not have associated language deficits have also been shown to benefit from PA intervention (Gillon, 2005). In a three-year longitudinal study, Gillon examined the effects of PA intervention on phonological skills of 12 children ages 3 years to 3 years, 11 months with speech impairment as compared to a control group of 19 children in the same age range without speech delays. PA intervention focused on improving speech intelligibility, teaching PA at the phoneme level, and letter-name and letter-sound knowledge. Children in the treatment group received an average of 25.5 sessions of PA therapy twice a week, with one individual 45-minute session and one group 45-minute session with two or three other children in the study. Using tasks of PA assessment from Bradley and Bryant's (1983) research with typically developing 3- and 4-year-old children, the authors assessed children at three 7- to 8-month intervals from age 3 to 5. In order to control for performance variability typical of young children, some tasks were administered twice a few days apart, and the higher score was included in the analysis. Children's scores on rhyme and letter recognition improved at a similar rate in both groups. An analysis of variance with gain scores on a phoneme matching task showed that children in the treatment group showed more growth from Time 1 to Time 2 than the group of typically developing children. No significant difference was noted in gain scores of the two groups from Time 2 to Time 3. Effect size of growth in phoneme matching scores from Time 1 to Time 2 was moderate (.5). Almost and Rosenbaum (1998) also investigated treatment for phonological (articulation) disorders focusing on inclusion of final consonants and production of fricatives (*s, sh, f*), velars (*f, g*), and consonant clusters. Treatment was administered twice weekly in 30-minute sessions for four months; one group received four months of treatment followed by four months without treatment, and a comparison group waited four months before receiving four months of treatment. The first group showed significant differences in scores from pretest to post-test on phonological measures including the Goldman Fristoe Test of Articulation (Goldman & Fristoe, 1969). The early treatment group made greater gains in conversational articulatory precision than the later treatment group. Significant differences in test scores only remained for a measure of percentage of consonants correct after completion of treatment. Segers and Verhoeven (2004) examined the effects of a computer-based PA intervention focusing on word, syllable, and phoneme analysis, rhyming, and phoneme and syllable synthesis on 24 kindergarten children with SLI in the Netherlands. The treatment groups received 3.5 hours of PA intervention over the course of five weeks through a computer program using either normal speech (N=12) or modified speech (N=12) with slower presentation and amplification of certain syllables, as in the Fast ForWord – Language program mentioned above. A control group of 12 kindergarten children with SLI played computer games with vocabulary lessons. Based on difference *z* scores, an analysis of variance showed that PA intervention was effective for the children in the normal speech PA treatment condition, with a small effect size of .29 as compared to the control group. This effect was not significant 18-weeks later, however.

Overall, it seems that clinicians can have a moderate degree of confidence when choosing phonological awareness interventions to remediate language deficits (Cirrin & Gillam, 2008). Activities designed to improve rhyming, sound identification, phoneme identification and manipulation, and phoneme-grapheme connections are effective with the language-disordered population. Results suggest that this type of intervention creates connections between oral and written language skills, which may ultimately decrease reading difficulties.

Language processing

Auditory-processing approaches to language intervention include computer-based programs of instruction such as Fast ForWord – Language software (FFW-L; Scientific Learning Corporation, 1998). Children with receptive language difficulties have been shown to struggle with language comprehension because of slower auditory processing speed; that is, they may not have completely processed one sound before the next sound is presented (Bishop & McArthur, 2005; Tallal & Piercy, 1974). Studies focusing on language processing in school-aged children have investigated the FFW-L program, a computer-based intervention with auditory tasks, specific language forms, and reading content that is intensive (≥ 100 minutes per day), extensive (5 days a week for 4 to 8 weeks), and adaptive, in that the difficulty of auditory stimuli is modified based on the person's performance (Agocs, Burns, De Ley, Miller, & Calhoun, 2006). The goal of intervention is to enhance auditory processing skills and improvement of working memory, syntax, and other skills important to oral language and reading. Merzenich and colleagues (1996) found that 22 children ages 5 years, 2 months to 10 years with mixed receptive-expressive language impairment showed significant improvement in auditory temporal processing skills according to the Tallal Repetition Test (Tallal, 1980) after playing the Circus Sequence and Phoneme Identification computer games from Fast FFW-L. These games focused on perceptual identification of tone sequences and phoneme recognition. Effect sizes were not calculated for this study. Tallal and colleagues (1996) also found significant improvement on measures of speech discrimination, language processing, and grammatical comprehension in the same group of children who played FFW-L games focusing on speech discrimination and on-line language comprehension. Two groups were compared: one that received training with modified speech (slowed down with amplified phonemes) and one that received training with normal speech. A repeated measures analysis of variance was used to compare performance from pretest to posttest on various measures of auditory processing. Both groups improved with training, but the modified speech group showed significantly greater improvements. No effect sizes were calculated for this study. Cohen and colleagues (2005) compared FFW-L exercises with the modified speech component to other computer programs without modified speech and a control group receiving traditional school language services. Based on a sample of 60 school-aged children, results showed that children in every group made similar language gains on the Clinical Evaluation of Language Fundamentals – 3 (Semel, Wiig, & Secord, 1995) at 9-week and 6-month follow-up periods, suggesting that the computer interventions plus school therapy were not more effective than school therapy alone. Gillam and colleagues (in press, as cited in Cirrin & Gillam, 2008) conducted a large randomized controlled trial with a sample of 216 school-aged children with SLI, comparing FFW-L, which uses modified speech stimuli, to a computer-assisted language intervention (CALI) without a modified speech component, individualized language intervention (ILI), and academic enrichment (AE), which involved non-language-related cognitive and academic skills. Similar outcomes were found on language and auditory processing measures for all groups. Moderate effects were found for language improvement in all four groups after six weeks of intervention, and large effects were found for language improvement at 6-month follow-up of all four groups. Cirrin and Gillam (2008) suggested that the intensive nature of the language interventions, which required sustained attention and immediate responses to items, in addition to social interaction with peers and invested adults contributed to the positive outcomes seen in all groups. Overall, FFW-L does not appear to contribute more to positive language outcomes than traditional speech-language services, academic enrichment, or other computer training programs. Therefore, it is likely not a necessary component of an intervention program that already incorporates another method of language intervention.

Table 1. Review of Empirically Studied Speech-Language Interventions (Adapted from Cirrin & Gillam, 2008; Law et al., 2004; McCauley & Fey, 2006)

Intervention	Reference	Participants	Description	Outcome Measures	Results	Critique
			Syntax & Morphology Interventions			
Computerized grammatical comprehension training	Bishop, Adams, & Rosen (2006)	-N=36 children with receptive LI ages 8 to 13	-Nonrandomized comparison -2 experimental groups were trained on computer tasks for 20 daily 15-min. sessions in a slow-speech or modified speech condition used in FFW-L; control group received no training	-TROG-2 (Bishop, 1989) -ERRNI (Bishop, 2004) -ERRNI Mean Length of Utterance (MLU)	No significant differences between any groups	No mention of whether children continued to receive regular school language instruction, which could be a confounding variable
Visualizing and Verbalizing (V&V) to improve receptive syntax	Dixon, Joffe, & Bench (2001)	-N=8 children ages 9 to 15:1	-2 children received traditional therapy & 2 received V&V therapy in 30-minute sessions over 10 weeks - 4 received traditional therapy for 5 weeks & V&V for 5 weeks	Analytic Reading Inventory (Woods & Moe, 1995)	-Performance improved significantly with both interventions -No significant contribution of Visualizing & Verbalizing beyond traditional therapy	-No effect size reported -Small sample size – limited statistical power -Only 4 children in purely traditional or experimental interventions included in analysis -No breakdown of group by gender available -Possible variation in SES of participants -No control group to rule out confounding variables

Intervention	Reference	Participants	Description	Outcome Measures	Results	Critique
Syntax & Morphology Interventions (continued)						
Modeling vs. imitation of invented morphemes	Connell & Stone (1992)	-N=32 children with SLI, 24 age-matched and 20 language-matched controls ages 5 to 6:11	-Split-plot factorial -LI & control children were assigned to a modeling only or elicited imitation condition of invented morphemes; conditions were counterbalanced	Pre- and posttest production & comprehension probes of invented morphemes developed by authors	-Control children performed well in both experimental conditions -LI children showed minimal use of the morphemes with modeling alone and significantly > use with imitation -Both interventions equally effective for LI children's comprehension of morphemes	Results confounded by large effects of order of treatment: the 1st intervention administered overrode effects of 2nd intervention
Modeling vs. modeling plus evoked production to improve syntax	Weismer & Murray-Branch (1989)	-N=4 children with SLI ages 5:5 to 6:11	-Single-subject (alternating treatment with baseline) -Each child was given individualized grammatical goals to achieve and received treatments in alternating order	-No. of attempts & correct productions of target grammatical form	-3 children with expressive delay showed improvement with both treatments, 1 child with mixed receptive-expressive delay showed no significant improvement -No significant difference between treatment types	Alternating treatment design may have caused interference from one treatment with the other, suggesting that different results may be attained if treatments were administered independently

Intervention	Reference	Participants	Description	Outcome Measures	Results	Critique
Syntax & Morphology Interventions (continued)						
2nd 5-month Phase of Parent- vs. Clinician-administered Grammar Facilitation	Fey, Cleave, & Long (1997)	-N=28 children ages 3:8 to 5:10 carried from above study	9 children received 10 total mos. of clinician intervention, 9 received 10 total mos. of parent intervention, & 10 children made up dismissal group who received essentially no additional intervention after 1st 5 mos. of study	DSS analysis	-Clinician- and parent-group DSS higher after Phase 2 of intervention; size of gain not sig. different from size of Phase 1 gain for each group -Clinician group's size of Phase 2 gain in mean main verb score of DSS sig. < size of Phase 1 gain -Dismissal group: no reliable change from post-Phase 1 to post-Phase 2	Treatment fidelity not consistently monitored in parent intervention, which may suggest poorer adherence to treatment techniques and thus poorer consistency of gains over time
Semantics & Vocabulary Interventions						
Communicative Reading Strategies (CRS) vs. traditional reading decoding	Crowe (2003)	-N=12 children ages 8 to 11 with LLD and scoring <50th percentile on national normed reading achievement test	-Nonrandomized controlled trial -4 children received 6 wks of CRS, 4 children received traditional decoding, and 4 children in control group -All children continued to get school-administered language services -Control children offered 6 wks of intervention following study	-Gray Oral Reading Test – Revised (GORT-4; Wiederholt & Bryant, 1986) -Comprehensive Receptive and Expressive Vocabulary Test (CREVT; Wallace & Hammill, 1994)	-Posttest reading comprehension sig. > for CRS group -No sig. differences on vocab measures between groups -All groups worse on receptive and general vocab posttests (an alternate form of the CREVT) -CRS group improved on alternate form of expressive vocab posttest (moderate ES of .5); other groups did worse	-Small sample size = limited statistical power -Decreases in receptive vocabulary skills suggest posttest alternate form may have been more difficult

Intervention	Reference	Participants	Description	Outcome Measures	Results	Critique
			Semantics & Vocabulary Interventions (continued)			
Three service delivery models of vocabulary intervention	Throneburg et al. (2000)	-N=32 children grades K to 3rd who were eligible for speech-language services based on lang/artic scores	-Nonrandomized controlled trial -Children received 1) group hands-on vocab lessons with 60 embedded vocab words collaboratively planned/taught by an SLP & teacher, 2) separately planned/taught classroom-based lessons, or 3) one 50-min. vocab session outside of class given by SLP using same techniques	Pretest/posttest performance on 20-item vocabulary test measuring ability to define words, use words in sentences, and recognize definition of word in a multiple-choice question	-Large ES's noted from pretest to posttest for every condition: Collaborative condition = 2.5, Classroom-based condition = 3.5, & Traditional pull out of class = 1.2 -Students in collaborative condition obtained highest posttest scores	No non-treatment control group used, so there is a possibility that effects were due to confounding variables, e.g., word learning outside of interventions
Three verbal presentation styles of novel words	Weismer & Hesketh (1993)	N=16 children (8 with SLI & 8 normal language controls) ages 5:1 to 6:7	-Nonrandomized comparison -9 invented words were presented using variation in rate (slow/normal/fast), stress (with/without emphasis on certain syllables), and visual aids (verbal presentation with/without iconic gesture)	No. of novel words produced and comprehended	-Slower speaking rate and additional use of gestures improved learning of novel words for both groups - ES's for SLI group: variation of rate for comprehension & production = 1.1 (large); variation of stress for comprehension = .12 (no effect)/production = .74 (moderate); variation of gesture for comprehension = .57 (moderate)/production = .33 (small)	Small sample size = limited statistical power

Intervention	Reference	Participants	Description	Outcome Measures	Results	Critique
			Semantic & Vocabulary Interventions (continued)			
Intensive vs. weekly speech therapy	Barratt, Littlejohns, & Thompson (1992)	- N=39: 27 males & 12 females ages 3:1 to 3:7 -21 children received weekly therapy, 18 received intensive therapy	-Randomized comparison -Clinician-administered interactive language therapy focusing on exp/rec skills -Intervention given 40 min. per week over 6 mos. or in 2 intensive therapy blocks (40 min., 4 days per week for 3 weeks in a 3 mo.-block)	Reynell Expressive & Receptive Scales (Reynell, 1977)	-Both groups showed improvement in comprehension score, but no sig. difference between groups -Intensive group showed sig. improvement in expression score, while weekly group improved, but not significantly (p=.18)	-No control group for comparison -Effect sizes not calculated
Semantic vs. phonological word-finding	Wing (1990)	-N=10 children ages 5:11 to 7:1 with severe LI	-Nonrandomized comparison -Half of children received semantic treatment, half received phonological treatment in 30 25-min. group therapy sessions over 2.5 mos.	Test of Word Finding (German, 1986)	-Phonological treatment produced sig. improvement, semantic did not -Cirrin & Gillam (2008) computed mod. large effect sizes for both groups: semantic (d=.6), phonological (d=.7)	-No random assignment -No control group for comparison

Intervention	Reference	Participants	Description	Outcome Measures	Results	Critique
			Phonological Interventions			
Rhyming & phoneme awareness intervention	Van Kleeck et al. (1998)	N=24 preschool aged children with LI	-Nonrandomized comparison with historical control -Group instruction 2x/wk for 2 semesters - fall incl. rhyming words; spring incl. blending sounds -Control group- children who attended same classes prev. year	Measures of rhyming and phoneme awareness	-Experimental group obtained scores 1.5 standard deviations above the control children on standardized measures of phoneme awareness, large ES	-No random assignment
Phonological awareness intervention for preschool children with LI	Gillon (2000)	-N=91 children ages 5:6 to 7:6 with speech and LI	-Matched treatment and control groups compared with normally developing children - Comparison of pull-out traditional therapy services, classroom instruction with PA intervention, or monthly consultation with parents	-Lindamood Auditory Conceptualization Test (LACT; Lindamood & Lindamood, 1979) -Queensland University Inventory of Literacy (Dodd et al., 1996)	-PA group made sig. > improvement than other intervention groups -No sig. diff. b/w PA group & typically developing children's scores on the LACT post-intervention. -PA group made sig.> improvement than other groups on reading measures. -PA cf. traditional: large ES on phoneme awareness (2.58), mod. ES on reading comprehension. -PA cf. consultation: large ES on phoneme awareness (1.77), mod. ES on rhyming (.67)	-No random assignment

Intervention	Reference	Participants	Description	Outcome Measures	Results	Critique
			Phonological Interventions (continued)			
11-month follow-up to above study	Gillon (2002)	-N=60 children from above sample	-Matched treatment and control groups compared with normally developing children -Comparison of 20 PA children, 20 who received either traditional or consultation treatment, & 20 typically developing children	Burt Word Reading Test (Gilmore et al., 1981)	-Improvements maintained for PA children in phoneme awareness and word-recognition -Children in PA group had reading at or > level expected for their age & maintenance of non-word spelling -Large ES (2.42) calculated by Cirrin and Gillam (2008) for PA intervention on word-recog. from pre-test in original Gillon study to 11-month f/u -Large ES of 1.52 shown for other intervention groups from pre-test to f/u	-No random assignment
Phonological intervention for speech impairment	Gillon (2005)	-N=12 children age 3:0 to 3:11 with speech impairment and 19 children without speech delays	-PA intervention to improve speech intelligibility, understanding at the phoneme level, and letter-name and letter-sound knowledge	PA assessment tasks from Bradley and Bryant (1983)	-Scores on rhyme and letter recog. improved at a similar rate in both groups -PA group showed > growth from Time 1 to Time 2 than typical children -No sig. diff. in gain scores of 2 groups from Time 2 to Time 3 -Mod. ES (.5) of growth in phoneme matching scores from Time 1 to Time 2	-No critique

Intervention	Reference	Participants	Description	Outcome Measures	Results	Critique
Phonological Interventions (continued)						
Speech intervention for phonological (articulation) disorders	Almost & Rosenbaum (1998)	-N=26 children ages 2:9 to 5:1 with phonological disorder	-RCT -Speech therapy focusing on accurate articulation & phonological components -Group 1 received 4 mos. treatment and 4 mos, no treatment, Group 2 waited 4 mos. before receiving 4 mos, treatment	Goldman Fristoe Test of Articulation (Goldman & Fristoe, 1969)	-Group 1: sig. differences in scores from pretest to post-test on phonological measures including Goldman-Fristoe -Group 1 made greater gains in conversational articulatory precision than Group 2 treatment group -Sig. differences in test scores only remained for a measure of percentage of consonants correct at posttest	-No critique
Computer-based phonological awareness vs. vocabulary training	Segers & Verhoeven, (2004)	N=24 kindergarten children with SLI	-Computer-based PA intervention using normal speech vs. modified speech, control group played vocabulary computer games	Researcher-developed tasks measuring word, syllabic, rhyme, and phonemic awareness	-Based on difference *z* scores, ANOVA showed that PA intervention was effective for the children in the normal speech treatment condition, with a small ES of .29 as compared to the control group. -Effect was not significant 18-weeks later	No critique

Intervention	Reference	Participants	Description	Outcome Measures	Results	Critique
			Language Processing Interventions			
Circus Sequence & Phoneme Identification games of FFW-L	Merzenich et al. (1996)	N=22 children age 5:2 to 10:0 with mixed rec/exp LI	-Games focusing on perceptual identification of tone sequences and phoneme recognition -8 to 16 hours of training during a 20-day period	Tallal Repetition Test (Tallal, 1980)	Children improved significantly in auditory temporal processing ability with training	No effect sizes calculated
Fast ForWord Language games with and without modified speech stimuli	Tallal et al. (1996)	N=22 children age 5:2 to 10:0 with mixed rec/exp LI	- Listening exercises and games designed to improve speech discrimination and on-line language comprehension; modified vs. normal speech conditions -Extensive daily training for 4 weeks	-The Token Test for Children (DiSimoni, 1978) - Goldman-Fristoe-Woodcock Diagnostic Auditory Discrimination Test (Goldman, Fristoe, & Woodcock, 1974) - Curtiss and Yamada Comprehensive Language Evaluation-Receptive (Curtiss & Yamada, Unpublished work) -Computerized Version of the Tallal Repetition Test (Tallal & Miller, 1994) - Goldman-Fristoe Test of Articulation (Goldman & Fristoe, 1986)	-Significant improvements in speech discrimination and language comprehension abilities were demonstrated in both groups -Modified speech group made greater gains	No effect sizes calculated

Intervention	Reference	Participants	Description	Outcome Measures	Results	Critique
			Language Processing Interventions (continued)			
RCT of FFW-L modified speech vs. other computer programs and school services	Cohen et al. (2005)	N=60 children age 6 to 10 with mixed rec/exp LI	-RCT with 3 intervention groups, incl. treatment as usual (TAU) -FFW-L group: tasks incl. discrimination of tones, phonemes, syllables, & words; memory for commands; & grammar w/modified speech -Other treatment group: tasks incl. listening, phonological awareness, reading, writing, vocabulary, syntax, no modified speech -Control group: school therapy services only (TAU)	Clinical Evaluation of Language Fundamentals-3rd Edition (Semel et al., 1995)	-Similar gains between groups on the CELF-3 at 9-wk and 6-mo Follow-up -Computer intervention plus school therapy was not more effective than school therapy alone -ESs: At 9 wks, d=-.09 for FFW vs. Control (no effect), d=.27 for FFW vs. Other (small); At 6 mos, d = .05 for FFW vs. Control (no effect), d = -.27 for FFW vs. Other (small)	No critique
RCT of FFW-L vs. other computer programs and academic enrichment	Gillam et al. (in press, as cited in Cirrin & Gillam, 2008)	N=216 children age 6 to 9 with SLI	-RCT with 4 intervention groups, incl. TAU -FFW-L group with modified speech -CALI group without modified speech -ILI (TAU) -AE -Interventions implemented for 6 weeks	Language and auditory processing measures (not specified)	-Similar gains in all groups -Moderate effects (d = .56 to .79) for language gains after 6 weeks of intervention, large effects (d = .93 to 1.34) for language gains at 6-month follow-up	No critique

Summary of Language Interventions

Empirically studied approaches to intervention with LI children vary widely. The current review organized a sample of interventions by the area of language to be remediated: grammar, vocabulary, phonological awareness, or processing skills. Overall, phonological awareness intervention appears to have the best evidence base, reporting moderate to large effect sizes of intervention. Several intervention packages such as the Hanen Program for Parents (Girolametto & Weitzman, 2006), Language is Key (Cole, Maddox, & Lim, 2006), and Enhanced Milieu Teaching (Hancock & Kaiser, 2006) purport to address deficits in various areas of language and are designed for children with specific language impairment as well as children with developmental delays and mental retardation. These interventions incorporate several techniques and service delivery models outlined above, including focused stimulation approaches, conversational recast, modeling and imitation, and training parents to facilitate language learning in the home.

STUDY INTERVENTIONS

The interventions reviewed next are those used in the current study to remediate language deficits. The Montessori Method has been empirically studied, but the more specialized Montessori Applied to Children At-Risk (MACAR) intervention has not. The DuBard Association Method has been investigated with pilot studies that are methodologically flawed and have not been statistically analyzed. Sensory integration therapy, which was a component of the occupational therapy used in the current study, has been shown not to produce significant effects on language or academics.

Montessori Applied to Children At Risk

Dr. Maria Montessori, an Italian physician in the late 1800s, developed the Montessori Method of teaching, which employs multisensory techniques that allow the child to choose which activities to master and learn at his or her own pace. Dr. Montessori observed that children go through several “sensitive periods” of learning from birth to age 5, in which they gravitate toward certain tasks (Montessori, 1988; Pickering, 1988; Seldin, 2006). The teacher may be known as a “director”, “directress”, or “guide”, emphasizing that the child intuitively knows which activities are appropriate for his or her developmental level and may use the teacher as a resource if needed (Seldin, 2006). One defining feature of the Montessori Method is the Prepared Environment, which involves a classroom tailored to the needs of the child that increases the potential for learning. The Prepared Environment consists of the correct scaling of furniture to fit the child (e.g., lowered counters and shelves and smaller chairs and desks), as well as trays of objects that can be easily manipulated by small hands to learn specific concepts (Pickering, 1988). Concepts to be learned are presented by the teacher, beginning concretely and progressing to abstract ideas (Pickering, 1988). The teacher may present several activities on trays to a small group of students, allowing them to choose what they would like to practice, and modifying the difficulty level accordingly. Another defining feature of Montessori is the multisensory nature of the curriculum, which teaches concepts in nine areas: Practical Life, Sensorial, Mathematics, Language, Social Studies, Physical Sciences, Art, Music, and Perceptual Skills. The Practical Life curriculum, which is specific to the Montessori Method, encourages the child's independence in taking care of herself and her environment, through activities such as using the restroom by herself, cutting food to prepare snacks, pouring liquids, and cleaning up the room (Pickering, 2004a). Children learn about each subject area through visual, auditory, tactile-kinesthetic, gustatory, and olfactory discrimination tasks (Pickering, 1988). They are taught to distinguish colors, shapes, and sizes through the visual domain; tone values and volume through the auditory domain; textures, weights, and temperatures through the tactile domain; and a variety of tastes and smells through activities involving food. Children are taught to attend to detail and develop perceptual acuity through presentation of stimuli with barely noticeable differences (Pickering, 1988).

The Montessori Method also presents language lessons using multisensory techniques. Oral language is always taught before progressing to written language, as it serves as a building block for reading and writing (Pickering, 1988). Children learn to label every object used in a lesson and describe its features and functions, and precise articulation is necessary to demonstrate mastery of oral language. The child is introduced to written language through the use of metal frames that have detachable shapes to trace and shade, which helps the child build the fine motor skills necessary for writing. The child's other activities in the classroom – painting, cleaning tables, and cutting fruit and vegetables – also help the child's fine motor skills to develop. The child may progress further toward reading and writing by tracing sandpaper letters; this task is multisensory because the child must observe the shape of the letter (visual), say the letter and sound it produces (auditory), and feel the texture of the letter while tracing its shape (tactile-kinesthetic) (Pickering, 1988). Reading progresses further through the use of the moveable alphabet, which is a set of letters that the child can manipulate to produce words, and picture cards that depict the words (Pickering, 1988).

Pickering (1988, 2004a, 2004b) asserts that children who are at-risk for a learning disability (Brutten, Richardson, & Mangel, 1973; Critchley, 1984; Shedd, 1967) can benefit from the Montessori Method of teaching. These children experience difficulty in the areas of attention, organization, gross and fine motor skills, and perception, and they often have accompanying weaknesses in language acquisition, reading, writing, and abstract mathematical concepts (Pickering, 2004a). Pickering claims that At Risk children benefit from the 1:1 ratio of teacher to student, which allows the teacher to respond more readily to the child's challenges than the traditional educational environment. The structure of the Montessori classroom, with various activity stations and small groups engaging quietly in activities, seems to provide an orderly system in which the child with attention difficulties can operate. The teacher models appropriate behavior by avoiding yelling across the room and instead approaches each child individually, speaking in a controlled and calm voice. The teacher offers the At Risk child choices and direction to encourage her to be productive. When the child chooses a particular activity to work on alone, she is given a mat to work on; children are taught to respect each other's individual work spaces, encouraging structure and responsibility for behavior. Children must set up activities, perform the activities, and replace them on the shelves when finished, encouraging order and organization. After work periods are over, children are encouraged to play the Silence Game by sitting on a taped line on the floor and seeing how long they can remain still and quiet. Many activities promote motor skills development of the At Risk child, including carrying trays to work mats, cutting food, and doing class exercises to music, such as marching, hopping, and skipping to the rhythm. Perceptual skills of the At Risk child are thought to develop through multisensory activities that promote integration of sensory input (Pickering, 2004a). Academic skills of the At Risk child in areas such as language and mathematics are also taught in a multisensory fashion, utilizing manipulative objects such as sandpaper letters, stencils, and number rods, as well as verbalization to describe the activities being executed. These techniques are designed to encourage the integration of visual, auditory, tactile, and kinesthetic input, which is believed to be essential to promote learning in the child who is At Risk for a learning disability (Pickering, 2004b).

Despite having widespread influence on educational practice and an established reputation as a holistic, integrative approach to teaching, very limited empirical research has been conducted on the Montessori teaching method. As Tim Seldin, President of the Montessori Foundation, puts it, “the research that has been done to date on Montessori’s effectiveness in the American educational context is far more limited than it should have been after its more than 90-year history in this country” (Seldin, 2002/03, p. 5). Additionally, many Montessori research articles are virtually inaccessible to the general public, given the fact that they are published in “fugitive” documents, such as journals of the Association Montessori Internationale or the American Montessori Society, which are not indexed in the Educational Resources Information Center (ERIC) or other research databases (Chattin-McNichols, as cited in Seldin, 2002/03, p. 5). Despite the minimal number of studies available and limited access to these studies, Lillard and Else-Quest (2006) estimate that 5000 schools in the United States, including 300 public schools and some high schools, use the Montessori teaching method. Frequently-cited studies by Duax (1995) and Dawson (1987) assert that Montessori education leads to superior results in academic achievement, but Lopata, Wallace, and Finn (2005) explain that evidence for this claim is limited by methodological flaws. Duax's study examined performance on annual standardized achievement tests of 36 children enrolled in a private Montessori school from second through eighth grades (1995). Results showed that these children outperformed national norms in the areas of total reading and total math, leading the author to conclude that "Montessori schools produce greater than expected academic achievement in students" (Duax, 1995). However, due to lack of a control group, no mention of appropriate statistical tests of significance, and the fact that the sample's second-grade testing results were already above average, Duax's (1995) claims do not indicate that achievement results are due to Montessori education specifically. Dawson's (1987) study looked at mean grade equivalence scores on the Iowa Test of Basic Skills and the Metropolitan Achievement Test (Hogan, Farr, Prescott, & Balow, 1986) in 88 first- through fifth-grade minority students in a Montessori magnet program. Test scores were significantly higher than national test norms or Houston Independent School District mean test scores, which were based on scores of first- through fourth-graders attending traditional schools in the district matched for ethnicity (Dawson, 1987). The author did not provide data on whether the groups differed in any way before beginning

Montessori education, demographic characteristics of the sample were not statistically controlled, and the factor of parental selection for type of school could not be ruled out based on this study (Dawson, 1987). Miller and Bizzell (1984) conducted research on long-term achievement outcome of Montessori preschool education versus traditional preschool education in ninth- and tenth-grade low-income African-American students. Based on the Comprehensive Tests of Basic Skills (McGraw-Hill, 1974), no significant differences were found between education programs. Higher scores were found for males who attended Montessori preschool programs, but since achievement scores were shown to be commensurate with IQ, achievement could not be attributed to the education programs themselves.

Lopata and colleagues (2005) studied the academic achievement outcomes of students attending New York urban schools with either a Traditional Non-Magnet (TNM), Structured Magnet (SM), Open Magnet (OM), or Montessori education program. A sample of 291 fourth-graders and 252 eighth-graders were chosen from one each of the four different types of public schools, which were matched on gender, ethnicity, and socioeconomic status. The types of schools chosen for comparison with Montessori were based on key differences in the educational environment. For example, Montessori's approach focuses on learning rather than work products such as tests and grades, and no negative consequences are used to discipline behavior. In the SM and TNM schools, however, classroom assignments, grades, and strict discipline are employed. The OM school utilizes a flexible approach to scheduling and classroom size, and school meetings are used to resolve conflicts. The researchers used math and language arts test scores from the New York State Mathematics and English/Language Arts exams and the Math and Language Arts portions of the TerraNova (McGraw-Hill, 2002) to determine whether Montessori education programs produce significantly better academic achievement scores than other traditional education programs (Lopata et al., 2005). The authors found that the fourth-grade Montessori students scored significantly better in mathematics than the OM students by .60 standard deviations, but they performed significantly worse than the TNM students by .37 standard deviations (Lopata et al., 2005). No significant differences were found between Montessori and other types of education on measures of language arts achievement for fourth-graders. In contrast, eighth-graders in the Montessori school scored significantly lower than SM students by .77 standard deviations and TNM students by .59 standard deviations in language arts achievement. No significant differences between Montessori and other programs were found for mathematics achievement. The authors did not find conclusive evidence for Montessori students' academic superiority over students educated in other types of programs. They also acknowledge significant limitations in their research, including the fact that only one school of each type of education system was chosen, meaning that differences could be attributable to other characteristics of the school itself rather than the type of education structure. Fidelity to the educational program structure was not monitored, and no information on the amount of time the students had been in their respective programs was available to the researchers (Lopata et al., 2005).

A more recent study conducted by Lillard and Else-Quest (2006) compared academic and social competence of Montessori-educated children to those educated in other types of systems. To control for effects of parental selection, the researchers chose their sample from children whose parents entered the Montessori school lottery in Milwaukee, WI, in 1997 and 2003. Children who were randomly chosen for the Montessori school were the experimental group, and children who were not chosen formed the control group and were dispersed among 27 public inner-city schools and 12 suburban public, private/voucher, or charter schools. Average income levels among parents were similar, but ethnicity was not surveyed because parental income has been shown to contribute more to child performance than ethnicity (Duncan et al., 1998, as cited in Lillard & Else-Quest, 2006). A total of 53 control students and 59 Montessori students participated in the study. A five-year-old group consisted of 25 controls and 30 Montessori children, and a 12-year-old group consisted of 28 controls and 29 Montessori children. While gender ratios were not quite balanced (Montessori: 50% girls in five-year-olds, 59% girls in 12-year-olds; Control: 60% boys in five-year-olds, 64% boys in 12-year-olds), the authors explain that gender was not shown to contribute significantly to any of their findings.

Based on the Woodcock-Johnson III Test Battery, Montessori five-year-olds showed significantly better performance than controls on three subtests: Letter-Word Identification (single word reading), Word Attack (phonological decoding of nonsense words), and Applied Problems (math skills) (Lillard & Else-Quest, 2006). No differences were found on the Picture Vocabulary (expressive vocabulary) subtest or the executive function subtests of Spatial Reasoning and Concept Formation. However, on a card sort test of executive functioning, Montessori five-year-olds performed better than controls.

On the Social Problem-Solving Test-Revised, Montessori five-year-olds were significantly more likely to find solutions to problems that involved higher-order reasoning and fairness than children in traditional educational settings. These children were also more likely to engage in shared peer play on the playground and less likely to be involved in rough play, but it remains unclear whether observers were blind to the treatment condition of the children. In the 12-year-old group, Montessori students were judged to write more creative essays with more advanced sentence structure than their traditionally educated peers; however, the criteria for rating these essays was not provided. Twelve-year-old Montessori students did not outperform traditional education students on tests of achievement. Montessori 12-year-olds were more likely to choose the positive assertive response on a multiple choice measure of social skills, and they expressed a greater sense of school community on a measure of attitudes about school (Lillard & Else-Quest, 2006). While this study demonstrated some isolated advantages of Montessori in reading and math skills for five-year-olds, creativity and complexity in essay-writing for 12-year-olds, and social skills with peers, it is difficult to say whether these findings would be replicated at other Montessori schools, since the study was based on one Montessori school campus. One criticism alleged that since the children not chosen for the Montessori school were dispersed to other traditional schools, they attended school with peers whose parents did not enter the original lottery, creating an indirect influence of parental selection for entry in the lottery (Kavanagh, 2007). The authors responded that while information on which specific traditional schools admit by lottery was not readily available, the control children who attended schools that were known as “more likely to” (Kavanagh, 2007, p. 597) admit by lottery still did not perform as well as Montessori children.

Another study of Milwaukee public schools measured the long-term effects of early Montessori education. Gartner and Kerzner-Lipsky (2003) chose 201 students who graduated from traditional public high schools between 1997 and 2001 as their sample. The experimental group attended Montessori schools from age three or four until fifth grade, and the control group, which was matched for age, gender, and ethnicity, attended traditional education programs in their elementary years. The groups were also matched based on which high school they attended. Based on the Wisconsin Knowledge and Concepts Evaluation and the ACT, the Montessori-educated students outperformed the traditionally-educated students on measures of mathematics and science. The groups did not differ significantly on English/Social Studies scores or GPA. The researchers acknowledge, however, that since the groups were not chosen randomly, other factors such as parental selection for Montessori schooling could have influenced the results.

Many studies on Montessori’s effectiveness are plagued by methodological flaws, namely comparing only one Montessori school to other education systems or failure to adequately control for the confounding variable of parental selection for Montessori education. While some evidence suggests that Montessori education may have isolated benefits over traditional or magnet education programs, these benefits are not consistently seen in the same areas. Lillard and Else-Quest’s (2006) study suggests that achievement of five-year-old Montessori students is higher than that of five-year-olds in other types of education programs, but these results are not maintained at age 12. Social skills of Montessori students were found to be superior in both age groups, but the factor of parental selection could have played a role in this finding, and it is unclear whether observations of these social skills were made objectively. Gartner and Kerzner-Lipsky's (2003) study appears to be the most well-controlled study, but they acknowledge that they are not able to completely rule out the effect of parental and family factors. Overall, Montessori has been consistently shown to produce similar achievement test results to other types of education. However, the above-mentioned studies did not investigate the effects of Montessori techniques on children with specific language delays or learning differences, and they limited their dependent measures to tests of purely academic ability or social skills. It is possible that the current population of students may demonstrate a different rate of improvement in areas other than academics, such as receptive and expressive verbal language ability, articulation, and oromotor skills.

The Montessori Method Applied to Children At-Risk (MACAR), which involves more one-on-one teacher-child interaction and special focus on attention and language skills, has not been empirically studied and analyzed with appropriate statistical procedures. The basis for using MACAR as an intervention in the current study is based on theoretical knowledge and a wealth of teaching, remediation, and clinical experience with a select group of learning-different students. Some *clinical* studies involving children at-risk for learning disorders have been conducted on MACAR (Jones, 1971a, 1971b; Pickering, 1990). One study conducted from 1967-1970 involved 101 children ages 7 to 18 receiving remediation with MACAR and the Alphabetic-Phonic Structural Linguistic

(APSL) Approach in a self-contained program and 500 children receiving the same remediation approaches one hour a day. The entire sample was reported to show improvement in mean language scores over time (Pickering, 1990), but these changes were not statistically analyzed using stringent procedures such as *t* tests or the reliable change index, and no apparent attempts were made to compare these groups. Another study conducted during the 1970-1971 school year involved 46 children ages 5 and 6 in either a regular kindergarten classroom or the Early Childhood Education (ECE) Program, which consisted of a Montessori education structure applied to children at-risk for learning difficulties (Jones, 1971a). The ECE group happened to consist of a majority of children at high-risk for learning disabilities, as compared to the regular kindergarten group, which consisted of a majority of children within normal limits in terms of visual and auditory perception and language skills. Due to baseline discrepancies between groups in mean IQ and other skill areas, comparisons could not feasibly be made over time. The ECE group showed improvement in language areas including verbal cognitive ability, auditory processing, letter/sound knowledge, and language encoding, but a much smaller proportion of the sample showed change in vocabulary. Once again, however, these results were not statistically analyzed for significance. From 1971 to 1976, further study was conducted on the Reading Study Foundation programs, which involved 154 children ages 3 to 5 in the low average to gifted range of cognitive ability in various schools (Pickering, 1990). Children received the MACAR instructional method and APSL and showed gains in verbal cognitive ability, auditory processing, pre-reading skills, letter-writing, and language encoding. MACAR with APSL was also studied from 1976 to 1984 in São Paulo, Brazil, with 154 preschool students in a private American school. Students were shown to make small but steady gains in mental and perceptual ability from year to year (Pickering, 1990). Unfortunately, none of these clinical studies on MACAR with the APSL language approach were statistically analyzed to determine either statistically significant or reliable change, nor were means between groups compared statistically. Therefore, although increases in mean scores were observed, it cannot be determined whether the *magnitude* of these improvements represented change beyond that due to chance. Due to the dearth of appropriately analyzed empirical research on MACAR, this study serves as a pioneering effort to increase our understanding of the possible benefits of this intervention to children with LI who are prone to learning difficulties.

Table 2. Empirical Evidence for the Traditional Montessori Method with Non-Learning Disabled Children

Reference	Participants	Description	Outcome Measures	Results	Critique
Duax (1995)	N=36 2nd – 8th graders	Children attended a private Montessori school	Annual standardized achievement tests	Children performed better than national norms in total reading & total math	No control group, no tests of statistical significance, 2nd grade test results already > average, so results cannot be attributed to Montessori education
Dawson (1987)	N=88 1st – 5th grade minority students	Children attended a Montessori magnet program	-Iowa Test of Basic Skills -Metropolitan Achievement Test	Test scores significantly > national test norms or Houston ISD mean test scores	No data on group differences before beginning Montessori, demographics not controlled, parental selection cannot be ruled out
Miller & Bizzell (1984)	N=variable due to attrition, absenteeism, & missing data in school records; 9th & 10th grade low-income African American students	Follow-up of children who attended Montessori vs. traditional preschool programs	Comprehensive Test of Basic Skills	No significant differences between Montessori & traditional preschool education	Although higher achievement scores were observed in Montessori-educated males, they were commensurate with IQ, so results cannot be attributed to Montessori education
Lopata et al. (2005)	N=291 4th graders, 252 8th graders in 4 types of urban NY schools matched on demographics, including 1 Montessori school	Test scores obtained for children attending Montessori, Structured Magnet (SM), Open Magnet (OM), or Traditional Non-Magnet (TNM) education programs	-NY State Mathematics & English/Language Arts Exams -TerraNova Math & Language Arts portions	4th grade Math scores: Montessori > OM, Montessori < TNM scores; 4th grade Lang. Arts: no significant differences between Montessori & others 8th grade Math scores: no significant differences between Montessori & others; 8th grade Lang. Arts: Montessori < SM & TNM	Results not convincing for superiority of Montessori, but only one school from each type of education was used, education structure fidelity not measured, & amount of time child was in program not measured

Reference	Participants	Description	Outcome Measures	Results	Critique
Lillard & Else-Quest (2006)	N=112 children chosen based on Milwaukee Montessori school lottery; 5-year-olds- 25 controls & 30 Montessori, 12-year-olds-28 controls and 29 Montessori	Test scores obtained for children attending either Montessori school or other educational programs	-Woodcock-Johnson III Testing Battery -Card sort test of executive functioning -Social Problem-Solving Test-Revised -Playground observations -Written essays, school attitudes measure	5-year-olds: Montessori scored better than controls on card sort test, WJ-III single & nonsense word-reading & applied math problems, & some measures of social skills 12-year-olds: Montessori scored better on creative essays, more likely to give mature response on a measure of social skills, & more likely to report sense of school community than controls	Results based on only one Montessori campus vs. 27 control school campuses, meaning quality of Montessori school may be responsible rather than the education type itself; control children attended school with others whose parents may not have entered Montessori lottery, resulting in indirect effect of parental selection
Gartner & Kerzner-Lipsky (2003)	N=201 high school grads who attended Montessori or traditional programs from age 3 to 5th grade	Test scores obtained for students attending same traditional public high schools	-Wisconsin Knowledge & Concepts Evaluation -ACT	Montessori > controls in math & science, no differences in English, Social Studies, or GPA	Parental selection for Montessori elementary school cannot be ruled out

The DuBard Association Method

The DuBard Association Method is a "phonetic, systematic, structured, incremental and cumulative multisensory approach for teaching language and speech to children with multiple difficulties in language learning" (DuBard & Martin, 2000, p. 44). The teaching method was initially developed for hearing-impaired individuals by Mildred McGinnis in the 1920s and 1930s at the Central Institute for the Deaf in St. Louis, Missouri. Children are taught to precisely articulate phonemes of words before progressing to blend these sounds into whole words. The requirement of exact articulation is supported by the Motor Theory of Speech Perception, which asserts that precise production of sounds leads to improved perception of the sound and, thus, more efficient recall of language in later instances (DuBard & Martin, 2000). This phenomenon is also known as proprioceptive feedback, which, in the case of language learning, is information about how to produce speech sounds learned from making the actual movements with the lips, tongue, and jaw (DuBard & Martin, 2000). Information Theory, which promotes organization of smaller pieces of information into a framework, also serves as a basis for the DuBard Association Method (DuBard & Martin, 2000). The current intervention enhances the Montessori environment by incorporating these theoretical principles into the classroom. For example, incidental language charts are posted in the room to expose children to the visual representations of words and sounds, and the level of language in Montessori materials is reduced to the important points of the subject being taught (Stanislav, 2007). The DuBard Association Method, in conjunction with the Montessori

Method, purports to provide the framework that children with LI need to enhance language-learning and increase their potential for success.

DuBard and Martin (2000) emphasize that the specific tools utilized in the DuBard Association Method make it a unique instructional system. First, there is not a required packaged program of textbooks to buy. The DuBard School for Language Disorders at the University of Southern Mississippi recently developed pre-fabricated vocabulary and picture noun cards, as well as a booklet of drop drills for rehearsal with students, but the success of the program is mainly dependent upon the appropriate training and expertise of the instructor in the DuBard Association principles, not the materials utilized. Second, the use of Northampton Symbols, or Yale Chart Spellings, provides clear symbols for word sounds that are designed to facilitate the child's association of visual representations with their pronunciation. This visually distinct system of symbols aims to encourage development of both oral and written language skills. Third, children are required to write in cursive because, according to DuBard and Martin, this method allows them to produce the word fluidly and see the word as a whole. This way of writing apparently also reduces letter reversals commonly seen in children with learning disabilities. Fourth, instructors use color to draw the child's attention and differentiate phonemes in words. Fifth, children learn these phonemes in isolation as building blocks that are more manageable and can be more easily recalled than words. Sixth, instructors in the DuBard Association Method use "spaced form" (DuBard & Martin, 2000, p. 57) to modify their temporal rate of speech. This means that they say words more slowly and may prolong vowel sounds with slight pauses between phonemes of words (e.g., *b-oa-t,* *boa-t*). Next, precise articulation of phonemes is emphasized based on the Motor Theory of Speech Perception, as described above. Finally, the child receives his or her own speech-language book to document lessons learned and practice language exercises at home.

DuBard and Martin (2000) outline certain factors that are believed to influence development with the DuBard Association Method of instruction. The authors assert that children who are diagnosed early on with LI and have higher innate cognitive ability will tend to show more improvement than children with a later diagnosis and lower intellectual ability. An environment that provides language-impaired children with intensive instruction and allows them to interact with family and community who will reinforce these methods of instruction is thought to be ideal. The same instructors over a long period of time are also thought to provide the best scenario for children to feel comfortable enough with their language skills to adjust to different teachers later on. While the DuBard Association Method has been used with clinically observable improvement in hearing-impaired, aphasic, and language-impaired individuals for over 60 years, no empirical, randomized, controlled studies have been conducted on its efficacy (Sullivan & Perigoe, 2004). In this way, the DuBard Association Method is similar to other educational interventions in that it is based solely on anecdotal evidence and teaching experience (Sullivan & Perigoe, 2004). Based on pre- and post-intervention language test scores, the DuBard School in southern Mississippi has conducted some pilot studies showing improvement in children who have received years of intervention with the DuBard Association Method (DuBard School for Language Disorders, 1998;

Schraeder, 2008). The Arizona Articulation Proficiency Scale, Wechsler Individual Achievement Test-II, Test of Written Language, and Gray Oral Reading Test-4, were used to assess articulation, achievement, writing skills, and reading skills, respectively. Unfortunately, these test results have not been statistically analyzed for significance or reliable change, and no control group was utilized. The state of burgeoning research on the DuBard Association Method calls for more studies to be conducted on its efficacy with a language-impaired population. The current study aims to provide information about whether language skills improve with the use of the DuBard Association Method in conjunction with the MACAR educational instruction method.

Occupational Therapy with Incorporated Sensory Integration

Occupational therapy (OT) is an intervention designed to help individuals function independently and participate in activities of daily living. OT is used with people of all ages who have some type of impaired motor functioning, and in children, the intervention incorporates motor skill development, exercises, games, and play. Sensory integration (SI) refers to "the ability to organize, integrate, and use sensory information from the body and the environment" (Mauer, 1999, p. 383). Dr. A. Jean Ayres, an occupational therapist with advanced training in educational psychology and neuroscience, developed the theory of Sensory Integration (SI) based on her OT work with learning disabled children (Ayres, 1972). She observed sensory, motor, and perceptual deficits in these children and hypothesized that these impairments also affected their learning. Ayres's theory presumes that the brain functions holistically, wherein an interdependent relationship exists between the sensory systems, and brain regions must interact with each other in order for an individual to function (Ayres, 1972). It has been proposed that the interaction of the limbic, vestibular, tactile, and proprioceptive neural networks of the central nervous system are involved in SI dysfunction (Mauer, 1999). SI theorists emphasize the roles of these systems in learning oral and written language skills (Ayres, 1978; Magrun, Ottenbacher, McCue, & Keefe, 1981; Trott, Laurel, & Windeck, 1993). While SI deficits frequently occur in conjunction with language difficulties, the relationship between sensory integration and language acquisition is not clear. The goal of SI therapy is not to modify language directly, but rather to have positive effects on attention and behavior, which are components involved in the process of learning language (Cermak & Mitchell, 2006). An example of a challenge presented in therapy is an obstacle course in which the child must crawl across a platform to a ball pit, climb a rope ladder, and jump into a pile of pillows (Schaaf & Miller, 2005). This type of activity presumably promotes balance (vestibular control), awareness of the body in space (proprioception), and tactile stimulation.

In terms of research evidence to validate SI therapy, Schaaf and Miller (2005) note that while approximately half of the 80 studies reviewing the efficacy of sensory integrative OT have shown a significant positive impact, these studies present a host of methodological limitations. The lack of standardized, consistently replicable treatment is an issue because of the individualized nature of sensory integrative OT to fit the child's needs (Schaaf & Miller, 2005). Past research on sensory integrative OT has also lacked appropriate outcome measures and theoretical bases for study designs, comprising a "fishing expedition" approach to find any outcome that may be statistically significant and resulting in a diluted treatment effect (Schaaf & Miller, 2005, p. 146). Samples used in the research have also been highly heterogeneous, making it difficult to find any statistically significant differences between treatment groups (Schaaf & Miller, 2005). Given the fact that research on SI therapy lacks a standardized treatment protocol or duration, consistently used outcome measure, or consistent sample characteristics, there is no wonder that results have been so diverse (Schaaf & Miller, 2005).

A more definitive statement on the current status of SI research literature comes from Shaw (2008), who concludes that several quality empirical studies have been conducted that find SI not to be an effective treatment for children with learning or developmental disabilities. For example, Hoehn and Baumeister's (1994) and Griffer's (1999) qualitative reviews of SI research document several well-controlled studies that found no significant differences between the SI treatment group and an alternative treatment group on academic or language measures (Carte, Morrison, Sublett, Uemura, & Setrakian, 1984; Polatajko, Law, Miller, Schaffer, & Macnab, 1991; Wilson, Kaplan, Fellowes, Gruchy, & Faris, 1992). Two other studies comparing SI therapy to both perceptual-motor training and a no-treatment control group found no significant effects for SI therapy on academic, cognitive, or language variables (Humphries, Wright, McDougall, & Vertes, 1990; Humphries, Wright, Snider, & McDougall, 1992). Based on post hoc power analyses of the aforementioned studies and two additional studies of SI therapy, Polatajko and colleagues (1992) reported a minimal effect size, concluding that SI therapy was not more effective than a placebo for the academic performance of children with learning disabilities. A meta-analysis of 16 studies reported a non-significant average effect size of SI treatment on language measures when compared to alternative methods of treatment such as perceptual motor therapy and academic tutoring (Vargas & Camilli, 1999). Shaw and colleagues also conducted a meta-analysis of 41 studies, finding no significant effect sizes for language improvement, behavior, or sensorimotor functions and small effect sizes for motor skills and psychoeducational performance (Shaw, Powers, Abelkop, & Mullis, 2002). Twelve studies that controlled for maturation were also meta-analyzed, resulting in negligible effect sizes for motor skills and psychoeducational performance.

The Shelton School EI program included OT to remediate motor skills deficits commonly seen in children with language delays, and several SI therapy components were included in these OT sessions with the overall goal of improving the SI that is presumably necessary to learn language. However, based on strong evidence to refute its use as a language intervention strategy, SI therapy components presumably do not have any statistically significant effect on language skills and any changes in language observed in the current study will not be attributed to SI therapy.

Table 3. Empirical Evidence for Sensory Integration Therapy (Adapted from Griffer, 1999; Hoehn & Baumeister, 1994)

Reference	Participants	Description	Outcome Measures	Results
Carte et al. (1984)	N=87 school-aged children with learning disabilities (LD)	Comparison of SI treatment (2 to 3 45-min. sessions a week over 9 mos.) vs. no-treatment control, pre/post-testing.	Academic: WISC-R (Wechsler, 1974), WRAT (Jastak & Jastak, 1978)	SI treatment had no significant effect on dependent academic measures
Humphries et al. (1990)	N=30 school-aged children with LD & SI dysfunction	Comparison of SI treatment, perceptual-motor (PM) treatment (each 1 hour a week for 24 wks.), & no-treatment control, pre/post-testing	Cognitive: WISC-R Academic: WRAT Language: TOLD-P (Newcomer & Hammill, 1982)	SI treatment had no significant effect on dependent academic or language measures
Humphries et al. (1992)	N=103 school-aged children with LD & SI dysfunction	Comparison of SI treatment, PM treatment (each 3 1-hour sessions a week for 8 mos.), & no-treatment control, pre/post-testing	Cognitive & academic: WISC-R, WPPSI (Wechsler, 1967), WRAT, K-ABC (Kaufman & Kaufman, 1983), BSSI (Goodman & Hammill, 1975) Language: ITPA (Kirk, McCarthy, & Kirk, 1968), CELF (Semel-Mintz & Wigg, 1982), DARD (Durrell & Catterson, 1980), Rosner Test of Auditory Analysis (Rosner, 1975)	SI treatment had no significant effect for psychoeducational variables
Polatajko et al. (1991)	N=67 school-aged children with LD	Comparison of SI and PM treatment (each 1 hour a week for 6 mos.), pretest, posttests at 6 & 9 mos.)	Academic: WJPEB clusters (Woodcock & Johnson, 1977)	Improvement over time for both groups, but no significant difference between groups
Wilson et al. (1992)	N=29 school-aged children with learning & motor deficits	Comparison of SI & traditional academic tutoring (each 2 50-min. sessions per week for 1 year), pretest, 6-mo. Midtest, 12-mo. posttest	Academic: WJPEB clusters	Improvement over time for both groups, but no significant difference between groups

Predictors of Language Impairment Outcomes

Several specific predictors of improvement in children with LI have been identified in the literature, but the particular aspects of functioning measured in these studies vary considerably. Some factors that have been shown to be related to prognosis include nonverbal IQ, narrative retelling ability, expressive syntax, age at time of intervention, incidence of problem behavior, and pervasiveness of language difficulties (expressive vs. mixed receptive-expressive). Bishop and Edmundson (1987) studied factors that distinguish children with transient impairment from those with persistent impairment. They tested 87 language-impaired children at age 4, 4½, and 5½ on various language measures. The best predictor of language outcomes at 5½ was narrative retelling ability based on the Bus Story Test (BS; Renfrew, 1991), which involves telling the child a short story about a bus while showing a book of pictures illustrating the story, then asking the child to recount the story as accurately as possible while looking at the pictures. The one language measure that did not correlate with language outcome was phonological competence, or articulation. The researchers noted that nonverbal cognitive ability contributed to the progress of their sample. Botting, Faragher, Simkin, Knox, and Conti-Ramsden (2001) also investigated predictors of language outcome by testing children on nonverbal IQ and language skills at age 7, then again at age 11. Demographics such as family income and maternal education were also examined. After post-testing, the sample of 117 children was divided into a "good outcome" group, defined as those children with less than three language test scores below the 16th percentile, and a "poor outcome" group, consisting of children with three or more impaired language test scores.

Children were tested on nonverbal IQ using Raven's Colored Matrices (Raven, 1986). Although participants with a nonverbal IQ of 70 or below were excluded, the researchers nonetheless found a significant difference between the good and poor outcome groups on this measure at age 7: the good outcome group demonstrated a significantly higher nonverbal IQ than the poor outcome group. Language ability was tested using the Test for Reception of Grammar (TROG; Bishop, 1989), which measures comprehension of grammatical constructions; the British Ability Scales naming vocabulary subtest (BAS-nv; Elliot, 1983), a measure of expressive vocabulary skills; the BS (Renfrew, 1991) to measure narrative retelling skills; and the Illinois Test of Psycholinguistic Abilities grammatic closure subtest (ITPA; Kirk et al., 1968), in which the child must finish an incomplete sentence in a way that is grammatically correct. Using logistic regression models, the researchers found that the individual measures significantly contributing to outcome at age 11 were the BAS-nv, BS, ITPA, and TROG at the < .001 level and nonverbal IQ at the < .05 level. When these language measures were entered into the regression model with nonverbal IQ, the language measures all emerged as sole predictors, indicating that they outweighed the predictive value of nonverbal cognitive ability. Narrative retelling skills based on the BS and expressive syntax based on the ITPA grammatic closure subtest were shown to be the two strongest individual predictors of overall prognosis. Family income and maternal education were not shown to predict language outcomes. As in the Bishop and Edmundson (1987) study, severity of phonology/articulation impairment, as measured by the Goldman-Fristoe Test of Articulation (Goldman & Fristoe, 1986) was also not found to be predictive of language outcome at age 11. A study using a sample of 373 children from 130 monozygotic and

109 same-sex dizygotic twin pairs in England investigated variables at age 2 that predict language outcomes at age 4½ (Oliver et al., 2004). At least one member of each twin pair scored below the 15th percentile on an objective language measure at age 4½. Using parent-report measures of the children's nonverbal ability, vocabulary, grammar, and semantic/pragmatic ability at age 2 and objective measures of children's language ability at age 4½, the authors found that low language status at 4 ½ is at least as genetically related to low nonverbal cognitive ability at age 2, 3, and 4 as it is to low-language scores at 2, 3, and 4. Thurm, Lord, Li-Ching, & Newschaffer (2007) followed 118 children from age two to five who had been diagnosed with either autism, pervasive developmental disorder - not otherwise specified (PDD-NOS), or non-PDD-spectrum developmental disabilities and found that non-verbal cognitive ability at age two was the strongest predictor of language skills at age five overall. Expressive and receptive communication skills at age three were the strongest predictors of language skills at age five in children with autism. One large randomized controlled trial investigated speech-language programs in Scotland on a sample of 152 children ages 6 to 11 receiving school services (Boyle et al., 2007). While the pervasiveness of language impairment affected results in that children with mixed receptive-expressive problems were not as likely to show improvement as those with expressive gains, nonverbal IQ was found not to be a moderating variable.

Cognitive ability has more often been investigated as a predictor of improvement with literacy interventions, as opposed to speech-language interventions specifically. In their review of reading interventions, Al Otaiba and Fuchs (2002) found that out of 23 studies, 15 studies investigated the effects of IQ on reading improvement. Five of these studies found that children with low vocabulary, low verbal ability, or low IQ were more likely to be nonresponders to reading intervention (Al Otaiba & Fuchs, 2002). Fuchs and Young (2006) reviewed 13 studies of children with reading deficits to determine whether IQ could predict success with reading intervention and found that in eight out of 13 studies, IQ accounted for unique variance. It served as a stronger predictor in older children with reading comprehension training than in younger children receiving a phonological awareness intervention. In their meta-analysis of 30 literacy intervention studies, Nelson, Benner, and Gonzalez confirmed that IQ influenced treatment responsiveness (2003). Other more robust predictors included rapid naming, phonological awareness, problem behavior, and memory. Demographic information, such as ethnicity and grade, was a weaker predictor than IQ. Beitchman and Young have emphasized that cognitive ability and initial severity of literacy problems are the most reliable predictors of literacy success in early adulthood (1997).

Age is a factor that has been implicated in several studies on predictors of language outcomes, but, in essence, these studies simply speculate that intervening early is better than intervening later due to children's capacity to make rapid language gains in the early years of development. None of them compare a group of language-impaired children receiving intervention to a control group of language-impaired children not receiving intervention, presumably due to the ethical dilemma of waiting to intervene. Bruce and Hansson (2008) emphasize the importance of identifying children at risk for language impairment early so that intervention can possibly improve chances of a positive outcome and indirectly decrease the likelihood that the child will exhibit social and behavioral problems later on. Schery (1985) examined correlates of language outcomes by examining archival testing data on 718 children with language disorders spanning eight years. Demographics and social, personality, development, language, and academic characteristics of the sample were used to establish variables as predictors of language outcome after two to three years in a language program. Age was the strongest predictor for all analyses. Overall, variables of IQ, language history, socioeconomic status, physical/neurological factors, and social-emotional background did not account very well for pretest or posttest language scores. However, when predicting pretest language performance, IQ and physical factors were the strongest variables. Social-emotional factors and IQ were the only factors that made significant contributions to levels of language improvement, although IQ was a weak predictor.

Children who perform better on assessments before starting an intervention tend to improve more than children with lower scores at the outset (Cook & Campbell, 1979). Shaywitz and colleagues (1995) argue that children who have higher initial IQ scores will tend to show improvements in IQ over time, and those with lower initial IQ scores will tend to show a decrease over time, constituting a Matthew Effect for IQ. They cite Stanovich's "cumulative advantage phenomenon" (1986, p. 381), in which children with initially higher cognitive resources tend to make more progress and learn more quickly. These children likely enjoy reading if they are good at it, so they will read more, further enhancing their vocabulary and language skills beyond those of children who do not read frequently (Stanovich, 1986). Shaywitz and colleagues followed a cohort of kindergarten children over the course of seven years to determine whether a Matthew Effect occurred for reading skills, such that the reading skills of originally strong readers would get better, while the skills of poor readers would get significantly worse without intervention. Using regression models to correlate the slope with the mean for both IQ and reading, they found a small but significant Matthew Effect for IQ, but they did not find a Matthew Effect for reading skills, which tended to remain relatively steady over time. The authors speculated that had the assessment of reading skills measured the more complex skill of comprehension versus decoding, a Matthew Effect may have been found for reading, reflecting the underlying language skills of the sample. These findings serve as a precursor for examining the relationship between IQ and language outcome. Some studies have suggested that the relationship between language and nonverbal IQ is dynamic and that the two constructs influence each other (Botting, 2005; Goorhuis-Brouwer & Knijff, 2001). Other investigators have confirmed that the relationship between nonverbal IQ and language is variable, with moderate correlations found in some cases and no significant correlations found in others, depending on the language measures used (Dethorne & Watkins, 2006).

Whether language deficits involve comprehension or language production can influence prognosis with intervention. Children with less pervasive language deficits tend to have a better prognosis in response to intervention than children with deficits in multiple areas. Children with low speech perception ability and poor receptive vocabulary skills are at greatest risk of delayed phonological awareness and reading skills (Rvachew & Grawburg, 2006). Rvachew and Grawburg (2006) therefore emphasize the importance of early intervention in the preschool years, if possible. In their meta-analysis of language interventions, Law, Garrett, and Nye (2004) concluded that there may be a differential effect of expressive syntax interventions, such that intervention is effective for children who do not have receptive language problems. Boyle and colleagues confirmed that school-aged children with specific expressive language delay were more likely to improve with intervention than those with mixed receptive-expressive delay, based on a large randomized controlled trial involving 130 children ages 6 to 11 with LI (2007). Children with speech or articulation impairment, which are not necessarily accompanied by underlying language deficits, also fare better than children with LI (Johnson et al., 1999; Young et al., 2002). In terms of a prescription for improvement, younger children receive the most benefit per unit of language therapy provided, and children with lower functional communication ability need more units of language therapy to make gains than children with higher initial ability (Jacoby, Lee, Kummer, Levin, & Creaghead, 2002).

Law, Garrett, and Nye (2003) explain that no universal rules exist for choosing which type of intervention to implement with language-impaired children, nor is there a set precedent for when to intervene, nor is there clear-cut evidence to serve as a basis for a decision. However, it appears that certain factors including cognitive ability, age, and pervasiveness of impairment may influence success with intervention, such that children with better skills at the beginning of an intervention will tend to improve at a faster rate, and the sooner the intervention can be implemented and the less pervasive the impairment, the better.

CHAPTER FOUR
Aims and Hypotheses

RATIONALE

A review of the literature in this area indicates that developmental LI profoundly affects children's academic performance and effective communication with others, necessitating intervention. To date, no empirical studies using appropriate statistical analyses have examined the effects of the Montessori Method modified for children at-risk for learning disorders or the DuBard Association Method. Therefore, this longitudinal study examines the degree of improvement in language, articulation, and oromotor skills during participation in an intensive intervention program over the course of three years. Research also suggests that specific variables, including nonverbal IQ, age, and pervasiveness of language impairment may be associated with degree of improvement with intervention. Consequently, this study will also analyze the possible association of baseline nonverbal IQ, age, and expressive vs. mixed receptive-expressive impairment with language skills outcome. The characteristics of children who left the study will also be examined and compared to those who remained. This study will make a valuable contribution to developmental language impairment research in that it is the only study to date to analyze the impact of this unique intervention.

AIMS AND HYPOTHESES

Aim I

To determine whether core language skills, expressive language skills, and receptive language skills improve significantly over the course of three years of participation in the Shelton Early Intervention program, as measured by the Expressive One-Word Picture Vocabulary Test (EOWPVT), Receptive One-Word Picture Vocabulary Test (ROWPVT), Peabody Picture Vocabulary Test – 3rd Edition (PPVT-III), and the Core Language, Expressive Language, and Receptive Language Composites of the Clinical Evaluation of Language Fundamentals – 4th Edition (CELF-4) and Clinical Evaluation of Language Fundamentals – Preschool – 2nd Edition (CELF-PS2).

Aim I Hypotheses

I. a. Participants' scores on the CELF-4 and/or CELF: PS-2 Core Language Composite will increase significantly from Baseline to follow-up Time 3, indicating improvement in general language ability. Scores on the Expressive Language and Receptive Language Composites will also increase significantly from Baseline to Time 3, suggesting improvement in expressive and receptive language skills, respectively.

I. b. Participants' scores on the EOWPVT will increase significantly from Baseline to Time 3, indicating improvement in expressive vocabulary skills.

I. c. Participants' scores on the PPVT-III will increase significantly from Baseline to Time 3, indicating improvement in receptive vocabulary skills.

Aim II

To determine whether articulation skills improve with intervention, as measured by the Arizona Articulation Proficiency Scale-3rd Revision (Arizona-3), and whether oromotor skills improve, as measured by the Focal Oromotor Control subtest of the Verbal Motor Production Assessment for Children (VMPAC).

Aim II Hypotheses

II. a. Participants' scores on the Arizona-3 will increase significantly from Baseline to Time 3, indicating improvement in articulation skills.

II. b. Participants' scores on the Focal Oromotor Control subtest of the VMPAC will increase significantly from Baseline to Time 3, suggesting gains in oromotor skills.

Aim III

To determine whether a significant interaction effect exists on dependent language measures (CELF-4, CELF: PS-2, EOWPVT, and PPVT-III) for the baseline independent variable of nonverbal cognitive ability group, which will be defined as a Leiter International Performance Scale – Revised (Leiter-R) test score of Average IQ (≥ 85) or Below Average IQ (<85).

Aim III Hypothesis

A significant interaction effect will be found for the independent baseline variable of Nonverbal IQ group, such that children in the Average IQ group (Leiter-R ≥ 85) will show a greater rate of improvement in overall language, expressive language, and receptive language skills over time than children in the Below Average IQ group (Leiter-R < 85), suggesting that children with higher nonverbal cognitive ability are more amenable to language intervention.

Aim IV

To determine whether a significant interaction effect exists between the independent variables of Age and Time. Age groups will be defined as Young (3 to 5 years) and Old (6 to 9 years). The dependent variables will be standardized scores on language measures (CELF-4, CELF: PS-2, EOWPVT, and PPVT-III).

Aim IV Hypothesis

A significant interaction effect will be observed between the independent variables of Age group and Time, such that children in the Young age group (3 to 5 years) will show a greater rate of improvement in core language, expressive language, and receptive language skills over Time than children in the Old age group (6 to 9 years). This interaction will suggest that children who begin the intervention earlier are more amenable to language intervention.

Aim V

To determine whether a significant interaction effect exists between the independent variables of Impairment Type and Time. Groups will be defined as Mixed Receptive-Expressive Impairment (Standard score more than 1 SD below the mean on Receptive Language Index and Expressive Language Index of the CELF-4) and Expressive Only Impairment (Standard score more than 1 SD below the mean on Expressive Language Index and standard score of ≥ 85 on the Receptive Language Index of the CELF-4 or CELF:PS-2). The dependent variables will be standardized scores on language measures (CELF-4, CELF:PS-2, EOWPVT, and PPVT-III).

Aim V Hypothesis

A significant interaction effect will be observed between the independent variables of Type of Impairment and Time, such that children in the Expressive Only Impairment group will show a greater rate of improvement in overall language, expressive language, and receptive language skills over Time than children in the Mixed Receptive-Expressive Impairment group. This interaction will suggest that children with less pervasive impairment in only the expressive domain are more amenable to language intervention.

Aim VI

To determine whether children who exit the study (Dropouts) significantly differ from remaining children (Completers) on the Leiter-R, CELF-4, CELF-PS2, EOWPVT, PPVT-III, Arizona-3, or VMPAC Focal Oromotor Control subtest at Baseline.

Aim VI Hypothesis

The Dropout participants will show significantly higher scores on the Leiter-R, CELF-4, CELF-PS2, EOWPVT, PPVT-III, Arizona-3, and VMPAC Focal Oromotor Control subtest at Baseline, suggesting that they entered the study with significantly higher ability and were able to progress and exit the intervention.

CHAPTER FIVE
Method

PARTICIPANTS

Participants who entered the Early Intervention (EI) study in the fall of 2004 are 20 children (8 females, 12 males) who were diagnosed with a primary oral language disorder based on extensive admissions testing to the Shelton School in Dallas, Texas. The age range of participants was 3 years, 8 months to 9 years, 3 months, with a mean age of 6.17 years (*SD*=1.54). Sixty percent (N=12) of the participants were male, and 40% (N=8) were female. Seventy-five percent (N=15) were Caucasian, 5% (N=1) African American, 5% (N=1) Hispanic, and 15% (N=3) were classified in the "Other" category. Experienced speech-language pathologists, psychologists, and educational diagnosticians administered admissions tests, which included measures of verbal and nonverbal cognitive ability, short-term and working memory, reading comprehension, decoding skill, expressive and receptive language, auditory processing, and articulation. Within the Shelton system, children are categorized according to a unique pattern of scores on standardized tests. Pattern 6 is assigned to children with a predominant Oral Language Disability or Dysphasia. These children typically obtain a profile of skills in the following approximate ranges: verbal IQ – low average (85 to 89) or below (<85); nonverbal IQ – average (85 to 115) or above (>115); auditory processing, processing speed, visual perceptual ability, reading comprehension, spelling, and handwriting - below average (<85); and reading rate and accuracy – average (85 to 115) or below (<85). Pattern 6 children may also have moderate to severe receptive or expressive difficulties as evidenced by below average (<85) scores on receptive or expressive tasks, as well as significant weaknesses in auditory memory and moderate to severe articulation impairment on the Arizona Articulation Proficiency Scale.

Specific inclusion criteria for the study are:

1) Participants must have a primary oral language disorder, namely phonological (articulation) disorder, expressive language disorder, or mixed receptive-expressive language disorder, as diagnosed by a certified speech-language pathologist. A primary oral language disability was defined in this study by a standard score of more than one standard deviation (SD) below the mean (84 or below) on an articulation test (Arizona-3) or a comprehensive standardized language test encompassing both expressive and receptive language skills (CELF: PS-2 or CELF-4).
2) Children must be in the age range of 3 years to 9 years, 11 months.

Exclusion criteria are:

1) Children who scored in the average or higher range (standard scores of 85 or above) on all language and articulation measures (i.e., CELF: PS-2, CELF-4, & Arizona-3) were excluded from the study, as they did not meet criteria for a diagnosis of a phonological, expressive language, or mixed receptive-expressive language disorder.
2) Participants could not be outside the age range of 3 years to 9 years, 11 months.
3) Participants could not have a primary behavioral or emotional problem, as evidenced by clinically significant scores on the BASC Clinical Scales (with the exception of the Attention scale); clinical observation by a speech-

language pathologist of the child’s being unable to receive a sample DuBard Association Method lesson; and/or prior diagnosis of primary behavior or emotional disorder (e.g., Oppositional Defiant Disorder, Disruptive Behavior Disorder, Conduct Disorder, or Mood Disorder).

Existing students of the Shelton School who met criteria for Pattern 6 were first considered for inclusion in the intervention based on the severity of their language and/or articulation problems. It should be noted that some existing students included in the study showed severe language deficits in addition to severe perceptual problems and characteristics on the Pervasive Developmental Disorder Spectrum. These children were included because of their language needs, which called for intensive intervention. Newcomers to the Shelton School were evaluated based on their profile of admissions test scores and included in the intervention if they met criteria for a Pattern 6 diagnosis and had severe language problems. Children with less severe language difficulties or other diagnoses including dyslexia who were not assigned to one of the Early Intervention classes were assigned to other language instruction methods offered at the Shelton School, including Alphabetic Phonics and Sequential English Education. Assignments are made for every student based on the child’s specific profile of learning needs. This Shelton School Model is called Profile to Prescription (Pickering, Coffman, Stanislav, Kneese, & Snyder, 2003). Selection of a no-treatment control group was not ethically merited, as the most severely impaired children were assigned to the Early Intervention study. In terms of attrition, a very careful descriptive consideration will be given to those children who exit the study.

PROCEDURES

Upon acceptance to the Shelton School, participants entering the EI program completed routine admissions testing, consisting of tests of intellectual ability, achievement, language, attention, and memory. Informed consent was obtained from parents, and informed assent was given by children entering the study. The intervention program began in the fall of 2004 and has been integrated into the Shelton curriculum as a unique program for children with severe language difficulties. The current study analyzes assessments administered during the 2004-2007 academic years. Prior to beginning the intervention, participants were assigned to one of two classrooms: eight children 3 to 4 years of age, with the exception of two delayed 5-year-olds, were assigned to the Beginner classroom, and the other 12 children 5 to 9 years of age were assigned to the Intermediate classroom. Each classroom was taught by a Montessori-trained teacher and a speech-language pathologist. These two experienced professionals were trained in the Montessori Method, the DuBard Association Method, or some combination of the two. The Beginner classroom teacher was certified in traditional Montessori education by Montessori Education Centers Associated (MECA) and trained in the Sequential English Education (SEE) language instruction method. The Beginner classroom speech-language pathologist (SLP) was a certified academic language therapist (CALT) at the Teaching level through the Academic Language Therapy Association (ALTA), received Beginner-level training in the DuBard Association Method, received MACAR training, and was trained in the SEE method. The Intermediate classroom teacher was a CALT and was certified in both primary (ages 3-6) and elementary (ages 6-9) Montessori teaching. She also co-taught training courses in MACAR, offered by the Shelton School. The Intermediate classroom SLP was a CALT at the Therapy level through ALTA and received an overview training course in the Montessori Method. Licensed occupational therapists provided occupational therapy (OT) with incorporated sensory integration techniques to groups of children twice a week in the children's classrooms. The OT's were also certified or received training in the administration of the Sensory Integration and Praxis Tests.

Assessments were administered at baseline in the early fall of 2004, and post-testing occurred from late spring to early summer in 2005, 2006, and 2007. Testing took place during the school day, usually across several sessions with different examiners due to the long assessment battery. Language evaluations were performed by Shelton School staff with a masters degree in speech-language pathology; cognitive and academic tests were administered by doctoral candidates in psychology and licensed psychologists; and motor and sensory-integration tests were given by licensed occupational therapists.

MEASURES

Upon enrollment in the EI study, participants received a comprehensive evaluation of cognitive ability, language/articulation skills, academic skills, attention and memory, perceptual skills, and motor skills, both at baseline and after one, two, and three years of intervention to assess the efficacy of the intervention program. The measures chosen were standardized, well-established in the field, and psychometrically sound. In those cases in which the child was either above or below the established age range for existing standardized scores for a particular measure, raw scores were recorded to track each child's progress over time. The following measures of nonverbal cognitive ability, language, articulation, and oromotor skills were selected for the current analysis.

Nonverbal Cognitive Ability

1. Leiter International Performance Scale—Revised (Leiter-R; Roid & Miller, 1997). The Leiter-R is an individually administered test of nonverbal intelligence, memory, and attention for ages two years to 20 years, 11 months. The examiner pantomimes and uses gestures to convey instructions to the examinee in order to eliminate the need for verbal communication. In the current study, this nonverbal test is compared with language assessments in order to evaluate the size of the discrepancy and contribute to appropriate diagnoses regarding LI. The Leiter-R Brief IQ Screener, with an average administration time of 25 minutes, was used to measure the child's ability to perform complex nonverbal manipulations of visualization and fluid reasoning. The Brief IQ Composite consists of four subtests: Figure Ground, Form Completion, Sequential Order, and Repeated Patterns. The Figure Ground subtest, also known as The Find It Game, measures ability to identify figures or designs hidden in a complex picture. Form Completion, or The Put It Together Game, measures the child's ability to choose the correct whole object in a complex scene based on a picture of the object in jumbled pieces. The Sequential Order subtest, or The Which Comes Next Game, involves the child's placing cards or foam objects of different sizes, shapes, colors, or patterns into a logical sequence. Repeated

Patterns, The Over and Over Game, requires the child to fill in the missing stimulus in a pattern of objects or cards.

The Leiter-R Brief IQ Screener yields standard scores with a mean of 100 and standard deviation of 15. The test was normed on a sample of 1,719 individuals that reflected the proportions of ethnic groups in the 1993 U.S. Census update survey. Reliability coefficients for the Brief IQ Screener are .88 for ages 2-5, .90 for ages 6-10, and .89 for ages 11-20. Test-retest reliability coefficients were obtained by administering the Leiter-R to 163 children and adolescents ages 2-20 on two occasions (time lapse between administrations was not reported in the manual). Test-retest reliability coefficients are .88 for ages 2-5, .91 for ages 6-10, and .96 for ages 11-20. Practice effects were observed, but the test creators assert that this is commonly seen on nonverbal measures (Roid & Miller, 1997).

In terms of validity for the Leiter-R, content, criterion, and construct validity were evaluated. Content validity was verified using item-response theory (IRT) and internal consistency analyses of individual items, as well as meticulous research on the items selected and their representation of intelligence theory. Nonverbal content validity was verified by 114 examiners in the standardization phase who determined that all items could be administered completely nonverbally. The Brief IQ Screener showed concurrent validity with the Fluid Reasoning subtests of the Woodcock-Johnson-Revised ($r=.62$ to .77 in ages 6-20), the WISC-III Full Scale IQ ($r=.85$), and the WISC-III Performance IQ ($r=.85$). Criterion validity was evaluated by determining specificity and sensitivity of the Leiter-R for classifying certain groups, with the best rates shown for classifying cognitive delay.

Language Skills

1. Clinical Evaluation of Language Fundamentals – 4th Edition (CELF-4; Semel et al., 2003). The CELF-4 is an individually administered test battery for the assessment and diagnosis of language disorders in students age five to 21. The current study examined the Core Language Score and Receptive and Expressive Language Indices to determine whether significant change occurred in these areas. The Core Language Score (CLS) consists of four subtests for ages five to eight: 1) Concepts and Following Directions (C&FD), in which the student listens to an increasingly complex sequence of directions and points appropriately to pictures in a stimulus book; 2) Word Structure (WS), which measures the student's ability to use morphological rules to identify inflection and use correct pronouns in reference to people and things; 3) Recalling Sentences (RS), which measures the student's ability to listen to increasingly complex sentences and repeat them verbatim; and 4) Formulated Sentences (FS), which requires the student to create and verbalize semantically and syntactically correct sentences of increasing complexity. The CLS for ages nine to 12 consists of the same subtests except for WS, which is replaced with Word Classes-Total (WC-T), a measure of the student's ability to comprehend and describe relationships between associated words. This subtest consists of a Receptive and Expressive part of each item, resulting in a WC-Receptive (WC-R) score and a WC-Expressive (WC-E) score; there are also two levels of difficulty to this subtest. The

Receptive Language Index (RLI) for ages five to eight consists of the C&FD, WC-R, and Sentence Structure (SS) subtests. The SS subtest measures the student's ability to listen to sentences and select the correct pictorial representation of each. The RLI for ages nine to 12 consists only of C&FD and WC-R subtests. The Expressive Language Index (ELI) for ages five to eight is made up of WS, RS, and FS subtests while the ELI for ages nine to 12 is made up of RS, FS, and WC-E subtests.

The CELF-4 was normed in 2002 using a sample of more than 4,500 individuals that was representative of the United States population of students age five to 21. The sample was stratified by age, sex, ethnicity, geographic region, and parent education level to ensure adequate representation from all groups. Test-retest reliability was calculated by administering the CELF-4 to a sample of 320 individuals from the standardization sample. The test was administered once and then repeated with each examinee within a range of seven to 35 days later. Test-retest coefficients for the overall sample were calculated using Fisher's *z* transformation and were corrected for variability of the standardization sample. The CLS showed a coefficient of .92, the RLI showed a coefficient of .89, and the ELI showed a coefficient of .92.

Internal consistency was measured using Cronbach's coefficient alpha and split-half reliability. The average coefficient alpha across age groups for the CLS was .95. The RLI showed a coefficient alpha of .89, and the ELI showed a coefficient alpha of .93, demonstrating good internal consistency of items. Split-half reliability was determined by splitting subtest items into odd- and even-numbered items and correlating them to obtain an estimate of distribution of the items. Correlations corrected using the Spearman-Brown formula were as follows: .95 for the CLS, .90 for the RLI, and .93 for the ELI, suggesting excellent split-half reliability. Interrater reliability was calculated for subtests that required some level of clinical judgment to score responses. The FS subtest showed a mean agreement of .90, WS showed a mean agreement of .98, and both levels of the WC subtest showed a mean agreement of .95 between independent raters.

Content validity evidence for the CELF-4 consists of extensive research on language skill development that served as a basis for test items. Subtests also correlate highly with their respective composites, suggesting that they each contribute measurement of a particular skill to the overall scores. The CLS shows correlations ranging from .65 to .97 with other composites. The RLI and ELI show a moderate correlation of .79, which is expected because they do not share any subtests. The CELF-4 was correlated with the CELF-3 to determine convergent validity. The CLS had a correlation of .84 with the corresponding composite on the CELF-3, and the RLI and ELI each showed a correlation of .79 with the corresponding composite.

2. Clinical Evaluation of Language Fundamentals; Preschool – 2nd Edition (CELF:PS-2; Wiig et al., 2004). The CELF:PS-2 is an individually administered test to assess and diagnose language disorders in children age three to six. Like the CELF-4, the CELF:PS-2 also includes a CLS, RLI, and ELI, which were administered in the current study to evaluate improvement in overall language ability, receptive language skills, and expressive language skills, respectively. The CLS for both age groups (three to four and

five to six) consists of the SS, WS, and Expressive Vocabulary (EV) subtests. The EV subtest measures the child's ability to correctly identify pictures of objects, people, or activities in a stimulus book. The RLI for ages three and four consists of the SS, C&FD, and Basic Concepts (BC) subtests. The BC subtest measures the child's ability to differentiate between concepts such as size, positioning, and number of objects. The RLI for ages five and six consists of the SS, C&FD, and WC-R subtests. The ELI for both age groups is made up of the WS, EV, and RS subtests.

The CELF:PS-2 was standardized using a sample of 1,150 children stratified by age into eight six-month age groups of 100 children each. The test composites yield standard scores with a mean of 100 and standard deviation of 15. Test-retest reliability was measured using a sample of 13-17 children from each age group for a total of 120 children; the second administration of the test was given between two and 24 days after the first. The following test-retest coefficients were calculated for the CLS, RLI, and ELI composites using Fisher's *z* transformation and were corrected for variability in the standardization sample: CLS – r=.91, RLI – r=.91, and ELI – r=.94. Cronbach's coefficient alpha was used to calculate internal consistency; the average coefficients across age ranges were .90 for the CLS, .91 for the RLI, and .92 for the ELI. Split-half reliability coefficients were .92 for the CLS, .92 for the RLI, and .94 for the ELI. Inter-scorer agreement was calculated for the WS, EV, and the expressive part of the WC subtest, which require a degree of clinical judgment on the examiner's part to score responses. The WS and EV subtests each yielded an interrater reliability coefficient of .97, and the WC-E yielded a coefficient of .95.

In terms of validity for the CELF:PS-2, content validity was confirmed through extensive research on language skill development that served as a basis for test items. Subtests correlate highly with their respective composites, suggesting that they each contribute measurement of a particular skill to the overall scores. The CLS shows correlations ranging from .85 to .93 with other composites. The RLI and ELI show a more moderate correlation of .76, which is expected because they do not share any subtests. The CELF:PS-2 was correlated with other measures of the same language constructs, including the CELF-Preschool, CELF-4, and the Preschool Language Scale-4th Edition (PLS-4). Correlations of the CLS, RLI, and ELI with corresponding composites on the CELF-Preschool ranged from .75 to .88. The CELF:PS-2 showed correlations ranging from .69 to .80 with corresponding composites on the CELF-4 and correlations ranging from .73 to .76 with corresponding composites on the PLS-4.

3. Expressive One-Word Picture Vocabulary Test – 3rd Edition (EOWPVT; Brownell, 2000a). The EOWPVT is a measure of expressive vocabulary for students age two to 18 years, 11 months. The examinee is asked to name objects, actions, and concepts that may or may not be familiar. The test was designed to be used in conjunction with other measures to obtain an overall picture of a child's language abilities. The test was normed in 1999 on 3,661 individuals who closely represented the United States population. The EOWPVT yields standard scores with a mean of 100 and standard deviation of 15, as well as percentile ranks and age equivalents. Reliability testing included Cronbach's coefficient alpha and split-half reliability coefficients for internal consistency estimates.

Coefficient alphas across age groups ranged from .93 to .98 with a median of .96, suggesting excellent internal homogeneity of items. Split-half reliability coefficients ranged from .96 to .99 across age groups, with a median of .98. Test-retest reliability gives evidence of temporal stability of the test. Corrected test-retest correlations range from .88 to .97 with a coefficient of .90 for the entire sample. Interrater reliability was calculated on a sample of 20 individuals based on scores from two different examiners, yielding a correlation of .93.

Content validity was confirmed by the rigorous procedures used for test item selection, to assure that test content represented the construct being measured. Concurrent validity was measured by correlating the EOWPVT with other expressive and receptive vocabulary tests. The EOWPVT showed a corrected correlation of .72 with the Expressive Vocabulary Test (EVT); .76 with the Peabody Picture Vocabulary Test-3rd Edition (PPVT-III), which is a measure of receptive vocabulary; a .72 with the Receptive One-Word Vocabulary Test (ROWPVT); and a .78 with the Wechsler Intelligence Scale for Children – 3rd Edition (WISC-III) Vocabulary subtest. The EOWPVT also showed correlations of .87 with the Expressive Language Score of the CELF-3 and .85 with the 1990 edition of the EOWPVT.

4. Peabody Picture Vocabulary Test – 3rd Edition (PPVT-III; Dunn & Dunn, 1997). The PPVT-III was designed to measure receptive, or hearing, vocabulary skills and serves as a screening test of verbal intellectual ability for individuals age two years, sixth months to 90+ years. The examinee chooses the picture that best represents a word spoken by the

examiner out of four picture choices. The test was standardized in 1995 and 1996 on a sample of 2,725 individuals that represented the correct demographic proportions of the United States population in 1994. Reliability data on internal consistency was obtained using alpha coefficients and split-half reliability. Alpha coefficients ranged from .92 to .98 across age groups, with a median of .95 for both forms of the test. Split-half reliability coefficients, in which the actual items taken by each examinee were split into odd- and even-numbered items then correlated, ranged from .86 to .97, with a median of .94 for both forms of the test. Alternate forms reliability was obtained by administering two parallel forms of the test to the entire standardization sample. Alternate forms reliability coefficients calculated using standard scores ranged from .88 to .96 across age groups, with a median of .94. Test-retest reliability was also obtained by administering the test a second time one month later to a sample of individuals; these coefficients ranged from .91 to .94.

Content validity is based on the rigorous item selection procedures used by the test developers to compile a sample of vocabulary words that could be clearly depicted using black-and-white line drawings. Construct validity is substantiated by content-related evidence that the test measures the underlying construct of hearing vocabulary, as well as evidence that vocabulary is the single most indicative test of intellectual ability. Criterion validity is based on the PPVT-III's correlation with other measures of vocabulary and intellectual ability. The PPVT-III was compared to the WISC-III on a sample of 41 children age seven years, 11 months to 14 years, four months, yielding correlations ranging from .82 to .92. When compared to the Kaufman Adolescent and Adult

Intelligence Test (KAIT) using a sample of 28 adolescents age 13 years through 17 years, eight months, the PPVT-III showed correlations ranging from .76 to .91. When the PPVT-III and the Kaufman Brief Intelligence Test (K-BIT) were administered to a sample of 80 adults age 18 years through 71 years, one month, correlations ranged from .62 to .82. Correlations were higher between the PPVT-III and vocabulary sections of the tests, as opposed to nonverbal sections, as expected by the test developers.

Articulation

1. Arizona Articulation Proficiency Scale – 3rd Revision (Arizona-3; Fudala, 2003). The Arizona-3 is an individually administered test of articulation skill in children 18 months through 18 years of age. The test was standardized on a sample of 5,515 individuals who roughly represented demographic characteristics of the United States population. The test yields standard scores with a mean of 100 and standard deviation of 15, as well as percentile ranks and levels of articulation impairment. Internal consistency estimates were calculated using Cronbach's alpha, yielding coefficients ranging from .96 to .78 across age groups, with a median of .925. Interrater reliability was calculated to determine stability of scoring between different examiners; a sample of 503 children in the 1st grade was followed until they reached 3rd grade (N=401 in 2nd grade, N=342 in 3rd grade), and a mean interrater reliability coefficient of .95 was obtained across testing points. Test-retest reliability was obtained with a sample of 259 individuals ranging from 18 months to 19 years, five months of age. The median reliability correlation was .97. Validity research for the Arizona-3 began in 1959, when recorded speech samples were

judged for articulation difficulties and a scoring system was developed. In a study of 45 children age six to 12, the Arizona test was administered, and additional spontaneous speech samples for each child were recorded and scored with an established 9-point defectiveness continuum. A high correlation of .92 was obtained between the recorded sample scores and the Arizona test scores, suggesting excellent construct validity (Barker & England, 1962). The Arizona-3 has also been correlated with other established measures of articulation to obtain criterion-related validity. The Arizona-3 shows a correlation of .88 with the Goldman-Fristoe Test of Articulation and a correlation of .91 with the Photo Articulation Test.

Oromotor Skills

1. Verbal Motor Production Assessment for Children (VMPAC; Hayden & Square, 1999). The VMPAC was designed to measure integrity of neuromotor abilities involved in dysarthria and apraxia of speech disorders. The Focal Oromotor Control subtest was used for the purposes of this study; it consists of 46 items that the examiner utilizes to determine the child's level of oromotor control in the mandibular (jaw), labial-facial (lips and face), and lingual (tongue) areas. The examiner assesses both speech phoneme production and non-speech movements in each area, as well as how the areas function together. The test was standardized in fall 1997 and spring 1998 using a sample of 1,040 children between the ages of 3 years and 12 years, 11 months. The sample was stratified by age, gender, ethnicity, parent education level, and region and adequately represented the 1995 United States population. The VMPAC yields percentage scores from 0 to 100

for each age group in the standardization sample. The 5th, 50th, and 95th percentile scores are also provided for corresponding percentage scores, and severity ranges of mild, moderate, or severe impairment can be assigned to children showing delay.

Test-retest reliability data were obtained for 115 children from the standardization sample, who took the test on two occasions 7 to 14 days apart. Test-retest correlation coefficients for each test area ranged from .56 to .90, with the Focal Oromotor Control subtest yielding the highest correlation of .90. However, the authors do not explicitly state whether this correlation is significant. Interrater reliability was assessed on a sample of 119 children from the standardization sample aged 3 to 7 years; coefficients ranged from .93 to .99 across subtests. Content validity for the VMPAC was established through extensively researched information on children's developmental oromotor skills and item selection. Construct validity is evidenced by obtaining correlations between the subtests of the VMPAC to determine whether their relationships reflect established research, e.g., since the Focal Oromotor Control and Sequencing subtests measure skills that are both cortical and not distinctly localized in the younger child's brain, these subtests are expected to show high correlation with each other, which is the case ($r=.84$ for age 3). Special group studies also confirmed construct validity, showing that the VMPAC effectively differentiates children with articulation, phonological, oromotor, and generalized motor difficulties from unimpaired children.

STATISTICAL ANALYSES

The current study, which began in 2004, was modeled as a repeated measures, within-subjects design with subject as own control. This design ensured that children with severe speech and language deficits were not denied the right to receive early intervention. A control group was not deemed ethically appropriate considering the severity of deficits seen in these children and the importance of intervening early. Demographic data including the age of each child, numerical codes for gender and ethnicity, and standardized test scores were double-entered into an electronic database and checked for discrepancies before transferring to the Statistical Package for the Social Sciences (SPSS) electronic software.

Statistical analyses include descriptive statistics on the demographic characteristics of the sample, including gender, age, and ethnicity, as well as baseline nonverbal IQ and language measures. Repeated measures analyses of variance (ANOVA) were used to determine whether significant improvements in language and articulation measures occurred over the course of three years of intervention. If the ANOVA F statistic was significant, pairwise t tests were conducted to determine which time points were significantly different. The Holm's Sequential Bonferroni correction was used to control for Type I error. The Reliable Change Index (RCI) (Christensen & Mendoza, 1986; Jacobson, Follette, & Revenstorf, 1984; Jacobson & Truax, 1991) was used as a comparison to ANOVA to investigate whether language scores improved reliably over time. Reliable change was examined for the difference between baseline, or first recorded, and Time 3, or last recorded, test scores. To calculate the RCI, the standard error of the difference between the pretest and posttest scores (S_{diff}) must first be calculated (Christensen & Mendoza, 1986):

$$S_{diff} = \sqrt{2(S_E)^2}$$

S_E is the standard error of measurement. The RCI is calculated by subtracting the pretest score from the posttest score for a subject and dividing that by S_{diff} between the two test scores (Jacobson & Truax, 1991):

$$RC = \frac{x_2 - x_1}{S_{diff}}$$

When the RCI > 1.96 (p<.05), the change from pretest to posttest is deemed to be reliable, as it represents change beyond what would be expected due to chance. In order to facilitate calculation of individuals' RCI scores, the Reliable Change Generator Version 2.0 software (Devilly, 2004) was used in place of hand calculations.

Additional independent variable and interaction terms for Age (3-5 years vs. 6-9 years), Nonverbal IQ (Leiter-R score < 85 vs. ≥ 85), and Pervasiveness of Impairment (CELF-4 RLI score < 85 & CELF-4 ELI < 85 vs. CELF-4 RLI score ≥ 85 & CELF-4 ELI and/or Arizona-3 < 85) were also calculated, due to these demographic and cognitive measures' being identified in the literature as predictors of improvement with language interventions. Appropriate post-hoc analyses were conducted for significant findings.

Aim I: Language Hypotheses

I. a. Participants' scores on the CELF-4 and/or CELF: PS-2 Core Language Composite will increase significantly from Baseline to Time 3, indicating improvement in general language ability. Scores on the Expressive Language and Receptive Language Composites will also increase significantly from Baseline to Time 3, suggesting improvement in expressive and receptive language skills, respectively.

I. b. Participants' scores on the EOWPVT will increase significantly from Baseline to Time 3, indicating improvement in expressive vocabulary skills.

I. c. Participants' scores on the PPVT-III will increase significantly from Baseline to Time 3, indicating improvement in receptive vocabulary skills.

Aim I Hypotheses were analyzed using the one-way repeated measures ANOVA statistic to determine overall significance over three years of intervention. Multivariate and univariate tests were examined for significance below .05, and if the ANOVA showed significant change, pair-wise *t* tests determined where significance occurred between baseline and time points 1, 2, and 3. Bonferroni corrections were conducted to adjust for Type I error since multiple *t* tests were conducted. The RCI was also computed to determine reliable change over time.

Aim II: Articulation and Oromotor Skills Hypotheses

II. a. Participants' scores on the Arizona-3 will increase significantly from Baseline to Time 3, indicating improvement in articulation skills.

II. b. Participants' scores on the Focal Oromotor Control subtest of the VMPAC will increase significantly from Baseline to Time 3, suggesting gains in oromotor skills.

Aim II Hypotheses were analyzed using the one-way repeated measures ANOVA statistic to determine overall significance. Multivariate and univariate tests were examined for significance below .05, and if the ANOVA showed significant change, pair-wise *t* tests determined where significance occurred between time points. Bonferroni corrections were conducted to adjust for Type I error, and the RCI was computed to determine reliable change.

Aim III: Interaction Effect of IQ Hypothesis

A significant interaction effect will be found for the independent baseline variable of Nonverbal IQ group with Time, such that children in the Average IQ group (Leiter-R ≥ 85) will show a greater rate of improvement in overall language, expressive language, and receptive language skills over time than children in the Below Average IQ group (Leiter-R < 85), suggesting that children with higher cognitive ability are more amenable to language intervention.

One-way repeated measures ANOVAs were conducted on CELF-4, CELF:PS-2, EOWPVT, and PPVT-III scores to determine the within-subjects effect of Time and the between-subjects effect of Nonverbal IQ group. Interactions and main effects were examined to determine significance at the .05 level.

Aim IV: Interaction Effect of Age Hypotheses

A significant interaction effect will be observed for the independent demographic variable of Age group with Time, such that children in the Young age group (3 to 5 years) will show a greater rate of improvement in overall language, expressive language, and receptive language skills over time than children in Old age group (6 to 9 years), suggesting that children who begin the intervention earlier are more amenable to language intervention.

Repeated measures ANOVAs were conducted on CELF-4, CELF:PS-2, EOWPVT, and PPVT-III scores to determine the within-subjects effect of Time and the between-subjects effect of Age group. Interactions and main effects were examined for significance at the .05 level.

Aim V: Relationship Between Impairment and Response Hypothesis

A significant relationship will be observed between the categorical variables of Impairment Type and Response, such that children in the Expressive/Phonological

Impairment Group will make up a greater proportion of Responders than children in the Mixed Receptive-Expressive Impairment Group on language measures. The Expressive/Phonological Impairment Group must show below average (< 85) standard scores on either the expressive or articulation tests at baseline (CELF-4 ELI, Arizona-3) and average scores on the CELF-4 RLI. The Mixed Receptive-Expressive Impairment Group must show below average standard scores at baseline on the CELF-4 ELI and the CELF-4 RLI. Response Group will be defined as Responders, Non-Responders, or Opposites as determined by prior RCI calculations for each measure. This relationship will suggest that children with less pervasive impairment in only the expressive domain are more likely to show response to intervention.

A two-way contingency table analysis was conducted to determine the relationship between Impairment Type and Response group. If Pearson's chi square (χ^2) statistic was significant, pairwise comparisons were conducted to determine which groups showed significantly different proportions in each category, and a Bonferroni correction was applied to reduce Type I error rate.

Aim VI: Attrition Hypothesis

The Dropout participants will show significantly higher scores on the Leiter-R, CELF-4, CELF: PS-2, EOWPVT, PPVT-III, Arizona-3, and VMPAC Focal Oromotor Control subtest at Baseline, suggesting that they entered the study with significantly higher ability and were able to progress and exit the intervention.

A series of independent *t* tests was conducted on the means of Baseline scores for Completers versus Dropouts. Bonferroni corrections were applied to reduce Type I error rate.

CHAPTER SIX
Results

ANALYSIS OF DATA

Demographic characteristics for the sample of 20 participants enrolled in the Shelton School's EI Program were calculated and are presented in the table below.

Table 4. Demographic Characteristics of the Sample

	Frequency (%)	Descriptive Statistics
Gender		
Male	12 (60)[a]	<u>M</u> age=5.62, SD=1.46
Female	8 (40)	<u>M</u> age=6.82, SD=1.45
Total	20 (100)	<u>M</u> age=6.17, SD=1.54
Ethnicity		
Caucasian	15 (75)	---
African American	1 (5)	---
Hispanic	1 (5)	---
Other	3 (15)	---
Age at Baseline		
3	1 (5)	---
4	4 (20)	---
5	5 (25)	---
6	2 (10)	---
7	6 (30)	---
8	1 (5)	---
9	1 (5)	---

[a]The first value in the column represents the number of participants; the value in parentheses represents the percentage of the overall sample.

Descriptive statistics were also calculated to determine the baseline characteristics of the sample on standardized measures and are presented in the table below.

Table 5. Descriptive Analyses of Baseline Nonverbal IQ and Language Skills

Measure	Skill Measured	N	Mean	SD	Range
Leiter-R Brief IQ	Nonverbal IQ	20	84.15	17.10	67
CELF-4 CLC	Core Language	15	59	15.01	47
CELF-4 RLI	Receptive Language	15	74.67	19.69	61
CELF-4 ELI	Expressive Language	15	59.27	13.41	40
CELF:PS-2 CLC	Core Language	12	77.67	14.34	49
CELF:PS-2 RLI	Receptive Language	3	78.67	10.21	19
CELF:PS-2 ELI	Expressive Language	3	77.33	17.67	32
Arizona 3	Articulation	20	73.55	12.98	47
PPVT 3	Receptive Vocabulary	20	92.6	8.13	32
EOWPVT	Expressive Vocabulary	20	81.85	11.01	42
VMPAC subtest[a]	Focal Oromotor Control	20	78.6	11.83	35

[a]Descriptive statistics are based on percentage of items correct on this subtest as standard scores are not available.

Concurrent deficits in the areas of motor coordination, visual perception, cognition, short-term memory, academics, and attention were noted throughout the sample. Several standardized measures representative of these skill areas were administered at baseline to characterize the variety of concomitant impairments present in the sample. Children who obtained standard scores of more than one SD below the mean (< 85) on the skill measure were considered to have a deficit in the corresponding area. Motor coordination deficits were measured by the Motor Coordination subtest of the Beery-Buktenica Developmental Test of Visual-Motor Integration (Beery VMI; Beery & Beery, 2004), and visual perception deficits were measured by the Visual Perception subtest of the Beery VMI. Verbal cognition deficits were measured by the Slosson Intelligence Test (Slosson, Nicholson, & Hibpshman, 1991), and nonverbal cognition deficits were measured by the Leiter-R Brief IQ Composite (Roid & Miller, 1997). Memory deficits were measured by the Short-term Memory Composite of the Woodcock-Johnson III Tests of Achievement

(WJ-III; Woodcock, McGrew, & Mather, 2001). Academic deficits were measured by the Letter-Word Identification (single-word reading), Spelling, Applied Problems (math word problems), and Writing Samples (creating and writing sentences) subtests of the WJ-III. Attention deficits were defined by "at-risk" elevations (T-scores of 60 or above) on the Attention Problems scale on the Parent or Teacher forms of the Behavioral Assessment System for Children (BASC; Kamphaus & Reynolds, 1992).

Frequency analyses were conducted on baseline measures representative of the above skill areas to convey the heterogeneous nature of the current sample and are summarized in Table 6.

Table 6. Frequency Analyses of Children with Concomitant Deficits in Auxiliary Skill Areas

Measure	Deficit Frequency (%)	No-deficit Frequency (%)	Missing Data (%)
Slosson Intelligence Test	18 (90)	1 (5)	1 (5)
Leiter-R Brief IQ Composite	11 (55)	9 (45)	---
Beery VMI Motor Coordination	14 (70)	5 (25)	1 (5)
Beery VMI Visual Perception	10 (50)	10 (50)	---
WJ-III Short-Term Memory Composite	10 (50)	7 (35)	3 (15)
WJ-III Letter-Word Identification	11 (55)	9 (45)	---
WJ-III Spelling	11 (55)	9 (45)	---
WJ-III Applied Problems	12 (60)	8 (40)	---
WJ-III Writing Samples	9 (45)	5 (25)	6 (30)
BASC-Parent: Attention	7 (35)	12 (60)	1 (5)
BASC-Teacher: Attention	7 (35)	11 (55)	2 (10)

Note: Missing data points were due to discontinued testing or failure of teacher or parent to complete BASC forms.

Modifications to Initially Proposed Analyses

In light of an examination of available data, certain statistical analyses that were initially proposed will be modified or will not be carried out due to limitations in interpretability. As proposed, one-way repeated measures ANOVAs were conducted on language measures to determine significant change over time. ANOVAs will be interpreted according to procedures described by Green and Salkind (2005). In addition to examination of multivariate and univariate *F* tests and Bonferroni-corrected pairwise comparisons, polynomial contrasts will be reported when appropriate. These within-subjects contrasts determine whether a linear, quadratic, or cubic trend is seen in the data over time. This post-hoc test is more appropriate for data with equal time intervals between administration of test measures (Green & Salkind, 2005). A linear trend indicates a steady increase or decrease in means over time; a quadratic trend indicates a single "bend" in the data such that the mean at one time point increases or decreases in opposition to the other time points; and a cubic trend indicates two "bends" in the data (Laija, 1997).

Although repeated-measures ANOVAs were to be conducted on the CELF: PS-2, very limited data prohibited appropriate statistical analysis on the Expressive Language Index (ELI) and Receptive Language Index (RLI). Since the CELF: PS-2 age norms encompass 3- to 6-year-olds and the CELF-4 age norms encompass 5- to 21-year-olds, both measures were administered to 5- and 6-year-old children who fell in the overlapping normative age group at baseline. Of the CELF: PS-2 language composite scores, only the Core Language Composite was administered to participants at baseline, with the exception of three children who also received the ELI and RLI at baseline. The reason that certain children received these indices while others did not remains unclear. Due to the very limited data available on the CELF: PS-2 ELI and RLI, these composite scores will not be included in any statistical analyses.

The psychometric properties of the VMPAC were deemed inappropriate for the initially proposed statistical analyses. This measure only produces raw scores and percentage of items correct on subtests, and only the 5th, 50th, and 95th percentiles are available for conversion of percentage scores. Due to lack of specificity with regard to the full range of percentages and equivalent percentile scores in each age group, not all percentage scores could be converted into equivalent percentiles or z-scores. The percentage scores are not standardized, so the same score may represent very different skill levels depending on the child's age. For example, a 3-year-old who obtains a score of 62% correct is considered to be within normal limits for his or her age range, whereas a 7-year-old who obtains a score of 62% is considered to have severe oromotor skill deficits. The VMPAC test manual does not provide information on how to assess change over time, and no information is provided as to whether test-retest reliability coefficients were statistically significant (McCauley & Strand, 2008). Therefore, the psychometric properties of the VMPAC do not appear to be suitable for ANOVA or RCI. A descriptive analysis will replace the initially proposed analyses to determine change in skill ranges over time on the VMPAC. The VMPAC uses four skill ranges: Within Normal Limits (WNL), Mild Deficit, Moderate Deficit, and Severe Deficit.

An additional hypothesis regarding the impact of the participants' impairment type at baseline on response to intervention was also added in light of research showing that children with expressive or articulation deficits only, as opposed to those with more pervasive difficulties in both receptive and expressive domains, are more amenable to intervention. Children were therefore categorized according to their first observed CELF-4 ELI and RLI scores and baseline Arizona-3 articulation scores to determine whether they exhibited below average (standard score < 85) scores in expressive and/or phonological (articulation) domains only, or expressive, phonological, and receptive domains. These categories were named Expressive/Phonological Impairment and Mixed Receptive-Expressive Impairment. First observed scores on the CELF-4 ELI and RLI were used because the participants who received the CELF: PS-2 CLC only received measures of expressive and receptive *vocabulary* (EOWPVT and PPVT-III, respectively) at baseline, as opposed to the more comprehensive expressive and receptive indices of the CELF: PS-2. Since the vocabulary measures alone do not represent the full range of expressive and receptive language functions, these measures were not deemed appropriate to serve as a basis for Impairment categories. The categories of Non-Responder, Responder, and Opposite were also created by the participants' RCI scores. A two-way contingency table analysis using crosstabulation will be conducted to determine whether a significant relationship exists between the categorical variables of Impairment Type and Response Type, with the expectation that children with less pervasive impairment will show a greater proportion of Responders than the group with more pervasive impairment. See Appendix A for an exploratory analysis excluding the participants who took the CELF: PS-2.

Calculation of Reliable Change Index

The RCI was calculated for all standardized measures using the Reliable Change Generator – Version 2.0 (Devilly, 2004). The standard deviations and test-retest reliability coefficients listed in Table 7 were used in the calculations, which are based on the aforementioned procedures outlined by Jacobson and Truax (1991). One RCI was calculated based on the number of participants whose scores were included in the ANOVA for the measure being examined. A second RCI was calculated for each child who had at least two scores that could be compared on the measure being examined. The baseline standard score, or first score available, was compared to the last recorded standard score for each child who participated in the intervention. Children who showed significant change using a 95% confidence interval (z-score cutoff point of ± 1.96) will be placed into a category called *Responders*, indicating a statistically significant response on the standardized measure from baseline to post-intervention testing. Those who did not show significant change are called *Non-Responders*, and those who showed significant decline are called *Opposites*, indicating that their scores changed in the opposite direction of that expected. Since some children received the CELF:PS-2 at the beginning of the intervention and later received the CELF-4 because of their age, some of the RCI scores represent change that occurred over the course of one or two years as opposed to three. This caveat will be discussed further in the Limitations section.

Table 7. Standard Deviations & Test Retest Reliability Coefficients used in Reliable Change Index Calculations

Measure	Standard Deviation	Test-Retest Reliability
Leiter-R Brief IQ Screener[a]	22.9	.92
CELF-4 Core Language Composite	13.6	.92
CELF-4 Expressive Language Index	14	.92
CELF-4 Receptive Language Index	13.6	.89
CELF:PS-2 Core Language Composite	13.6	.91
Arizona Articulation Proficiency Scale-3[b]	8.93	.97
Peabody Picture Vocabulary Test-3[c]	12.75	.92
Expressive One-Word Picture Vocabulary Test	15.72	.90

Note: Unless otherwise noted, all values are based on all ages used in the normative sample for standardization of each measure.
[a]Leiter-R SD and test-retest reliability values represent statistical averages of values provided in the test manual for standardization groups encompassing the age range of this study (2-5, 6-10, and 11-20).
[b]Arizona-3 SD and test-retest reliability values represent statistical averages of values provided in the test manual for standardization groups encompassing the age range of this study (3-4:11, 5-7:11, and 8 or higher).
[c]PPVT-3 SD and test-retest reliability values represent statistical averages of values provided in the test manual for standardization groups encompassing the age range of this study (2:6-5:11, 6-10:11, and 12-17:11).

Power Analysis

Although three years of intervention and data collection had occurred prior to the design of the current data analysis, an a priori power analysis was conducted to determine the theoretical sample size that would be needed to obtain an 80% level of statistical power with F tests (ANOVAs), chi squares (contingency table analyses), and t tests using the GPOWER computer program (Faul & Erdfelder, 1992). All power analyses were conducted using the desired medium effect size (0.25 for F tests, 0.3 for chi square, and 0.5 for t tests) and the .05 alpha level as a criterion for significance. Based on these values, 128 participants would be needed for 80% statistical power with F tests. For chi square tests with two degrees of freedom, as in the current study, a sample size of 108 participants would be needed. For t tests, a sample size of 102 participants would be needed. A post-hoc power analysis was also conducted based on medium effect sizes and the .05 level of significance to determine the level of statistical power of the current analyses. For the ANOVAs conducted in the current study, the current sample size of 20 yields statistical power of 18.5%. The contingency table analysis yields 18.96% statistical power. The *t* tests conducted between Completers (N=13) and Dropouts (N=7) carry 26.8% statistical power. Due to the extremely limited power of the subsequent analyses, all results should be interpreted with great caution.

Aim I Hypotheses and Results

Hypothesis I. a. predicted that participants' scores on the CELF-4 and/or CELF: PS-2 Core Language Composite (CLC) would increase significantly from Baseline to follow-up Time 3, indicating improvement in general language ability. To test this hypothesis, one-way repeated measures ANOVAs and RCI analyses were conducted. Significance of ANOVAs was based on a p value less than .05.

To determine significant change in CELF-4 CLC scores, the repeated measures ANOVA was conducted with the within-subjects factor being number of years participating in the intervention and the dependent variable being standardized scores on the CELF-4 CLC. The ANOVA automatically truncated cell values to include only those subjects who completed the measure at all four time points, resulting in a small N of 10 participants. Neither multivariate nor univariate F tests indicated a significant time effect on the CELF-4 CLC (see Table 9). An RCI analysis conducted on these same 10 participants showed that 2 of these children (20%) were Responders, and 8 of these children (80%)

were Non-Responders. An RCI analysis including all 18 participants who had at least two scores on the CELF-4 CLC over the course of three years of intervention showed that 4 out of 18 participants (22%) were Responders on the CELF-4 CLC during participation in the intervention, and 14 participants (78%) were Non-Responders (see Table 10). The remaining two participants only had one score on the CELF-4 CLC because they received the CELF: PS-2 for the first two years of intervention and left the study after the second year.

To determine change over time in core language skills of preschool-age children, a one-way repeated measures ANOVA was conducted on the CELF: PS-2 CLC, which was only administered at three time points (Baseline, Year 1, and Year 2) because all children who completed the study were older than the CELF: PS-2 normative age ceiling (6 years old) at Year 3. The ANOVA truncated cell values to 6 participants who received the test at all three time points. Neither multivariate nor univariate *F* tests indicated a significant time effect (see Table 9). The RCI conducted on these 6 participants showed that all 6 (100%) were Non-Responders. An RCI conducted on 11 participants who had at least two scores on the CELF: PS-2 showed that 2 out of 11 participants (18%) were Responders, and 9 out of 11 (82%) were Non-Responders (see Table 10). Of the 11 children who took both the CELF-4 CLC and the CELF: PS-2 CLC, 2 were Responders on both measures (18%), 1 was a Responder on the CELF-4 CLC and a Non-Responder on the CELF: PS-2 CLC (9%), 6 were Non-Responders on both measures (55%), and 2 had only one score on the CELF-4 and were Non-Responders on the CELF: PS-2 (18%).

Hypothesis I. a. also predicted that scores on the Expressive Language Index (ELI) and Receptive Language Index (RLI) of the CELF-4 would increase significantly from Baseline to Time 3, suggesting improvement in expressive and receptive language skills, respectively. To test this hypothesis, a one-way repeated measures ANOVA was conducted on each composite score. According to multivariate tests, the ANOVA on the ELI indicated a significant time effect, Wilks' $\Lambda = .15$, $F\ (3, 7) = 13$; $p < .01$, multivariate $\eta^2 = .85$. The multivariate η^2 is an indicator of effect size that ranges in value from 0 to 1 and represents the strength of the relationship between the repeated measures variable (time) and the dependent variable (language test performance) (Green & Salkind, 2005). Univariate tests did not show a significant time effect on the CELF-4 ELI. Bonferroni-corrected pairwise comparisons were not significant, and polynomial contrasts did not show any significant trends in the data over time, although the linear effect approached significance (see Table 9). The RCI analysis including the 10 participants whose scores were analyzed in the ANOVA revealed that 2 participants out of 10 (20%) were Responders, and 8 participants (80%) were Non-Responders. An RCI analysis was also conducted with the 18 participants who had at least two scores on the CELF-4 ELI over the course of three years of intervention; the remaining two participants only had one score on the CELF-4 ELI because they received the CELF: PS-2 for the first two years of intervention and left the study after the second year. The RCI analysis showed that 4 out of the 18 participants (22%) were Responders during participation in the intervention, and 14 participants (78%) were Non-Responders (see Table 10).

A one-way repeated measures ANOVA conducted on 10 participants who received the CELF-4 RLI at all four time points was significant according to multivariate *F* tests, Wilks' Λ = .27, F (3, 7) = 6.23, p = .02, multivariate η^2 = .73 Univariate tests showed a significant effect as well, F (3, 27) = 3.42, p = .03, partial η^2 = .28. Bonferroni-corrected pairwise comparisons showed a significant difference between Year 1 and Year 3, mean difference = 9.6, p = .04. Polynomial contrasts revealed a significant quadratic effect, F (1, 9) = 16.89, $p < .01$, partial η^2 = .65 (see Table 9). The quadratic effect shows means increasing at Baseline, Time 1, and Time 2, with a significant decline from Time 2 to Time 3, visible on the line graph in Figure 1. An RCI analysis conducted on the 10 participants included in the ANOVA yielded no Responders, 7 Non-Responders (70%), and 3 Opposites (30%). An RCI conducted on the 18 participants with at least two scores on the CELF-4 RLI yielded 2 Responders (11%), 13 Non-Responders (72%), and 3 Opposites (17%). See Table 10 for an overview of RCI results.

Based on the above statistical analyses, Hypothesis I. a. appears to be generally unsupported, with the exception of the CELF-4 ELI results, which showed a significant increase over time according to multivariate tests. Univariate tests, however, were not significant. Core language skills did not show significant change over time in children ages 3 to 6 on the CELF: PS-2 or in children age 6 and up on the CELF-4. Receptive language skills on the CELF-4 showed significant decline over time, contrary to hypothesized results. According to RCI analyses, the overwhelming majority of participants showed no response on the CELF language composites over time.

Hypothesis I. b. predicted that participants' scores on the EOWPVT would increase significantly from Baseline to Time 3, indicating improvement in expressive vocabulary skills. One-way repeated measures ANOVA multivariate tests showed no significant time effect, but univariate tests did, F (3, 36) = 3.46, p = .03, partial η^2 = .22. Pairwise comparisons were not significant, and polynomial contrasts did not reveal significant trends in the data, although the quadratic effect approached significance. An RCI conducted with the same 13 participants included in the ANOVA revealed that 3 participants (23%) were Responders, 9 participants (69%) were Non-Responders, and 1 participant (8%) showed decline over time and was thus placed in the Opposite category. An RCI including all 20 participants revealed that 8 participants (40%) were Responders, 11 (55%) were Non-Responders, and 1 participant (5%) showed decline over time and is thus placed in the Opposite category.

It appears that Hypothesis I. b. is generally unsupported due to non-significant multivariate tests of significance. Univariate tests were significant but follow-up pairwise comparisons and polynomial contrasts were not. The RCI analyses showed that a majority of participants did not respond or declined with intervention.

Hypothesis I. c. predicted that participants' scores on the PPVT-III would increase significantly from Baseline to Time 3, indicating improvement in receptive vocabulary skills. One-way repeated measures ANOVA multivariate tests showed a significant time effect, Wilks' Λ = .46, F (3, 10) = 3.94, p = .04, partial η^2 = .54. Univariate tests also showed a significant time effect, F (3, 36) = 5.40, $p < .01$. Bonferroni-corrected pairwise comparisons showed a significant difference between Baseline and Year 2, mean change = 7.08, $p = .02$. An examination of the means over time indicates that PPVT-III scores declined significantly (see Table 8). Polynomial contrasts indicated a significant linear effect, $F(1, 12) = 10.13$, $p < .01$, reaffirming the decline over time (see Table 9). Figure 2 shows the linear trend in data for the PPVT-III. An RCI conducted with the same 13 participants included in the ANOVA revealed no Responders, 7 (54%) Non-Responders, and 6 (46%) Opposites. An RCI including all 20 participants revealed 3 Responders (15%), 10 Non-Responders (50%), and 7 Opposites (35%) on the PPVT-III (see Table 10).

Hypothesis I. c. is clearly not supported by the results, in that scores showed a significant decline over time according to the ANOVA and polynomial contrasts. The RCI analyses showed that only 3 children showed a reliable response over time, with the remainder showing no response or a decline.

Aim II Hypotheses and Results

Hypothesis II. a. predicted that articulation skills would improve over time, as measured by the Arizona Articulation Proficiency Scale-3rd Revision (Arizona-3). A one-way repeated measures ANOVA was conducted on 13 participants with Arizona-3 scores at every time point, and multivariate tests showed no significant time effect, although the *p* value approached significance at .08. Univariate tests showed a significant time effect, $F(3, 36) = 3.5$, $p = .03$. Bonferroni-corrected pairwise comparisons indicated a mean difference of 5 between Baseline and Year 3 that approached significance at $p = .05$. Polynomial contrasts of within-subjects effects indicated a significant linear effect over time, F (1, 12) = 8.74, $p = .01$, partial $\eta^2 = .42$ (see Table 9). Means displayed in Table 8 indicate an increasing linear trend over time. Figure 3 shows the linear trend graphically. An RCI analysis of the 13 participants included in the ANOVA yielded 6 Responders (46%), 6 Non-Responders (46%), and 1 Opposite (8%). An RCI conducted on all 20 participants yielded 12 Responders (60%), 7 Non-Responders (35%), and 1 Opposite (5%). Table 10 presents results of all RCI analyses.

Hypothesis II. a. appears to be marginally supported by the significant univariate F tests and the finding that the majority of participants' reliable change scores showed significant response on the Arizona-3.

Hypothesis II. b. predicted that oromotor skills would improve over time, as measured by the Focal Oromotor Control subtest of the Verbal Motor Production Assessment for Children (VMPAC). The aforementioned limitations in psychometric properties of the VMPAC prohibited appropriate statistical analyses with the ANOVA or RCI. An examination of skill ranges of all 20 participants over time indicated that 2 participants (10%) scored Within Normal Limits at Baseline and remained in this skill range over time, 6 participants (30%) improved by at least one skill range from Baseline to their last observed score, 11 participants (55%) did not improve, and 1 participant (5%) showed a decline. Overall, 40% of participants maintained scores in the highest skill range or improved at least one skill range over time, and 60% of participants showed no change or a decline.

Hypothesis II. b. is not supported, in that the majority of participants either showed no change or declined in skill range from baseline to last observation.

Table 8. Means and Standard Deviations at Each Time Point for Language Measures

Measure	Time Point	N[a]	Mean	Standard Deviation
CELF-4 CLC	Baseline	10	55.90	11.08
	Time 1	10	60.90	14.44
	Time 2	10	59.60	10.22
	Time 3	10	59.70	10.30
CELF:PS-2 CLC[b]	Baseline	6	83.17	17.75
	Time 1	6	83.00	16.99
	Time 2	6	85.50	18.56
CELF-4 ELI	Baseline	10	56.60	10.23
	Time 1	10	60.80	13.58
	Time 2	10	61.30	12.06
	Time 3	10	62.40	8.59
CELF-4 RLI	Baseline	10	72.40	16.90
	Time 1	10	75.90	12.51
	Time 2	10	78.30	9.55
	Time 3	10	66.30	6.38
EOWPVT	Baseline	13	82.54	10.27
	Time 1	13	86.08	10.54
	Time 2	13	88.85	7.89
	Time 3	13	87.77	8.16
PPVT-III	Baseline	13	92.54	7.20
	Time 1	13	90.46	8.01
	Time 2	13	85.46	8.36
	Time 3	13	84.46	8.29
Arizona-3	Baseline	13	76.53	12.87
	Time 1	13	78.08	14.22
	Time 2	13	79.77	11.25
	Time 3	13	81.54	9.84

[a]Cell sizes represent values truncated by ANOVA calculation.
[b]The CELF: PS-2 was only given at Baseline, Time 1, and Time 2 due to the participants' being older than established age norms for the measure at Time 3.

Table 9. Repeated Measures ANOVA Results for Language Measures

Measure	Multivariate Significant?[a]	Univariate Significant?[b]	Pairwise Comparisons[c]	Polynomial Contrasts[d]
CELF-4 CLC	NS	NS	NS	NS
CELF:PS-2 CLC	NS	NS	NS	NS
CELF-4 ELI	Wilks' Λ = .15, F (3, 7) = 13, $p < .01$, multivariate η^2 = .85	NS	NS	NS
CELF-4 RLI	Wilks' Λ = .27, F (3, 7) = 6.23, $p = .02$, multivariate $\eta^2 = .73$	F (3, 27) = 3.42, $p = .03$, partial $\eta^2 = .28$	Year 1 > Year 3, mean difference = 9.6, p = .04	Quadratic effect, F (1, 9) = 16.89, $p < .01$, partial $\eta^2 = .65$
EOWPVT	NS	F (3, 36) = 3.46, $p = .03$, partial $\eta^2 = .22$	NS	NS
PPVT-III	Wilks' Λ = .46, F (3, 10) = 3.94, $p = .04$, multivariate $\eta^2 = .54$	F (3, 36) = 5.40, $p < .01$, partial $\eta^2 = .31$	Baseline > Year 2, mean difference = 7.08, $p = .02$	Linear decline, F (1, 12) = 10.13, $p < .01$, partial $\eta^2 = .46$
Arizona-3	NS	F (3, 36) = 3.5, $p = .03$, partial $\eta^2 = .23$	NS	Linear increase, F (1, 12) = 8.74, $p = .01$, partial $\eta^2 = .42$

[a]Based on multivariate tests of significance: NS = not significant, p value reported if significant (< .05).
[b]Based on univariate tests of significance: NS = not significant, p value reported if significant (< .05).
[c]Bonferroni-corrected pairwise comparisons of estimated marginal means between time points.
[d]Trend analysis for equal time intervals indicates linear or quadratic effect and p value.

Table 10. Response Patterns Based on RCI scores for Language Measures

Measure	N[a]	Responder (%)	Non-Responder (%)	Opposite (%)
CELF-4 CLC	10	2 (20)[b]	8 (80)	---
	18	4 (22)	14 (78)	---
CELF:PS-2 CLC	6	---	6 (100)	---
	11	2 (18)	9 (82)	---
CELF-4 ELI	10	2 (20)	8 (80)	---
	18	4 (22)	14 (78)	---
CELF-4 RLI	10	---	7 (70)	3 (30)
	18	2 (11)	13 (72)	3 (17)
EOWPVT	13	3 (23)	9 (69)	1 (8)
	20	8 (40)	11 (55)	1 (5)
PPVT-III	13	---	7 (54)	6 (46)
	20	3 (15)	10 (50)	7 (35)
Arizona-3	13	6 (46)	6 (46)	1 (8)
	20	12 (60)	7 (35)	1 (5)

[a]First number represents number of participants included in ANOVA; second number represents total number of participants with at least two scores to be compared using RCI calculation.
[b]First number represents number of participants in the category; number in parentheses represents percentage of total participants included in RCI analysis.

Aim III Hypotheses and Results

Hypothesis III predicted that a significant interaction effect would exist on language measures between the variables of Time and Nonverbal IQ group, such that children with a nonverbal IQ of 85 or above would show a greater rate of improvement over time than children with a nonverbal IQ below 85. IQ Groups will be defined as a Leiter International Performance Scale – Revised (Leiter-R) Brief IQ Composite score of Average IQ (≥ 85) or Below Average IQ (<85). A one-way repeated measures ANOVA was conducted for each language measure, using time as the within-subjects factor and IQ Group as the between-subjects factor. No interaction effect was found between Time and IQ Group, and no main effects were found for either variable on the CELF-4 CLC or

CELF: PS-2 CLC. According to multivariate tests, a significant main effect was found for Time on the CELF-4 ELI, F (3, 6) = 15.35, $p < .01$, multivariate $\eta^2 = .89$, but this effect was not demonstrated by univariate tests. No main effect was found for the between-subjects factor of IQ Group, and no interaction effect was found between Time and IQ Group. A significant main effect was found for Time on the CELF-4 RLI according to multivariate tests, F (3, 6) = 9.6, $p = .01$, multivariate $\eta^2 = .83$, as well as univariate tests, F (3, 24) = 3.32, $p = .04$, partial $\eta^2 = .29$. No main effect was found for IQ Group, and the interaction approached significance at $p = .067$. Multivariate tests did not show a main effect for time on the EOWPVT, but univariate tests did, F (3, 33) = 5.41, $p < .01$, partial $\eta^2 = .33$. No main effect was found for IQ Group, and no significant interaction effect was observed on the EOWPVT. Multivariate tests found a significant main effect for Time on the PPVT-III, F (3, 9) = 5.11, $p = .02$, multivariate $\eta^2 = .63$, as did univariate tests, F (3, 33) = 6.69, $p < .01$, partial $\eta^2 = .38$. No significant main effect was observed for IQ Group, and no significant interaction was found. No significant main or interaction effects were found on the Arizona-3.

Hypothesis III is unequivocally unsupported, as none of the hypothesized interactions were significant between IQ Group and Time. While some of the measures showed significant main effects of Time, the between-subjects factor of IQ Group showed no significant main effect.

Aim IV Hypotheses and Results

Hypothesis IV predicted that a significant interaction effect would exist on language measures between the variables of Age Group and Time, such that children in the Young age group would show a greater rate of improvement than the Old age group. Age groups will be defined as Young (3 to 5 years) and Old (6 to 9 years). The dependent variables will be standardized scores on language measures over time. A one-way repeated measures ANOVA was conducted for each language measure, using time as the within-subjects factor and Age Group as the between-subjects factor. No significant main effects were found for Time or Age Group, and no significant interactions were observed for the CELF-4 CLC. The CELF: PS-2 CLC could not be analyzed because all participants who took this measure at baseline were in the Young age group. Multivariate tests showed a significant main effect of Time on the CELF-4 ELI, F (3, 6) = 12.48, $p < .01$, multivariate $\eta^2 = .86$, but univariate tests did not. No significant main effect was found for Age Group, and no interaction effect was found. No interaction effect was found for the CELF-4 RLI. A significant main effect was found for Time according to multivariate tests on the CELF-4 RLI, F (3, 6) = 5.12, $p = .04$, multivariate $\eta^2 = .72$, but this effect was not significant according to univariate tests. No main effect was found for Age Group. While multivariate tests were not significant, a significant main effect for Time was observed on the EOWPVT according to univariate tests, F (3, 33) = 4.40, $p = .01$, partial $\eta^2 = .29$. No significant main effect was found for Age Group, and the interaction between Time and Age Group approached significance at $p = .065$. Multivariate tests were not significant, but univariate tests showed a significant main effect for Time on the PPVT-III, F (3, 33) = 5.23, $p < .01$, partial $\eta^2 = .32$. No significant main effect was found for Age Group, and the interaction between Time and Age Group was not significant. Univariate, but not multivariate, tests showed a significant main effect for Time on the Arizona-3, F (3, 33) = 3.21, p = .04, partial η^2 = .23. No significant main effect for Age Group was found, and the interaction between Time and Age Group was not significant.

No significant interactions were found between the within-subjects factor of Time and the between-subjects factor of Age Group on any of the language measures, meaning that Hypothesis IV is unequivocally unsupported.

Aim V Hypothesis and Results

Hypothesis V predicted that the categorical variables of Impairment Type and Response Type would have a significant relationship with each other on language measures, such that children with Expressive/Phonological Impairment at baseline would show a higher proportion of Responders than children with Mixed Receptive-Expressive Impairment at baseline. A two-way contingency table analysis using crosstabulations was conducted to determine the relationship between categories. The Pearson χ^2 was not significant for the CELF-4 CLC, CELF-4 ELI, CELF: PS-2, Arizona-3, PPVT-III, or EOWPVT. Impairment Type and Response Type were shown to be significantly related on the CELF-4 RLI, Pearson χ^2 (2, N=18) = 6.19, $p < .05$, Cramér's V = .57. The proportions of children in the Expressive/Phonological Group who were Responders, Non-Responders, and Opposites were .14, .43, and .43, respectively. The proportions of children in the Mixed Receptive-Expressive Group who were Responders, Non-Responders, and

Opposites were .09, .91, and .00, respectively. Post-hoc pairwise comparisons were conducted to evaluate the difference between these proportions. Table 10 presents the results of these analyses. The Holm's sequential Bonferroni procedure was used to control for Type I error at the .05 level across comparisons. The only pairwise comparison that was significant following the Bonferroni procedure was between the Non-Responders and Opposites, Pearson $\chi^2(1, N=16) = 6.15$, $p = .01$, Cramér's $V = .62$. The probability of a child's being in the Mixed Receptive-Expressive Group was significantly higher for Non-Responders than for Opposites. Figure 4 presents a bar graph of these results. It should be noted that 67% of cells in this 2 x 3 contingency table have an expected frequency of less than 5 participants. Green and Salkind (2005) caution that the validity of results should be called into question when a table of this size has expected frequencies of less than 5 in more than 20% of cells.

Hypothesis V is largely unsupported in that children with Expressive/Phonological Impairment did not show significantly greater proportions of Responders on any measure. In fact, pairwise comparisons on the only significant chi square for the CELF-4 RLI showed a significant difference in proportions for Impairment category between Non-Responders and Opposites.

Table 11. CELF-4 RLI Results for Crosstabulation of Impairment Type and Response Type

			Response Type			
			Non-Responders	Responders	Opposites	Total
Impairment Type	Exp/Phon. Impairment	Count	3	1	3	7
		Expected Count	5.1	.8	1.2	7.0
		% within Exp/Phon. vs. Mixed	42.9%	14.3%	42.9%	100.0%
		Residual	-2.1	.2	1.8	
	Mixed Rec-Exp. Impairment	Count	10	1	0	11
		Expected Count	7.9	1.2	1.8	11.0
		% within Exp/Phon. vs. Mixed	90.9%	9.1%	.0%	100.0%
		Residual	2.1	-.2	-1.8	
	Total	Count	13	2	3	18
		Expected Count	13.0	2.0	3.0	18.0
		% within Exp/Phon. vs. Mixed	72.2%	11.1%	16.7%	100.0%

Aim VI Hypothesis and Results

Hypothesis VI predicted that children who exited the study (Dropouts) would show significantly higher scores than remaining children (Completers) on the Leiter-R, CELF-4, CELF: PS-2, EOWPVT, PPVT-III, Arizona-3, and VMPAC Focal Oromotor Control subtest at Baseline, suggesting that they entered the study with significantly higher ability and were able to progress and exit the intervention. As previously mentioned, the VMPAC Focal Oromotor Control subtest was excluded from the analyses because standard scores were not available. A series of independent-samples t-tests was conducted to determine significant differences between group means for Completers and

Dropouts. T-tests were not significant for the Leiter-R or any of the language measures once Bonferroni corrected. Therefore, Hypothesis VI is unsupported due to the fact that no significant differences were found in group means for Completers vs. Dropouts.

An examination of skill ranges of the entire sample at the last observed time point is presented in the following table to characterize level of functioning at *outcome* according to each measure.

Table 12. Skill Ranges for All Participants at Exit

	Participant	CELF-4 CLC	CELF-4 ELI	CELF-4 RLI	CELF:PS2 CLC	Arizona3	PPVT3	EOWPVT	VMPAC FOC %	LeiterR
Dropouts	1	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	WNL	Avg.
	2	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Sev.	Avg.
	3	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	WNL	Avg.
	4	Ext. Low	Ext. Low	Ext. Low	---	Ext. Low	Bel. Avg.	Ext. Low	Sev.	Ext. Low
	5	Ext. Low	Ext. Low	Bdl.	Ext. Low	Ext. Low	Avg.	Avg.	Sev.	Bdl.
	6	Avg.	Avg.	Avg.	Avg.	Ext. Low	Avg.	Avg.	WNL	Sup.
	7	Avg.	Avg.	Avg.	---	Bdl.	Avg.	Avg.	Sev.	Avg.
Completers	8	Bdl.	Bdl.	Ext. Low	---	Avg.	Bdl.	Avg.	Sev.	Ext. Low
	9	Ext. Low	Ext. Low	Ext. Low	---	Bdl.	Avg.	Bdl.	Sev.	Bdl.
	10	Ext. Low	Ext. Low	Ext. Low	---	Bdl.	Bdl.	Bel. Avg.	Sev.	Ext. Low
	11	Ext. Low	Ext. Low	Bdl.	---	Avg.	Bdl.	Avg.	Sev.	Avg.
	12	Ext. Low	Ext. Low	Ext. Low	---	Avg.	Avg.	Avg.	Sev.	Bel. Avg.
	13	Ext. Low	Bdl.	Bdl.	---	Bdl.	Bdl.	Bel. Avg.	Mild	Avg.
	14	Ext. Low	Ext. Low	Avg.	---	Bdl.	Avg.	Avg.	WNL	Avg.
	15	Ext. Low	Ext. Low	Ext. Low	---	Avg.	Avg.	Avg.	Sev.	Bdl.
	16	Ext. Low	Ext. Low	Ext. Low	---	Avg.	Avg.	Avg.	WNL	Ext. Low
	17	Ext. Low	Ext. Low	Avg.	---	Ext. Low	Avg.	Avg.	WNL	Avg.
	18	Bdl.	Bdl.	Bdl.	---	Avg.	Avg.	Avg.	WNL	Avg.
	19	Ext. Low	Ext. Low	Ext. Low	---	Ext. Low	Bel. Avg.	Bdl.	Sev.	Ext. Low
	20	Ext. Low	Ext. Low	Ext. Low	---	Avg.	Avg.	Avg.	Sev.	Avg.

Note: All skill ranges based on last observed standard scores except VMPAC ranges, which are based on percentage scores. Standard score range key: Sup.=Superior (120-129), Ab.Avg. = Above Average (116-119), Avg. = Average (85-115), Bel.Avg. = Below Average (80-84), Bdl. = Borderline (70-79), Ext.Low = Extemely Low (<70). VMPAC skill ranges based on those specified in the test manual. Percentage score range key: WNL=Within Normal Limits, Mild = Mild deficit, Mod = Moderate deficit, Sev = Severe deficit.

Table 13. Overall Hypotheses & Results

Hypothesis	Results
Hypothesis I.a. – CELF-4 & CELF:PS-2 scores will improve	• Generally unsupported, with exception of expressive language improvement • Core language: No change • Receptive language: Decline • Majority showed no response
Hypothesis II.b. – EOWPVT scores will improve	• Generally unsupported in that majority showed no response
Hypothesis II.c. – PPVT-III scores will improve	• Not supported – significant decline in scores
Hypothesis II.a. – Arizona-3 scores will improve	• Marginally supported in that higher proportion of total sample showed positive response
Hypothesis II.b. – VMPAC Focal Oromotor Control scores will improve	• Not supported – majority showed no change or declined
Hypothesis III – Interaction NVIQ x Time	• Not supported – no significant interactions
Hypothesis IV – Interaction Age x Time	• Not supported – no significant interactions
Hypothesis V – Relationship Impairment x Response, higher % of Responders for less pervasive Impairment	• Not supported – no significantly higher proportions of Responders in Expressive/Phonological Group on any measure
Hypothesis VI – Dropouts will show higher scores than Completers at Baseline	• Not supported – no significant differences once Bonferroni-corrected

CHAPTER SEVEN
Discussion

The current study was designed to evaluate effectiveness of the Shelton EI program in the areas of expressive and receptive language and vocabulary, articulation, and oromotor skills. Based on literature pointing to various predictors of improvement with intervention, nonverbal IQ, age, and pervasiveness of impairment were also examined as factors potentially associated with outcome. Participants were 20 children with LI who were enrolled in the Shelton School Early Intervention (EI) Program, a specialized curriculum designed to improve language, motor, and sensory-integration skills. Thirteen children completed three years of intervention in the program – six children left the study after two years, and one child left the study after one year.

The language, articulation, and vocabulary measures included in this study were chosen from a more comprehensive battery of 26 measures of language, phonological skills, perceptual skills, academics, short-term memory, motor skills, and behavioral/social/emotional functioning. The CELF-4 and CELF: PS-2 were selected based on their established status in the speech-language field as a comprehensive assessment of overall language skills (Core Language Composite), as well as expressive and receptive domains of language. The EOWPVT and PPVT-III were chosen for their straightforward and quick assessment of expressive vocabulary and receptive vocabulary, respectively. The Arizona-3 was chosen based on its established history as a reliable and valid measure of articulation skills and the fact that its norms encompassed the age range of the participants from baseline to outcome. While the Leiter-R Brief IQ Composite is considered an *estimate* of Nonverbal IQ in that only a select number of representative subtests from the subtest battery are given in the interest of time, it was selected due to the fact that it is the only nonverbal IQ measure with purely nonverbal instructions that encompasses the age range of the current study. The VMPAC Focal Oromotor Control subtest was chosen based on its measurement of mandibular, facial, labial, and lingual skills implicated in apraxia of speech, a deficit that is closely tied to articulation and was noted in several of the children in the current study. Although its standardization methods were shown to be lacking, it provides some index of the oromotor skill base necessary for producing language.

Core, Expressive, and Receptive Language Skills

The hypothesis that participants' scores on the CELF-4 CLC, CELF: PS-2 CLC, and CELF-4 ELI and RLI would improve over time was generally unsupported by the data, with the exception of one significant increase in the CELF-4 ELI based on multivariate tests. It seems likely that the severely limited statistical power of analyses due to small sample size greatly affected the chance of finding significant changes. Hence, the RCI was calculated to determine clinically significant change from pre- to post-intervention. The RCI generally echoed the overall findings of the ANOVAs in that the majority of participants, ranging from 70 to 100% across measures, showed no significant change while enrolled in the intervention program. Those children that did show reliable improvement made up 11 to 22% of the sample across measures. However, the lack of a control or comparison group renders these findings of improvement negligible in terms of attributing change to the EI Program itself. Disconcertingly, 17 to 30% of children showed a decline in receptive language skills, which was unexpected given the heavy emphasis on language development in the intervention.

One possible explanation for the discrepancy between expressive and receptive change over time may be that the language skills emphasized in the DuBard Association Method and Montessori techniques were mainly expressive and articulatory as opposed to receptive in nature. One of the tenets of the DuBard Association Method is that precise articulation is required, and phoneme production is emphasized in visual presentation of words. These skills may have been more prevalent in lessons than sentence comprehension or understanding of word meanings. In addition, the fact that the tests administered were age-normed using a non-impaired population of children begs the question of whether children with LI are actually making smaller improvements in language that cause their scores to appear static over time. Theoretically, the gains they are making may not be reflected as an increase in standard scores, but rather in the stability of their standard scores over time as compared to typically-developing children of the same age. In the same vein, a child with typical language development may not be expected to show a significant increase in language standard scores over time, either. The maintenance of language scores in the average range over time may be viewed as perfectly acceptable in the eyes of the child's parents and teachers. The observed decline in receptive language scores, on the other hand, may represent a stall or plateau in learning that serves to increase the discrepancy between language-impaired and typically developing children and translates into declining standard scores.

Expressive and Receptive Vocabulary

Findings for the EOWPVT and PPVT-III (expressive and receptive vocabulary measures) were very similar to findings for the Expressive and Receptive Language Indices of the CELF-4 in that expressive vocabulary skills were shown to increase significantly according to one set of *F* tests (univariate), whereas receptive vocabulary skills showed significant decline over time. Once again, the above explanations apply to account for the discrepancy between expressive and receptive vocabulary progress. What is more, children have been shown to score differently on expressive and receptive vocabulary measures depending on the administration order (Llorente, Sines, Rozelle, Turcich, & Casatta, 2000). Using the EOWPVT-Revised (expressive vocabulary) and the PPVT-Revised (receptive vocabulary), children who received the receptive measure first performed better on the expressive measure because of administration procedures: the examiner provides words on the receptive measure that prime the child to produce more words on the expressive measure. Examiners in the current study did not keep a record of the administration order of the EOWPVT and PPVT-III for each child, a factor that

Articulation

Results for the articulation measure showed a significant increase in articulation skills over time according to univariate tests, and the RCI analyses that included all participants showed that the majority of children made reliable improvement. It is possible that articulation skills are simply less difficult to remediate than other language skills. Another potential explanation for this finding is the heavy emphasis on precise articulation skill of the DuBard Association Method. Reinforcement of precise articulation and pronunciation according to phonemic visual representations of words using color and special symbols is essential to this method of language intervention (DuBard & Martin, 2000). However, it must be noted that the majority of children in the EI Program received some type of outside speech or language therapy before or during the study, either from Shelton's Speech, Language, and Hearing Clinic or a private SLP provided by the child's parents. Children received these services to help them continue to make progress in language skills, and it is impossible to determine whether any improvements observed in the current data were due to the Shelton EI Program or another language intervention. Additionally, the majority of speech-language measures were administered by an SLP who was also one of the child's classroom instructors. This confound introduces examiner bias, in that the test administrators were not blind to the goals of the study and likely wanted the children that they were teaching to succeed. Even though the examiners were thoroughly trained in standardized testing procedures and presumably made efforts to remain objective, they may have inadvertently affected testing outcomes through slight alterations in their behaviors during testing.

Oromotor Skills

While a truly quantitative analysis could not reasonably be conducted on raw or percentage scores from the VMPAC Focal Oromotor Control subtest, a qualitative description of change in oromotor skill ranges showed that the majority of participants did not show any change from Baseline to their last observed score. Notably, however, 40% of participants either maintained scores in the expected range or did show improvement. The DuBard Association Method focuses on precise articulation and remediation of oromotor skills necessary to produce speech accurately, but its impact was apparently variable across the sample. It is possible that three years of the EI Program was not sufficient to produce changes in the entire sample, and children with more severely impacted oromotor skills may need a longer period of intervention to show change or may have reached a threshold in development of oromotor skills. In all likelihood, children in the study with verbal dyspraxia who received outside intervention were identified by their individual SLPs to have deficits in focal oromotor control and received remediation for these skills, but information on specific duration, timing, and content of outside intervention for each child was not available. It should also be noted that the questionable test-retest reliability of the VMPAC, which was not reported to reach significance in the test manual, may have played a role in skill range classifications. Several children jumped from the Severe to the Within Normal Limits range and back again from year to year, and these ranges were only differentiated by a few percentage score points.

Impact of Nonverbal IQ

The current data showed no significant impact of Nonverbal IQ on improvement over time. Once again, it is very likely that the small sample size, with some between-subjects cell values truncated to only 3 participants for the higher IQ group, greatly limited statistical power and the chance of finding any significant interaction effects. The repeated measures ANOVAs only included 10 participants for the CELF-4 Composites, 6 participants for the CELF: PS-2 CLC, and 13 participants for each of the other language measures. Although the initial sample size of 20 was indeed small, attrition probably negatively affected the chance of finding significant results that much more. In addition, Nonverbal IQ Group was used in the ANOVA as a dichotomous categorical variable. It is possible that a regression model would have been more sensitive to effects of IQ on language because regression treats variables as continuous, so equal cell sizes are unnecessary. Regression was not used in the current study because the overall language outcome variable, the CELF CLC, was correlated with other composites from the same measure, and three out of its four subtests measured expressive skills. Therefore, it was not deemed appropriate as an outcome variable representative of overall progress. Those limitations aside, the findings of no significant interaction effects are not consistent with reviewed studies showing that at least a proportion of variance in language scores could be explained by initial nonverbal IQ (Bishop & Edmundson, 1987; Botting et al., 2001; Oliver et al., 2004; Schery, 1985; Thurm et al., 2007). Then again, most of these studies also showed that language tests were predictors of outcome, and language variables often outweighed the predictive value of nonverbal IQ. It is possible that nonverbal IQ and language ability simply load onto the common factor of general intellectual ability, which may account for most of the variance in outcome. Depending on the sample size, which tests were used to measure the constructs of nonverbal and language ability, and which type of statistical analyses were used (i.e. ANOVA vs. regression, intention-to-treat analyses vs. not), slight variations in results may be found from study to study. Boyle and colleagues (2007), who used intention-to-treat procedures and ANCOVA methods, for example, found nonverbal IQ, as measured by the Wechsler Abbreviated Scale of Intelligence (WASI), not to be a moderating variable in language outcome. However, the WASI uses verbal instructions, as opposed to Raven's Colored Matrices, which uses both verbal and nonverbal instructions and was used in the study by Botting and colleagues (2001). The extent to which the nonverbal IQ measure taps into verbal comprehension ability could affect how much variance is attributed to the constructs of nonverbal IQ and language ability in statistical analyses. As previously mentioned, DeThorne and Watkins (2006) sum up the relationship between nonverbal IQ and language as highly variable due to inconsistency in measures across studies. Unfortunately, this type of inconsistency is prevalent across the language literature does not make for a straightforward explanation of the current results.

Impact of Age

Age was not found to impact language outcomes in the current analyses. This Age by Time analysis was subject to the same pitfalls as the IQ by Time analysis: limited statistical power and dichotomous categories instead of continuous variables. It is likely that these factors accounted for the lack of significance, in that these findings are inconsistent with those of Jacoby and colleagues (2002), who reported that younger children with LI generally required fewer units (15-minute increments) of speech-language therapy to show significant improvement in functional communication than did children who started receiving intervention at an older age. A study by Schery (1985) also found age to be correlated with language outcomes. It seems that while the literature reports a relationship between age and responsiveness to intervention, the current study did not reflect the same findings.

Relationship Between Impairment and Response

The current findings did not support a relationship between pervasiveness of LI and response to intervention, with the exception of the CELF-4 RLI, in which children who did not respond to intervention showed a higher probability of having Mixed Receptive-Expressive impairment than those who showed significant decline over time. These findings are inconsistent with past research showing that children with more pervasive impairment across expressive and receptive domains are less amenable to intervention, and those with phonological problems only fare best (Boyle et al., 2007; Jacoby et al., 2002; Johnson et al., 1999; Law et al., 2004; Young et al., 2002). Receptive LI virtually never occurs without expressive impairment because expressive skills rely on the development of receptive skills (American Psychiatric Association, 2000), so receptive problems may always be considered a signal for a more pervasive disorder than expressive or articulation problems alone. Since baseline impairment status was not shown to have a relationship to response, it is likely that the small, heterogeneous sample and consequential lack of statistical power negatively affected the chances of finding a significant result.

Comparison of Completers to Dropouts

No significant differences were found between Completers and Dropouts at baseline in the current study on any measure. This result is likely due to the fact that the 7 children who left the study did so for a variety of reasons, not simply because all of them improved to the point of no longer needing intervention. Of these students, three of them remain at the Shelton School and receive some type of language intervention. One student transitioned to another classroom for children with less severe language difficulties for the last year of the EI program because her instructors found that her skills had progressed beyond that of her classmates in EI. Another student transitioned to a different classroom for children with less severe language problems but continued to receive language therapy from the Intermediate EI classroom SLP, who was trained to address this student's language problems. One child remains at Shelton and is now in a regular DuBard Association Method class, which is tailored to children with less severe language problems than those in EI. Of the children who no longer attend Shelton, one child who had low cognitive ability became increasingly emotionally distressed by the demands of the Intermediate EI classroom and displayed problematic behavior in the classroom, which interfered with his/her own learning and that of other students. His/her parents and instructors agreed that he/she should transition to a special education program in a public school. Another child's parent decided to move the child to another school for personal reasons. This particular child had medical needs, low cognitive ability, and little to no verbal expression ability upon exiting the study. Another child with severe articulation problems demonstrated strong language skills and cognitive ability, and his/her parents enrolled him/her in another private school that his/her siblings were attending. Finally, one child who demonstrated substantial progress and improved skills was enrolled in a public school by his/her parents for financial reasons. Since none of these children were considered to have fully remediated language skills when they left the study, there is not much reason to expect that their baseline scores would be significantly different than those of children who remained in the study. The initial assumption behind the hypothesis that dropouts would be shown to perform at higher levels than completers at baseline was that dropouts left the study because they began with higher ability and progressed to the point of no longer needing intervention. This supposition turned out not to be correct once information about the status of each child's language skills at the time they left the study was discovered upon school file review. In actuality, the fact that completers' baseline scores were not significantly different from those of dropouts shows that attrition did not affect the extent to which the reduced sample of 13 was representative of the original sample of 20 children. It is interesting to note, however, that while no differences were discovered at baseline, five of the seven children who left the study did show language skills and nonverbal cognitive ability in the average range on nearly every measure at *outcome*, while few of the children who completed the study progressed to the average range. Therefore, future studies should aim to determine which factors at baseline predict faster rates of improvement over time.

Conclusions

A limited number of conclusions can be drawn from this study of an intervention designed to improve language skills in 20 children with LI. Based on one-way repeated measures ANOVA analyses, children's overall expressive language and expressive vocabulary skills showed significant improvement over time according to multivariate or univariate tests, although the exact cause for this improvement cannot be determined due to a host of confounding variables and lack of a control group. Conversely, overall receptive skills and receptive vocabulary skills showed a significant decline over time according to both univariate and multivariate tests. Articulation skills showed significant improvement according to univariate tests. According to RCI analyses, relatively small proportions of the sample improved on language measures, with the exception of the expressive vocabulary measure, on which nearly half of the sample improved, and the articulation test, in which a majority of the sample showed reliable improvement. Most children showed no response to language measures, and few children showed decline, with the most children showing decline on the receptive language and receptive vocabulary measures. No significant relationships were found between nonverbal IQ and language improvement or age and language improvement. The only significant relationship between impairment type and response was revealed on the receptive language measure, in which children who showed no reliable response to intervention had a higher likelihood of having Mixed Receptive-Expressive impairment than those children who showed reliable decline. In other words, all children who declined on this measure had Expressive/Phonological impairment at baseline, and the number of Non-

Responders was high for those children with Mixed Receptive-Expressive impairment. Completers and Dropouts were also shown not to differ significantly at baseline, ensuring that the group of children who completed the study was representative of the original sample of 20 children on all measures.

Of these limited findings, the discrepancy between improvement in expressive/articulation measures and receptive measures warrants further investigation. While past research has shown baseline impairment status to predict response over time, this was not the case here. Improvement on expressive/articulation measures was not isolated to those children with less pervasive impairment, and decline on receptive measures was not isolated to those with more pervasive impairment. Future studies should therefore examine the differences in prognosis for these particular skills regardless of initial type of impairment, and interventions may need to be modified to place more emphasis on remediation of receptive skills, if possible.

Limitations

Several methodological limitations hinder the ability to draw meaningful conclusions from study results. First and foremost, no definitive statement about the impact of the intervention itself can be made due to the lack of a matched comparison or control group. Without a basis for comparison, any change observed in children's scores could be due to confounding factors that have nothing to do with the Shelton EI Program. Several potentially confounding variables can be enumerated. The unique structure of the

Shelton School environment, which is a supportive culture in which faculty exhaust all resources available to meet the individual child's needs, may have been a factor that influenced outcome. This environment was not isolated to the EI Program, but rather it extended throughout the school's programs and activities. The EI classroom environments themselves consisted of a mixture of ages, and older children likely had a positive influence on younger children, which was a feature not specific to the language intervention itself. Parental involvement may also have affected results, in that parents of these children likely helped them with homework, offered constant support and encouragement, and focused more on helping with language difficulties than the typical parent would. The closely related factor of parental selection may also have contributed to performance; parents who choose Shelton over another school may possess certain characteristics that distinguish them from other parents, i.e., socioeconomic status, education level, and learning disability status. These factors play an important role in each child's upbringing and home environment, which may influence progress over time. Medication changes likely influenced performance on testing to a certain extent. At least nine of the 20 children were prescribed some type of medication during the study for problems related to ADHD, mood or anxiety symptoms, or comorbid medical conditions for unspecified amounts of time. Medications included Ritalin, Concerta, Focalin, Focalin XR, Strattera, and Adderall for ADHD symptoms; imipramine for symptoms of obsessive-compulsive disorder; Lexapro, Prozac, and Remeron for depressive or anxiety symptoms, Phenobarbital and Trileptal for epileptic seizures, and insulin to control Type I diabetes. One child was also born with Joubert syndrome, a genetic disorder affecting balance, motor coordination, and other cognitive abilities. Other comorbid diagnoses included Reading Disorder (dyslexia), Developmental Coordination Disorder, auditory processing deficits, severe visual perception difficulties, oculomotor dysfunction, and mild mental retardation. Medication issues and associated deficits in functioning likely influenced results to an extent. All children involved in the study also had some type of outside speech-language intervention, either before the study began or during the study, and the type, frequency, intensity, and duration of therapy varied from child to child. This variable likely had a significant impact on the results. The aforementioned involvement of teachers as examiners is also a confounding variable because examiner bias may have influenced the child's performance. Different examiners also administered different subtests, so any type of subjectivity in scoring without an inter-rater reliability analysis could have introduced measurement error. Any combination of these variables may have been responsible for improvement or deterioration observed. Therefore, significant changes in performance cannot be attributed unequivocally to the intervention. The RCI was calculated to compensate for the lack of a control group in terms of determining reliably significant change in each individual before and after intervention, but the RCI does not account for confounding variables that may have contributed to variance in test scores. The RCI also represented change resulting from different lengths of intervention, as the last observed score was used as a comparison to baseline scores.

Another formidable limitation to the current study was the lack of statistical power. The small sample size of 20 children and attrition of 7 children severely limited statistical power of the analyses and the chance of finding any significant effects. The power analyses conducted using the GPOWER program demonstrated that a sample of at least

102 children would have been necessary to achieve acceptable 80% statistical power, and the current analyses achieved at most a quarter of the acceptable degree of power. Therefore, it may not be appropriate to draw meaningful conclusions about the current findings or apply them to the language-impaired child population without further research. It would be presumptuous and unethical to assume that the small sample of 20 children, which was reduced to as few as 6 participants for some analyses, was representative of the population being studied. The analyses in which significance was demonstrated should be interpreted within the context of the heterogeneous sample and wealth of potentially confounding variables.

The measurements chosen to monitor progress over time may also have limited interpretability. The CELF language measures, while they possess solid reliability and validity and are widely used in the field, did not encompass the broad age range of participants, which of course changed over time as participants grew older. Both the CELF-4 CLC and the CELF: PS-2 CLC were given to participants in the overlapping age range (5 to 6 years old). Although the CELF: PS-2 consists of less difficult items, a 6-year-old child should theoretically obtain a standard score within the same confidence interval on this measure as she would on the CELF-4, which consists of more difficult items, because of norming procedures that account for age. Appendix A presents results of an exploratory paired-samples *t* test that was conducted on the 7 participants who received both measures at baseline to determine whether a significant difference would be found between baseline scores. Psychometric properties of the measures used may also have played a role in the scores obtained. Evaluation of the psychometric properties of the Leiter-R indicates that the 8, 10, and 11-20-year-old age groups in the normative sample included fewer than the suggested 100 subjects needed to minimize error (DeThorne & Schaefer, 2004). The aforementioned problems with VMPAC standardization also prohibited statistical analyses, and administration order of the EOWPVT and PPVT-III may have affected performance on the expressive measure.

External validity of the study is hindered by the fact that the children chosen made up a convenience sample and were not selected randomly for intervention from a larger group of language-impaired children. This sample was not representative of various ethnicities because the majority of children were Caucasian, so results should be interpreted with caution when applied to samples with a higher proportion of other ethnicities. The EI Program was also highly specialized because it involved the MACAR educational structure and the DuBard Association Method, and both an instructor and an SLP were always present. Neither of the intervention methods has been researched with empirical, randomized, controlled studies, and no other programs that integrate both types of language intervention are known to exist. Therefore, results based on the unique and singular structure of the intervention cannot be generalized to other schools that do not have the same type of educational structure. It is also impossible to attribute any of the effects observed to one intervention component over another, or to the combination of the two.

Future Directions

Future research conducted on the DuBard Association Method and MACAR educational structure should address several limitations that could not be overcome in the current study. A much larger sample of at least the number of subjects suggested by power analyses would be necessary to increase the extent to which the sample is representative of the population and increase the chance of finding significant effects over time. A randomized, controlled trial should also be conducted so that any change seen in scores can be attributed to the intervention itself. While a true control group of children with LI who are not receiving intervention may not be ethically advisable, a comparison group could consist of language-impaired children matched for gender, ethnicity, socioeconomic status, and type of impairment (Expressive, Mixed Receptive-Expressive, and Phonological Disorder) who are receiving treatment-as-usual (TAU), or the regular standard of care. The TAU group may receive a pre-determined optimal amount of traditional speech-language therapy units provided by a school SLP. The DuBard Association Method could then be compared to traditional speech-language therapy, and any significant differences in performance or rate of improvement between the two groups could be attributed to the DuBard intervention itself. Intervention components (DuBard Association Method and MACAR) should also be separated to determine which intervention produces the greatest improvement. Conducting this study at multiple sites would also improve external validity by ensuring that the Shelton School environment itself is not responsible for changes seen over time. Other private and public schools could be included that administer the intervention and TAU to separate groups.

Threats to internal validity should be considered in sample selection, although the inherent nature of LI is that it presents a different picture of deficits from child to child. Ideally, concomitant nonverbal cognitive deficits, mood disorders, and attention deficits would be exclusionary in future research, as these variables increase the heterogeneity of the sample, affect testing results, and decrease the probability of finding the "real" effect of language intervention. Outside intervention should also be controlled, which may be done more easily with very young children who have not yet begun receiving language services. The age range of the sample should also be narrowed to decrease variation in the sample, and the confounding variable of natural maturation in language skills should be controlled.

The selection of measures should also be modified to include language measures that encompass the entire age range of the sample for the duration of the study, if possible. This change would overcome problems with comparing measures over time. Measures with strong reliability and validity should certainly be chosen, and the test-retest reliability coefficients in particular should be significant, since measures will be administered at several time points. Test-retest reliability may also be improved if the same examiners give the same tests to every participant in order to minimize possible subjectivity in scoring across examiners. It would also be important to consider the fact that the astounding variation of language measures used in LI research has contributed to general inconsistency in findings and very narrow conclusions about isolated areas of functioning, making it difficult to understand the central points about the nature of the disorder and the impact of intervention. Therefore, future studies should limit the number of measures administered to one comprehensive test that is the “gold-standard” in each skill area to be studied. This modification would not only reduce exhaustion and time commitment for the child and the examiner, but it would reduce the Type I error rate in statistical analyses and increase the likelihood that the research could be compared or meta-analyzed with other studies in the field, contributing to the research base.

Current findings should be considered in future research as well. The discrepancy between response found on expressive measures versus receptive measures should be investigated more fully with a larger sample. Interventions should be held to rigorous standards for evidence-based practice in future treatment of children with LI in order to save unnecessary burden on children and instructors in terms of time, energy, and resources required for intervention and testing.

Practical Implications

Due to the host of limitations involved with interpretation of the current results, the clinical implications of these findings are warranted to determine what contribution this study makes to the literature. While the majority of the sample did not show improvement overall, some reliable improvement was demonstrated in expressive and articulation skills. Whether this improvement was attributable to the current intervention or not, the fact that a noteworthy proportion of children improved (i.e., up to 60% in articulation skill) is encouraging and calls for further investigation of the factors responsible for this improvement. Although some children declined on receptive measures, this proportion was small, and it would seem that a lack of response on standardized measures is more desirable than widespread decline. Standardized measures may not capture more subtle improvements made by this particular clinical population. Certainly, without an appropriate control group, no definitive statements may be made about the current intervention itself, but trends in the data with this group of children is informative to the Shelton School and may be indicative of similar trends in small, heterogeneous groups of children at other schools who receive school-based intervention services, may see a psychiatrist for ADHD medications, and may receive outside intervention.

If nothing else, other school educators and clinicians may gain awareness that years of intervention may be needed to observe small improvements in children with LI, some areas may show decline over time according to tests normed on typically developing children, and several limitations to objective study design are difficult to overcome in a school setting in which parents want every resource possible made available to their children. Separate but equal treatment groups may not always be possible to create with such a small number of children who have LI in a given school, and funding and cooperation for a multiple-site research study between schools may be difficult to come by. However, with the push for evidence-based practice of language interventions (e.g., (Cirrin & Gillam, 2008), building the evidence base remains an important task. A reconfiguration of objectives may be necessary within school systems to address the need for appropriate research analysis to back up clinical approaches. Ultimately, children with LI are a heterogeneous group and their clinical needs for individualized language goals and intervention may overshadow the possibility for rigorously-designed research in the school setting.

APPENDIX A
Exploratory Analyses

Two exploratory analyses are warranted due to the fact that some children received the CELF-4 while others received the CELF: PS-2 at different time points depending on their age.

First, a paired-samples *t* test was conducted on the seven participants who received both CELF measures at baseline to determine whether they differed on baseline scores for the two measures. A significant difference was indeed found between the mean for the CELF-4 CLC (M = 57.57, SD = 13.76) and the mean for the CELF: PS-2 CLC (M = 71.71, SD = 7.45), $t(6) = 3.52$, $p = .01$. While both sets of scores were included in prior analyses, this exploratory analysis reveals that choosing between measures may have affected overall ANOVA results. This caveat should be taken into consideration when interpreting results for these measures.

Second, the two-way contingency table analysis reported above was based on two categories: Response Type, formed by classification of Responder, Non-Responder, and Opposite based on RCI analyses, and Impairment Type, formed by children's first observable language score on the CELF-4 ELI and RLI and baseline score on the Arizona-3. Not all of the CELF-4 composite scores represented *baseline* scores because five children only received the CELF: PS-2 CLC at baseline since they were too young to receive the CELF-4. Unfortunately, the CELF: PS-2 ELI and RLI were not administered to these five children at baseline, limiting available measures of expressive and receptive language skills at baseline to their expressive and receptive *vocabulary* scores on the EOWPVT and PPVT-III, respectively. Since these scores do not encompass a full range of expressive and receptive skills, they were deemed inappropriate to represent these skill domains in the five children who received the CELF: PS-2. Therefore, additional contingency table analyses were conducted just for the 15 participants who received the CELF-4 at baseline and had at least two comparable data points on each measure to determine response category using the RCI. The goal of this analysis was to determine whether using the CELF-4 ELI and RLI at subsequent time points to determine Impairment category affected earlier results. Of course, the resultant smaller sample size will affect validity of the exploratory contingency table analysis even more. Two of the contingency table analyses were found to be significant. The CELF-4 RLI chi square was again found significant, Pearson χ^2 (1, N=15) = 7.5, $p < .01$, Cramér's $V = .71$, and removing participants who only received the CELF: PS-2 at baseline consequently removed all Responders on this measure. Thus the 2 x 2 table comparing Non-Responders to Opposites again revealed that Non-Responders had a higher chance of having Mixed Receptive-Expressive impairment. Table 14 presents contingency table results for the CELF-4 RLI.

Table 14. CELF-4 RLI Results for Exploratory Crosstabulation of Impairment Type and Response Type

			Response Type		
			Non-Responders	Opposites	Total
Impairment Type	Exp/Phon. Impairment	Count	2	3	5
		Expected Count	4.0	1.0	5.0
		% within Exp/Phon. vs. Mixed	40.0%	60.0%	100.0%
	Mixed Rec-Exp. Impairment	Count	10	0	10
		Expected Count	8.0	2.0	10.0
		% within Exp/Phon. vs. Mixed	100.0%	.0%	100.0%
	Total	Count	12	3	15
		Expected Count	12.0	3.0	15.0
		% within Exp/Phon. vs. Mixed	80.0%	20.0%	100.0%

The chi square for the CELF-4 ELI was also found significant in contrast to previously reported results, Pearson χ^2 (1, N=15) = 6.52, p = .01, Cramér's V = .66. Results of the 2 x 2 analysis (there were no Opposites on this measure) compared Responders to Non-Responders and showed that Non-Responders had a higher probability of having Mixed Receptive-Expressive impairment than Responders. This finding is presumably different from the previous lack of significance because the number of Non-Responders with Expressive/Phonological impairment was reduced by three, making the discrepancy between this cell and the others even greater. Table 15 presents the results of this analysis.

Table 15. CELF-4 ELI Results for Exploratory Crosstabulation of Impairment Type and Response Type

			Response Type		
			Non-Responders	Responders	Total
Impairment Type	Exp/Phon. Impairment	Count	1	3	4
		Expected Count	2.9	1.1	4.0
		% within Exp/Phon. vs. Mixed	25.0%	75.0%	100.0%
	Mixed Rec-Exp. Impairment	Count	10	1	11
		Expected Count	8.1	2.9	11.0
		% within Exp/Phon. vs. Mixed	90.9%	9.1%	100.0%
	Total	Count	11	4	15
		Expected Count	11.0	4.0	15.0
		% within Exp/Phon. vs. Mixed	73.3%	26.7%	100.0%

The exploratory contingency table analyses thus reveal that children who showed no response on either the receptive or expressive language measures had a higher chance of belonging to the Mixed Receptive-Expressive category of impairment. These results should be considered with great caution due to small cell sizes and questionable validity.

APPENDIX B
Figures

Figure 1. Significant Quadratic Trend of Data for CELF-4 RLI

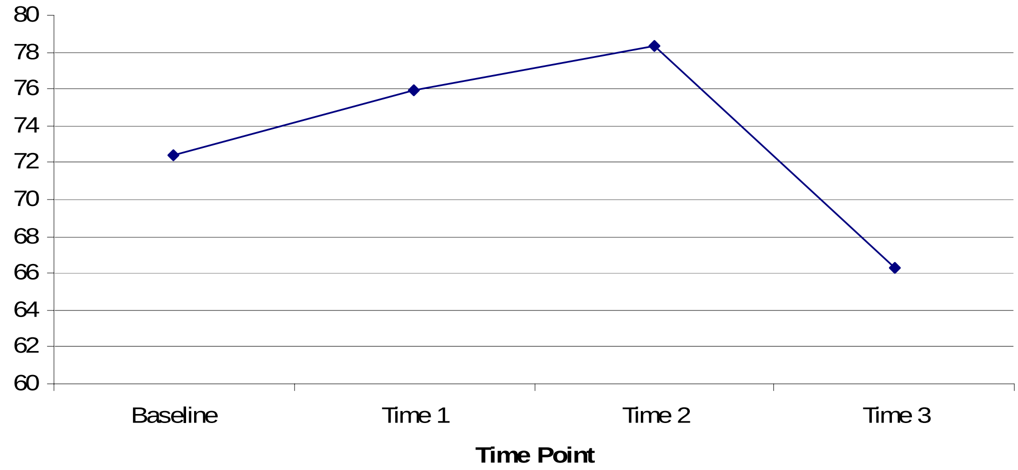


Figure 2. Significant Linear Trend of Data for PPVT-III

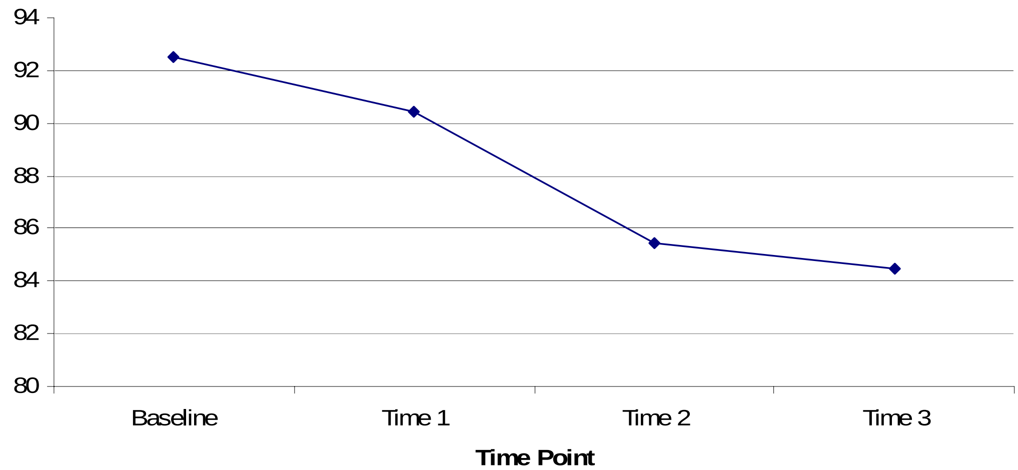


Figure 3. Significant Linear Trend of Data for Arizona-3

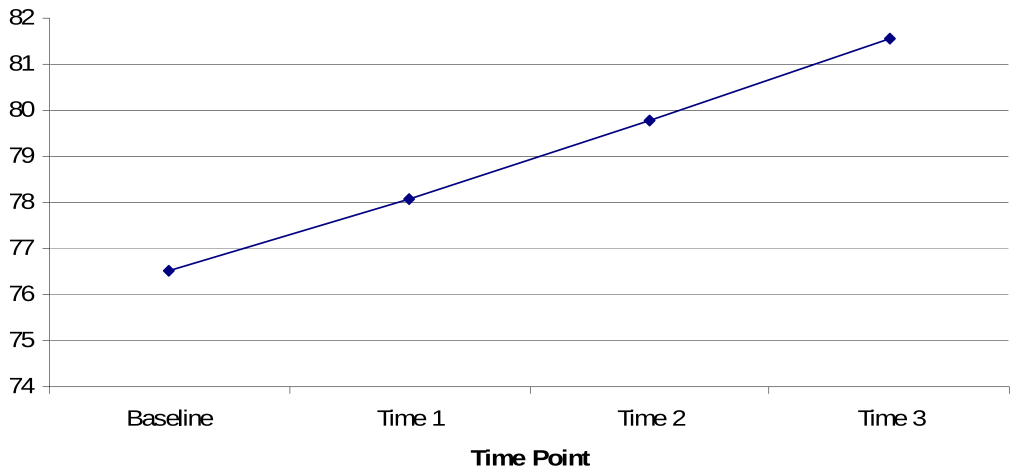
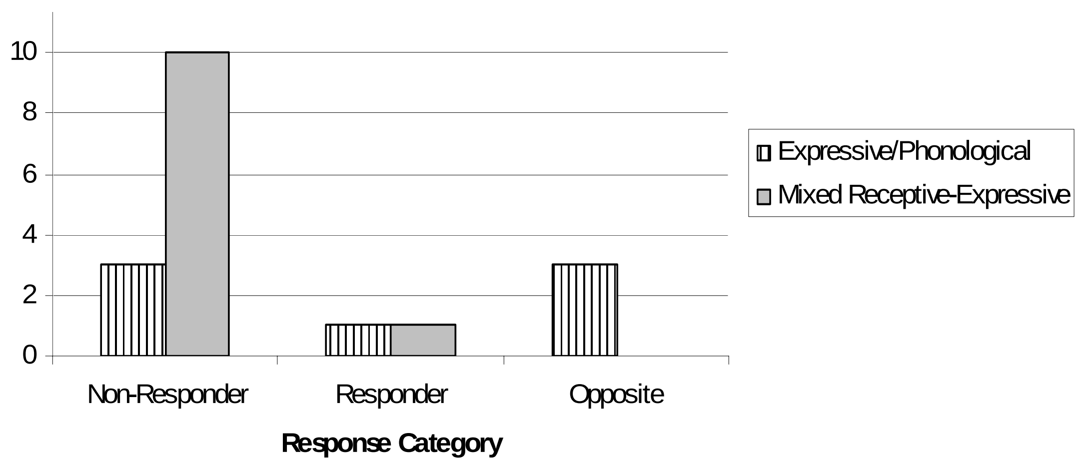


Figure 4. Response vs. Impairment Type Frequency for CELF-4 RLI



BIBLIOGRAPHY

Agocs, M., Burns, M. S., De Ley, L. E., Miller, S. L., & Calhoun, B. M. (2006). Fast ForWord Language. In R. J. McCauley & M. E. Fey (Eds.), *Treatment of Language Disorders in Children*. Baltimore, MD: Paul H. Brookes Publishing.

Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention. *Remedial and Special Education, 23*, 300-316.

Almost, D., & Rosenbaum, P. (1998). Effectiveness of speech intervention for phonological disorders: a randomised controlled trial. *Developmental Medicine and Child Neurology, 40*, 319-325.

American Academy of Child and Adolescent Psychiatry. (1998). Practice parameters for the assessment and treatment of children and adolescents with language and learning disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 37*(10 Supplement), 46S-62S.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders-Text Revision* (4th ed.). Washington, DC.

American Speech-Language Hearing Association. (2007). Late Blooming or Language Problem? Retrieved June 25, 2007, from http://www.asha.org/public/speech/disorders/Late-Blooming.htm

Aram, D. M., Morris, R., & Hall, N. E. (1993). Clinical and research congruence in identifying children with specific language impairment. *Journal of Speech & Hearing Research, 36*(3), 580-591.

Ayres, A. J. (1972). *Sensory Integration and Learning Disorders*. Los Angeles: Western Psychological Services.

Ayres, A. J. (1978). Learning disabilities and the vestibular system. *Journal of Learning Disabilities, 11*, 30-41.

Baker, L., & Cantwell, D. P. (1992). Attention deficit disorder and speech/language disorders. *Comprehensive Mental Health Care, 2*, 3-16.

Barker, J., & England, G. (1962). A numerical measure of articulation: further developments. *Journal of Speech and Hearing Disorders, 27*, 23-27.

Barratt, J., Littlejohns, P., & Thompson, J. (1992). Trial of intensive compared to weekly speech therapy in preschool children. *Archives of Disease in Childhood, 67*(1), 106-108.

Bashir, A. S., & Scavuzzo, A. (1992). Children with language disorders: Natural history and academic success. *Journal of Learning Disabilities, 25*(1), 53-65.

Beery, K. E., & Beery, N. A. (2004). *Beery-Buktenica Developmental Test of Visual-Motor Integration* (5th ed.). Los Angeles: Western Psychological Services.

Beitchman, J., Wilson, B., Brownlie, E., Walters, H., & Lancee, W. (1996). Long-term consistency in speech/language profiles: I. Developmental and academic outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry, 35*, 804-814.

Beitchman, J., & Young, A. (1997). Learning disorders with a special emphasis on reading disorders: a review of the past 10 years. *American Academy of Child and Adolescent Psychiatry, 36*, 1020-1032.

Beitchman, J. H., Hood, J., & Inglis, A. (1990). Psychiatric risk in children with speech and language disorders. *J Abnorm Child Psychol, 18*(3), 283-296.

Bell, N. (1987). *Visualising and Verbalising for language comprehension and thinking*. Paso Robles: Academy of Reading Publications.

Bell, N. (1991). Gestalt imagery: A critical factor in language comprehension. *Annals of Dyslexia, 41*, 246-260.

Bishop, D. (1994). Grammatical errors in specific language impairment: Competence or performance limitations? *Applied Psycholinguistics, 15*(4), 507-550.

Bishop, D., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders, and reading retardation. *Journal of Child Psychology and Psychiatry, 21*, 1027-1050.

Bishop, D., & Edmundson, A. (1987). Language impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders, 52*, 156-173.

Bishop, D. V. M. (1989). *Test for Reception of Grammar* (2nd ed.). Manchester, U. K.: Age and Cognitive Performance Research Centre, University of Manchester.

Bishop, D. V. M. (1997). *Uncommon understanding: Development and disorders of language comprehension in children*. Hove, England: Psychology Press/Erlbaum (UK) Taylor & Francis.

Bishop, D. V. M. (2004). *Expression, Reception, and Recall of Narrative Instrument (ERRNI)*. London: Psychological Corporation.

Bishop, D. V. M., Adams, C. V., & Rosen, S. (2006). Resistance of grammatical impairment to computerized comprehension training in children with specific and non-specific language impairments. *International Journal of Language and Communication Disorders, 41*(1), 19-40.

Bishop, D. V. M., & McArthur, G. (2005). Individual differences in auditory processing in specific language impairment: A follow-up study using event-related potentials and behavioural thresholds. *Cortex, 41*, 327-341.

Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: same or different? *Psychological Bulletin, 130*(6), 858-886.

Botting, N. (1998). Semantic-pragmatic disorder (SPD) as a distinct diagnostic entity: making sense of the boundaries. *Int J Lang Commun Disord, 33*(1), 87-90; discussion 95-108.

Botting, N. (2005). Non-verbal cognitive development and language impairment. *Journal of Child Psychology and Psychiatry, 46*(3), 317-326.

Botting, N., Faragher, B., Simkin, Z., Knox, E., & Conti-Ramsden, G. (2001). Predicting pathways of specific language impairment: what differentiates good and poor outcome? *J Child Psychol Psychiatry, 42*(8), 1013-1020.

Botting, N., Simkin, Z., & Conti-Ramsden, G. (2006). Associated reading skills in children with a history of Specific Language Impairment (SLI). *Reading and Writing, 19*, 77-98.

Boyle, J., Gillham, B., & Smith, N. (1996). Screening for early language delay in the 18-36 month age-range: The predictive validity of tests of production, and implications for practice. *Child Language Teaching & Therapy, 12*(2), 113-127.

Boyle, J., McCartney, E., Forbes, J., & O'Hare, A. (2007). A randomised controlled trial and economic evaluation of direct versus indirect and individual versus group modes of speech and language therapy for children with primary language impairment. *Health Technology Assessment, 11*(25), 1-139.

Bradley, L., & Bryant, P. (1983). Categorizing sounds and learning to read - A causal connection. *Nature, 301*, 419-421.

Brownell, R. (Ed.). (2000a). *Expressive One-Word Picture Vocabulary Test–Second Edition* (2nd ed.). Novalto, CA: Academic Therapy Publications.

Brownell, R. (Ed.). (2000b). *Receptive One-Word Picture Vocabulary Test-Second Edition* (2nd ed.). Novalto, CA: Academic Therapy Publications.

Bruce, B., & Hansson, K. (2008). Early Communication Skills; Important in Screening for Language Impairment and Neuropsychiatric Disorders. *Current Pediatric Reviews, 4*(1), 53-57.

Brutten, M., Richardson, S., & Mangel, C. (1973). *Something's wrong with my child*. New York: Harcourt Brace Jovanovich.

Cantwell, D., & Baker, L. (1987). *Developmental Speech and Language Disorders*. New York: Guilford.

Cantwell, D., & Baker, L. (1991). *Psychiatric and developmental disorders in children with communication disorder*. Washington, DC: American Psychiatric Press.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, England: Cambridge University Press.

Carroll, J. M., & Snowling, M. J. (2004). Language and phonological skills in children at high risk of reading difficulties. *Journal of Child Psychology and Psychiatry, 45*(3), 631-640.

Carte, E., Morrison, D., Sublett, J., Uemura, A., & Setrakian, W. (1984). Sensory integration therapy: A trial of specific neurodevelopmental therapy for the remediation of learning disabilities. *Journal of Developmental and Behavioral Pediatrics, 5*(4), 189-194.

Catts, H. W., Adlof, S. M., Hogan, T. P., & Ellis Weismer, S. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech Language & Hearing Research, 48*(6), 1378-1396.

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (1999). Language bases of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading, 3*, 331-362.

Cermak, S. A., & Mitchell, T. W. (2006). Sensory integration. In R. J. McCauley & M. E. Fey (Eds.), *Treatment of Language Disorders in Children*. Baltimore: Paul H. Brookes Publishing.

Christensen, L., & Mendoza, J. L. (1986). A method of assessing change in a single subject: An alteration of the RC index. *Behavior Therapy, 17*, 305-308.

Cirrin, F. M., & Gillam, R. B. (2008). Language intervention practices for school-age children with spoken language disorders: A systematic review. *Language, Speech, and Hearing Services in Schools, 39*, S110-S137.

Cohen, N. (2001). *Language imipairment and psychopathology in infants, children, and adolescents*. Thousand Oaks, CA: Sage.

Cohen, N., Vallance, D., Barwick, M., Im, N., Menna, R., Horodezky, N., et al. (2000). The interface between ADHD and language impairment: An examination of language, achievement, and cognitive processing. *Journal of Child Psychology & Psychiatry, 41*, 353-362.

Cohen, W., Hodson, A., O'Hare, A., Boyle, J., Durrani, T., McCartney, E., et al. (2005). Effects of computer-based intervention through acoustically modified speech (Fast ForWord) in severe mixed receptive-expressive language impairment: outcomes from a randomized controlled trial. *Journal of Speech, Language & Hearing Research, 48*(3), 715-729.

Cole, K. N., Coggins, T. E., & Vanderstoep, C. (1999). The influence of language/cognitive profile on discourse intervention outcome. *Language, Speech, & Hearing Services in Schools, 30*(1), 61-67.

Cole, K. N., & Dale, P. S. (1986). Direct language instruction and interactive language instruction with delayed preschool children: a comparison study. *Journal of Speech and Hearing Research, 29*, 206-217.

Cole, K. N., Dale, P. S., & Mills, P. E. (1990). Defining language delay in young children by cognitive referncing - Are we saying more than we know? *Applied Psycholinguistics, 11*(3), 291-302.

Cole, K. N., Maddox, M. E., & Lim, Y. S. (2006). Language is the Key: Constructive interactions around books and play In R. J. McCauley & M. E. Fey (Eds.), *Treatment of Language Disorders in Children*. Baltimore, MD: Paul H. Brookes Publishing.

Connell, P. J., & Stone, C. A. (1992). Morpheme Learning of Children with Specific Language Impairment under Controlled Instructional Conditions. *Journal of Speech and Hearing Research, 35*(4), 844.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues in field settings*. Chiacgo: Rand-McNally.

Critchley, M. (1984). *Developmental dyslexia*. London: W. Heineman Medical Books, Ltd.

Crowe, L. K. (2003). Comparison of two reading feedback strategies in improving the oral and written language performance of children with language-learning disabilities. *American Journal of Speech-Language Pathology, 12*(1), 16-27.

Curtiss, S., & Yamada, J. (Unpublished work). Curtiss and Yamada Comprehensive Language Evaluation-Receptive (CYCLE-R).

Dawson, M. (1987). Minority student performance: Is the Montessori magnet school effective? Retrieved March 25, 2008, from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED309881&site=ehost-live

DeThorne, L. S., & Schaefer, B. A. (2004). A guide to child nonverbal IQ measures. *American Journal of Speech-Language Pathology, 13*, 275-290.

Dethorne, L. S., & Watkins, R. V. (2006). Language abilities and nonverbal IQ in children with language impairment: inconsistency across measures. *Clinical Linguistics and Phonetics, 20*(9), 641-658.

Devilly, G. J. (2004). The Reliable Change Generator for Windows: Version 2.0: Centre for Neuropsychology, Swinburne University, Australia

DiSimoni, F. (1978). *The Token Test for Children*. Boston, MA: Teaching Resources Corporation.

Dixon, G., Joffe, B., & Bench, R. J. (2001). The efficacy of visualizing and verbalizing: Are we asking too much? *Child Language Teaching and Therapy, 17*(2), 127-141.

Dodd, B., Crosbie, S., MacIntosh, B., Teitzel, T., & Ozanne, A. (2000). *Preschool and Primary Inventory of Phonological Awareness*. London: Psychological Corporation.

Dodd, B., Holm, A., Oerlemans, M., & McCormick, M. (1996). *Queensland University Inventory of Literacy*. University of Queensland, Australia: Department of Speech Pathology and Audiology.

Duax, T. (1995). Report on academic achievement in a private Montessori school. *The National American Montessori Teachers' Association Journal, 20*(2), 145-147.

DuBard, N. E., & Martin, M. K. (2000). *Teaching language-deficient children: Theory and application of the association method for multisensory teaching.* Cambridge, MA: Educators Publishing Service, Inc.

DuBard School for Language Disorders. (1998). Multisensory approach to reading in lower elementary classes. Unpublished Pilot study. The University of Southern Mississippi DuBard School for Language Disorders, Mississippi Power Foundation, and Petal Elementary School.

Dunn, L., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test - Revised* Circle Pines, MN: American Guidance Service.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test* (3rd ed.). Circle Pines, MN: American Guidance Service.

Durrell, D. D., & Catterson, J. H. (1980). *Durrell Analysis of Reading Difficulty* (3rd ed.). San Antonio, TX: Psychological Corporation.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic Awareness Instruction Helps Children Learn to Read: Evidence from the National Reading Panel's Meta-Analysis. *Reading Research Quarterly, 36*(3), 250-287.

Elliot, C. D. (1983). *British Ability Scales*. Windsor, U. K.: NFER-Nelson.

Estil, L., Whiting, H., Sigmundsson, H., & Ingvaldsen, R. (2003). Why might language and motor impairments occur together? *Infant and Child Development, 12*(3), 253-265.

Faul, F., & Erdfelder, E. (1992). GPOWER: A priori, post-hoc, and compromise power analyses for MS-DOS Bonn, FRG: Bonn University Department of Psychology.

Fey, M. E., Cleave, P. L., & Long, S. H. (1997). Two models of grammar facilitation in children with language impairments: Phase 2. *Journal of Speech and Hearing Research, 40*, 5-19.

Fey, M. E., Cleave, P. L., Long, S. H., & Hughes, D. L. (1993). Two approaches to the facilitation of grammar in children with language impairment: An experimental evaluation. *Journal of Speech and Hearing Research, 36*, 141-157.

Flax, J., Realpe-Bonilla, T., Hirsch, L. S., Brzustowicz, L. M., Bartlett, C. W., & Tallal, P. (2003). Specific language impairment in families: Evidence for co-occurrence with reading impairments. *Journal of Speech, Language & Hearing Research, 46*, 530-543.

Foster, L. M., Hynd, G. W., Morgan, A. E., & Hugdahl, K. (2002). Planum temporale asymmetry and ear advantage in dichotic listening in developmental dyslexia and attention-deficit/hyperactivity disorder (ADHD). *Journal of the International Neuropsychological Society, 8*(1), 22-36.

Fuchs, D., & Young, C. L. (2006). On the Irrelevance of Intelligence in Predicting Responsiveness to Reading Instruction. [Original]. *Exceptional Children, 73*(1), 8-30.

Fudala, J. B. (2003). *Arizona Articulation Proficiency Scale – Third Revision*. Los Angeles: Western Psychological Services.

Gallagher, T. (1998). Treatment research in speech, language and swallowing: Lessons from child language. *Folia Phoniatrica Logopaedica, 50*(165-182).

Gartner, A., & Kerzner-Lipsky, D. (2003). Outcomes for students in a Montessori program: A longitudinal study of the experience in the Milwaukee public schools. Retrieved March 30, 2008, from http://www.montessori-ami.org/research/outcomes.pdf

Gauger, L. M., Lombardino, L. J., & Leonard, C. M. (1997). Brain morphology in children with specific language impairment. *Journal of Speech, Language & Hearing Research, 40*, 1272-1284.

German, D. (1986). *Test of Word Finding*. Allen, TX: DLM Teaching Resources.

Gillberg, C. (1998). Hyperactivity, inattention, and motor control problems: prevalence, comorbidity and background factors. *Folia Phoniatrica et Logopedica, 50*, 107-117.

Gillon, G. T. (2000). The efficacy of phonological awareness intervention for children with spoken language impairment. *Language, Speech, and Hearing Services in Schools, 31*, 126-141.

Gillon, G. T. (2002). Follow-up study investigating the benefits of phonological awareness intervention for children with spoken language impairment. *International Journal of Language & Communication Disorders, 37*(4), 381-400.

Gillon, G. T. (2005). Facilitating phoneme awareness development in 3- and 4-year-old children with speech impairment. *Language, Speech, and Hearing Services in Schools, 36*, 308-324.

Gillon, G. T. (2006). Phonological awareness intervention: A framework for preschool children with specific speech and language impairments. In R. J. McCauley & M. E. Fey (Eds.), *Treatment of Language Disorders in Children* (pp. 279-308). Baltimore, MD: Paul H. Brookes Publishing.

Gillon, G. T., & Moriarty, B. C. (2007). Childhood apraxia of speech: Children at risk for persistent reading and spelling disorder. *Seminars in Speech and Language, 28*(1), 48-57.

Gilmore, A., Croft, C., & Reid, N. (1981). *Burt Word Reading Test*. Wellington, NZ: NZCER.

Girolametto, L., & Weitzman, E. (2006). It takes two to talk - The Hanen Program for parents: Early language intervention through caregiver training. In R. J. McCauley & M. E. Fey (Eds.), *Treatment of Language Disorders in Children*. Baltimore, MD: Paul H. Brookes Publishing.

Goldman, R., & Fristoe, M. (1969). *Goldman-Fristoe Test of Articulation*. Circle Pines, MN: American Guidance Service.

Goldman, R., & Fristoe, M. (1986). *Goldman-Fristoe Test of Articulation*. Minnesota: American Guidance Association.

Goldman, R., Fristoe, M., & Woodcock, R. W. (1974). *Goldman-Fristoe-Woodcock Diagnostic Auditory Discrimination Test* Circle Pines, MN: American Guidance Service.

Goodman, L., & Hammill, D. D. (1975). *Basic School Skills Inventory*. Chicago, IL: Follett.

Goorhuis-Brouwer, S. M., & Knijff, W. A. (2001). Efficacy of speech therapy in children with language disorders: Specific language impairment compared with language impairment in comorbidity with cognitive delay. *International Journal of Otorhinolaryngology, 63*, 129-136.

Green, S. B., & Salkind, N. J. (2005). *Using SPSS for Windows and Macintosh: Analyzing and understanding data*. Upper Saddle River, NJ: Pearson Prentice Hall.

Griffer, M. R. (1999). Is sensory-integration effective for children with language-learning disorders?: A critical review of the evidence. *Language, Speech, and Hearing Services in Schools, 30*, 393-400.

Grigorenko, E. L. (2001). Developmental dyslexia: An update on genes, brains, and environments. *Journal of Child Psychology and Psychiatry, 42*(1), 91-125.

Grigorenko, E. L., Wood, F. B., Golovyan, L., Meyer, M., Romano, C., & Pauls, D. (2003). Continuing the search for dyslexia genes on 6p. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics), 118B*, 89-98.

Grigorenko, E. L., Wood, F. B., Meyer, M. S., Hart, L. A., Speed, W. C., Shuster, A., et al. (1997). Susceptibility loci for distinct components of developmental dyslexia on chromosomes 6 and 15. *American Journal of Human Genetics, 60*(1), 27-39.

Hancock, T. B., & Kaiser, A. P. (2006). Enhanced Milieu Teaching. In R. J. McCauley & M. E. Fey (Eds.), *Treatment of Language Disorders in Children*. Baltimore, MD: Paul H. Brookes Publishing.

Hayden, D., & Square, P. (1999). *Verbal Motor Production Assessment for Children: Examiner's Manual* The Psychological Corporation.

Hill, E. L. (2001). Non-specific nature of specific language impairment: A review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders, 36*(2), 149-171.

Hill, E. L., Bishop, D. V. M., & Nimmo-Smith, I. (1998). Representational gestures in developmental coordination disorder and specific language impairment: error-types and the reliability of ratings. *Human Movement Science, 17*, 655-678.

Hoehn, T. P., & Baumeister, A. A. (1994). A critique of the application of sensory integration therapy to children with learning disabilities. *Journal of Learning Disabilities, 27*(6), 338-350.

Hogan, T. P., Farr, R. C., Prescott, C. A., & Balow, I. H. (1986). *Metropolitan Achievement Test 6*. San Antonio, TX: The Psychological Corporation.

Horwitz, S. M., Irwin, J. R., Briggs-Gowan, M. J., Bosson Heenan, J. M., Mendoza, J., & Carter, A. S. (2003). Language delay in a community cohort of young children. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*(8), 932-940.

Humphries, T., Wright, M., McDougall, B., & Vertes, J. (1990). The efficacy of sensory integration therapy for children with learning disability. *Physical and Occupational Therapy in Pediatrics, 10*(3), 1-17.

Humphries, T., Wright, M., Snider, L., & McDougall, B. (1992). A comparison of the effectiveness of sensory-integrative and perceptual-motor training in treating children with learning disabilities. *Developmental and Behavioral Pediatrics, 13*(1), 31-40.

Hynd, G. W., Semrud-Clikeman, M., Lorys, A. R., Novey, E. S., & Eliopulos, D. (1990). Brain morphology in developmental dyslexia and attention deficit disorder/hyperactivity. *Archives of Neurology, 47*(8), 919-926.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15*, 336-352.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12-19.

Jacoby, G., Lee, L., Kummer, A., Levin, L., & Creaghead, N. (2002). The number of individual treatment units necessary to facilitate functional communication improvements in the speech and language of young children. *American Journal of Speech-Language Pathology, 11*, 370-380.

Jancke, L., Siegenthaler, T., Preis, S., & Steinmetz, H. (2007). Decreased white-matter density in a left-sided fronto-temporal network in children with developmental language disorder: Evidence for anatomical anomalies in a motor-language network. *Brain and Language, 102*, 91-98.

Jastak, J. F., & Jastak, S. (1978). *Wide Range Achievement Test*. Wilmington, DE: Jastak Associates, Inc.

Johnson, C., Beitchman, J., Young, A., Escobar, M., Atkinson, L., Wilson, B., et al. (1999). Fourteen-year follow-up of children with and without speech/language impairments: Speech/language stability and outcomes. *Journal of Speech, Language, & Hearing Research, 42*, 744-760.

Jones, J. H. L. (1971a). *Early Childhood Education: An Objective Study*. Lafayette, LA: Reading Disability Clinic of the Reading Research Foundation.

Jones, J. H. L. (1971b). *ESEA Title III Project Evaluation of Specific Identification/Instruction: Early Childhood Education*: Natchez Special Municipal Separate School District.

Kamphaus, R., & Reynolds, C. (1992). *Behavioral Assessment System for Children*. New York: AGS Publishing.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.

Kavanagh, E. (2007, February 2). Letters. *Science, 315,* 596-598.

Kirk, S. A., McCarthy, J. J., & Kirk, W. D. (1968). *Illinois Test of Psycholinguistic Abilities*. Chicago, IL: University of Illinois.

Lahey, M. (1990). Who shall be called language disordered? Some reflections and one perspective. [Journal
Peer-Reviewed Status-Unknown]. *Journal of Speech & Hearing Disorders, 55*(4), 612-620.

Laija, W. (1997). *Conducting ANOVA trend analyses using polynomial contrasts*. Paper presented at the Annual Meeting of the Southwest Educational Research Association

Larrivee, L. S., & Catts, H. W. (1999). Early reading achievement in children with expressive phonological disorders. *American Journal of Speech-Language Pathology, 8*, 118-128.

Larsen, J. P., Høien, T., Lundberg, I., & Odegaard, H. (1990). MRI evaluation of the size and symmetry of the planum temporale in adolescents with developmental dyslexia. *Brain and Language, 39*(2), 289-301.

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *International Journal of Language & Communication Disorders, 35*, 165-188.

Law, J., Garrett, Z., & Nye, C. (2003). Speech and language therapy interventions for children with primary speech and language delay or disorder. *Cochrane Database of Systematic Reviews*(3).

Law, J., Garrett, Z., & Nye, C. (2004). The efficacy of treatment for children with developmental speech and language delay/disorder: A meta-analysis. *Journal of Speech, Language & Hearing Research, 47*, 924-943.

Lee, L. (1974). *Developmental sentence analysis*. Evanston, IL: Northwestern University Press.

Leonard, L. (1990). Language disorders in preschool children. In G. H. Shames & E. H. Wiig (Eds.), *Human Communication Disorders* (3rd ed., pp. 607). Columbus, OH: Merrill Publishing Company.

Leonard, L. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Lillard, A., & Else-Quest, N. (2006). Evaluating Montessori education. *Science* Retrieved February 18, 2008, from http://www.sciencemag.org/cgi/reprint/313/5795/1893.pdf

Lindamood, C., & Lindamood, P. (1979). *Lindamood Auditory Conceptualization Test*. Allen, TX: DLM Teaching Resources.

Llorente, A., Sines, M. C., Rozelle, J. C., Turcich, M. R., & Casatta, A. (2000). Effects of test administration order on children's neuropsychological performance: Emerging one-word expressive and receptive language skills. *The Clinical Neuropsychologist, 14*(2), 162-172.

Lopata, C., Wallace, N. V., & Finn, K. V. (2005). Comparison of academic achievement between Montessori and traditional education programs. *Journal of Research in Childhood Education, 20*(1), 5-13.

Magrun, W. M., Ottenbacher, K., McCue, S., & Keefe, R. (1981). Effects of vestibular stimulation on spontaneous use of verbal language in developmentally delayed children. *American Journal of Occupational Therapy, 35*, 101-104.

Mandelbaum, D. E., Stevens, M., Rosenberg, E., Wiznitzer, M., Steinschneider, M., Filipek, P., et al. (2006). Sensorimotor performance in school-age children with autism, developmental language disorder, or low IQ. *Developmental Medicine and Child Neurology, 48*, 33-39.

Mauer, D. (1999). Issues and applications of sensory integration theory and treatment with children with language disorders. *Language, Speech, and Hearing Services in Schools, 30*, 383-392.

McCauley, R. J., & Demetras, M. (1990). The identification of language impairment in the selection of specifically language-impaired subjects. *Journal of Speech & Hearing Disorders, 55*(3), 468-475.

McCauley, R. J., & Fey, M. E. (Eds.). (2006). *Treatment of Language Disorders in Children*. Baltimore: Paul H. Brookes Publishing.

McCauley, R. J., & Strand, E. A. (2008). A review of standardized tests of nonverbal oral and speech motor performance in children. *American Journal of Speech-Language Pathology, 17*(1), 81-91.

McGinnis, M. (1939). *Aphasic children: Identification and education by the association method.* Washington, D.C.: Alexander Graham Bell Association for the Deaf.

McGraw-Hill. (1974). *Comprehensive Tests of Basic Skills*. Monterey, CA: CTB/McGraw-Hill.

McGraw-Hill. (2002). *TerraNova, 2nd edition: Technical Bulletin 1*. Monterey, CA: McGraw-Hill.

Merzenich, M., Jenkins, W., Johnston, P., Schriener, C., Miller, S., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science, 271*, 77-80.

Miller, L. B., & Bizzell, R. P. (1984). Long-term effects of four preschool programs: Ninth- and tenth-grade results. *Child Development, 55*, 1570-1587.

Montessori. (2005). Retrieved October 2, 2005, from http://www.montessori.edu

Montessori, M. (1988). *The Montessori Method*. New York: Schocken.

Muter, V., Hulme, C., & Snowling, M. (1997). *Phonological Abilities Test*. London: Psychological Corporation.

Nelson, J., Benner, G., & Gonzales, J. (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review. *Learning Disabilities Research and Practice, 18*, 255-267.

Newcomer, P. L., & Hammill, D. D. (1982). *Test of Language Development - Primary*. Austin, TX: Pro-ed.

Norris, J. A. (1988). Using communication strategies to enhance reading acquisition. *The Reading Teacher, 41*, 368-373.

Oliver, B., Dale, P. S., & Plomin, R. (2004). Verbal and nonverbal predictors of early language problems: An analysis of twins in early childhood back to infancy. *Journal of Child Language, 31*(3), 609-631.

Pickering, J. S. (1988). *Montessori applied to children at high risk.* Paper presented at the meeting of the Orton Dyslexia Society.

Pickering, J. S. (1990). Montessori applied to children at high-risk. Unpublished work. The Shelton School.

Pickering, J. S. (1992). Successful applications of Montessori methods with children at risk for learning disabilities. *Annals of Dyslexia, 42*, 90-109.

Pickering, J. S. (2004a). The At Risk child: How the Montessori classroom enhances learning: Part 1. *Montessori Life* Retrieved November 14, 2007, from http://findarticles.com/p/articles/mi_qa4097/is_200401/ai_n9398943

Pickering, J. S. (2004b). The At Risk child: How the Montessori classroom enhances learning: Part 2. *Montessori Life* Retrieved November 14, 2007, from http://findarticles.com/p/articles/mi_qa4097/is_200404/ai_n9399221

Pickering, J. S., Coffman, N., Stanislav, A., Kneese, L., & Snyder, S. (2003). *Profile to prescription: Individualizing MSLE instruction - The Shelton School model*. Paper presented at the International Dyslexia Association.

Plante, E. (1991). MRI findings in the parents and siblings of specifically language-impaired boys. *Brain and Language, 41*, 52-66.

Plante, E., Swisher, L., Vance, R., & Rapcsak, S. (1991). MRI findings in boys with specific language impairment. *Brain and Language, 41*, 52-66.

Polatajko, H. J., Kaplan, B. J., & Wilson, B. N. (1992). Sensory integration treatment for children with learning disabilities: Its status 20 years later. *Occupational Therapy Journal of Research, 12*, 323-341.

Polatajko, H. J., Law, M., Miller, J., Schaffer, R., & Macnab, J. (1991). The effect of a sensory integration program on academic achievement, motor performance, and self-esteem in children identified as learning disabled: Results of a clinical trial *Occupational Therapy Journal of Research, 11*, 155-176.

Polka, L., & Rvachew, S. (2005). The Impact of Otitis Media With Effusion on Infant Phonetic Perception. *Infancy, 8*(2), 101-117.

Prathanee, B. M. A., Thinkhamrop, B. P., & Dechongkit, S. P. (2007). Factors Associated With Specific Language Impairment and Later Language Development During Early Life: A Literature Review. *Clinical Pediatrics, 46*(1), 22-29.

Preis, S., Engelbrecht, V., Huang, Y., & Steinmetz, H. (1998). Focal grey matter heterotopias in monozygotic twins with developmental language disorder. *European Journal of Pediatrics, 157*, 849-852.

Preis, S., Lancke, L., Schittler, P., Huang, Y., & Steinmetz, H. (1998). Normal intrasylvian anatomical asymmetry in children with developmental language disorder. *Neuropsychologia, 36*(9), 849-855.

Preis, S., Steinmetz, H., Knorr, U., & Jancke, L. (2000). Corpus callosum size in children with developmental language disorder. *Cognitive brain research, 10*, 37-44.

Rapin, I. (1996). Practitioner review: Developmental language disorders: A clinical update. *Journal of Child Psychology and Psychiatry, 37*(6), 643-655.

Rapin, I., & Allen, D. (1983). Developmental language disorders: Nosologic considerations. In U. Kirk (Ed.), *Neuropsychology of language, reading and spelling* (pp. 155-184). New York: Academic Press.

Raven, J. C. (1986). *Coloured Progressive Matrices*. London: H. K. Lewis.

Renfrew, C. (1991). *The Renfrew Language Scales: Bus Story Test*. Bicester, UK: Winslow Press.

Reynell, J. (1977). *Reynell Developmental Language Scales - Revised*. Windsor: NFER Publishing Company.

Rice, M. L., Sell, M. A., & Hadley, P. A. (1991). Social interactions of speech- and language-impaired children. *Journal of Speech & Hearing Research, 34*(6), 1299-1307.

Roid, G. H., & Miller, L. J. (1997). *Leiter International Performance Scale – Revised: Examiner's Manual* Wood Dale, IL: Stoelting Co.

Rosner, J. (1975). *Helping children overcome learning difficulties: A step-by-step guide for parents and teachers*. New York: Walker.

Rutter, M., & Mawhood, L. (1991). The long-term psychosocial sequelae of specific developmental disorders of speech and language. In M. Rutter & P. Casaer (Eds.), *Biological risk factors for psychosocial disorders* (pp. 233-259). Cambridge: Cambridge University Press.

Rvachew, S., Creighton, D., Feldman, N., & Suave, R. (2005). Vocal development of infants with very low birth weight. *Clinical Linguistics & Phonetics, 19*(4), 275-294.

Rvachew, S., & Grawburg, M. (2006). Correlates of phonological awareness in preschoolers with speech sound disorders. *Journal of Speech, Language, and Hearing Research, 49*(1), 74-87.

Scarborough, H. (1990). Very early language deficits in dyslexic children. *Child Development, 61*, 1728-1743.

Schaaf, R. C., & Miller, L. J. (2005). Occupational therapy using a sensory integrative approach for children with developmental disabilities. *Mental Retardation and Developmental Disabilities, 11*, 143-148.

Schery, T. K. (1985). Correlates of language development in language-disordered children. *Journal of Speech and Hearing Disorders, 50*(1), 73-83.

Schraeder, M. (2008). Apraxia, Reading, and the Association Method. Unpublished Case studies. The University of Southern Mississippi Dubard School for Language Disorders

Scientific Learning Corporation. (1998). Fast ForWord - Language. Berkeley, CA: Author.

Segers, E., & Verhoeven, L. (2004). Computer-supported phonological awareness intervention for kindergarten children with specific language impairment. *Language, Speech, & Hearing Services in Schools, 35*(3), 229-239.

Seldin, T. (2002/03). What research says about Montessori's effectiveness. *Tomorrow's Child,* 5-11.

Seldin, T. (2006). Montessori 101: Some basic information that every Montessori parent should know. *Journal*. Retrieved from http://www.montessori.org/sitefiles/Montessori_101_nonprintable.pdf

Semel-Mintz, E., & Wigg, E. H. (1982). *Clinical Evaluation of Language Functions: Diagnostic Battery*. Columbus, OH: Merrill.

Semel, E., Wiig, E. H., & Secord, W. A. (1995). *The Clinical Evaluation of Language Fundamentals - Third Edition*. San Antonio, TX: Psychological Corporation.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals – Fourth Edition*. San Antonio: Harcourt Assessment, Inc.

Shaw, S. R. (2008). A school psychologist investigates sensory integration therapies: promise, possibility, and the art of placebo. Retrieved May 23, 2008, from http://www.autismtoday.com/articles/School_Psychologist_Investigates_Sensory_Integration.htm

Shaw, S. R., Powers, N. R., Abelkop, S., & Mullis, J. (2002). *Sensory integration therapy: Panacea, placebo, or poison?* Paper presented at the Annual Convention of the National Association of School Psychologists.

Shaywitz, B. A., Holford, T. R., Holahan, J. M., Fletcher, J. M., & et al. (1995). A Matthew effect for IQ but not for reading: Results from a longitudinal study. *Reading Research Quarterly, 30*(4), 894-906.

Shedd, C. L. (1967). Some characteristics of a specific perceptual motor disability. *Journal of the Medical Association of Alabama, 37*, 150-162.

Siegel, L. S. (1982). Reproductive, perinatal, and environmental factors as predictors of the cognitive and language development of preterm and full-term infants. *Child Development, 53*, 963-973.

Slomkowski, C. L., Nelson, K., Dunn, J., & Plomin, R. (1992). Temperament and language: Relations from toddlerhood to middle childhood. *Developmental Psychology, 28*, 1090-1095.

Slosson, R. L., Nicholson, C. L., & Hibpshman, T. L. (1991). *Slosson Intelligence Test-Revised.* . Columbia, MO: Hawthorne Educational Services, Inc.

Snowling, M., Bishop, D. V., & Stothard, S. E. (2000). Is preschool language impairment a risk factor for dyslexia in adolescence? *J Child Psychol Psychiatry, 41*(5), 587-600.

Snowling, M. J., & Hayiou-Thomas, M. E. (2006). The Dyslexia Spectrum: Continuities Between Reading, Speech, and Language Impairments. *Topics in Language Disorders, 26*(2), 110-126.

Spitz, R., Tallal, P., Flax, J., & Benasich, A. (1997). Look who's talking: A prospective study of familial transmission of language impairments. *Journal of Speech, Language, & Hearing Research, 40*, 990-1001.

Stackhouse, J., & Wells, B. (1997). How do speech and language problems affect literacy development? In C. Hulme & M. J. Snowling (Eds.), *Dyslexia: Biology, cognition, and intervention* (pp. 182-211). London: Whurr Publishing.

Stanislav, A. (2007). Reasons behind the structure of the EI classrooms and clinical impressions [Personal communication] Electronic mail received October 29, 2007.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360-406.

Stanton-Chapman, T. L., Chapman, D. A., Bainbridge, N. L., & Scott, K. G. (2002). Identification of early risk factors for language impairment. *Research in Developmental Disabilities, 23*(6), 390-405.

Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B., & Kaplan, C. A. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language & Hearing Research, 41*, 407-418.

Sullivan, A., & Perigoe, C. B. (2004). The Association Method for children with hearing loss and special needs. *The Volta Review, 104*(4), 339-348.

Swisher, L., & Plante, E. (1993). Nonverbal IQ tests reflect different relations among skills for specifically language-impaired and normal children. *Journal of Communication Disorders, 26*, 65-71.

Swisher, L., Plante, E., & Lowell, S. (1994). Nonlinguistic deficits of children with language disorders complicate the interpretation of their nonverbal IQ scores. *Language, Speech, & Hearing Services in Schools, 25*, 235-240.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language, 9*, 182-198.

Tallal, P., & Miller, S. (1994). Computerized Version of the Tallal Repetition Test (unpublished).

Tallal, P., Miller, S., Bedi, G., Byma, G., Wang, X., & Nagarajan, S. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science, 271*, 81-84.

Tallal, P., & Piercy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception *Neuropsychologia, 12*, 83-93.

Tallal, P., Ross, R., & Curtiss, S. (1989). Family aggregation in specific language impairment. *Journal of Speech, Hearing and Language Research, 27*, 987-998.

Tallal, P., Townsend, J., Curtiss, S., & Wulfeck, B. (1991). Phenotypic profiles of language-impaired children based on genetic/family history. *Brain and Language, 41*(1), 81-95.

Throneburg, R. N., Calvert, L. K., Sturm, J. J., Paramboukas, A. A., & Paul, P. J. (2000). A comparison of service delivery models: effects on curricular vocabulary skills in the school setting. *American Journal of Speech-Language Pathology, 9*(1), 10-20.

Thurm, A., Lord, C., Li-Ching, L., & Newschaffer, C. (2007). Predictors of Language Acquisition in Preschool Children with Autism Spectrum Disorders. *Journal of Autism & Developmental Disorders, 37*(9), 1721-1734.

Tomblin, J., Records, N., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 40*, 1245-1260.

Tomblin, J., Smith, E., & Zhang, X. (1997). Epidemiology of specific language impairment: Prenatal and perinatal risk factors. *Journal of Communication Disorders, 30*, 325-344.

Tomblin, J. B., & Zhang, X. (1999). Language patterns and etiology in children with specific language impairment In H. Tager-Flusberg (Ed.), *Neurodevelopmental disorders* (pp. 361-382). Cambridge, MA: MIT Press.

Trauner, D., Wulfeck, B., Tallal, P., & Hesselink, J. (2000). Neurological and MRI profiles of children with developmental language impairment. *Developmental Medicine and Child Neurology, 42*, 470-475.

Trott, M. C., Laurel, M. K., & Windeck, S. L. (1993). *Understanding sensory integration*. Tucson, AZ: Therapy Skill Builders.

Ullman, M. T., & Pierpont, E. I. (2005). Specific Language Impairment Is Not Specific To Language: The Procedural Deficit Hypothesis. *Cortex, 41*(3), 399-433.

Van der Lely, H., & Stollwerk, L. (1996). A grammatical specific language impairment in children. An autosomal dominant inheritance? *Brain and Language, 52*, 484-504.

van Kleeck, A., Gillam, R. B., & McFadden, T. (1998). A study of classroom-based phonological awareness training for preschoolers with speech and/or language disorders. *American Journal of Speech-Language Pathology, 7*(3), 65-76.

Van Weerdenburg, M., Verhoeven, L., & van Balkom, H. (2006). Towards a typology of specific language impairment. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 47*(2), 176-189.

Vargas, S., & Camilli, G. (1999). A meta-analysis of research on sensory-integration treatment. *American Journal of Occupational Therapy, 53*, 189-198.

Voci, S. C., Beitchman, J. H., Brownlie, E., & Wilson, B. (2006). Social anxiety in late adolescence: The importance of early childhood language impairment. *Journal of Anxiety Disorders, 20*(7), 915-930.

Wallace, G., & Hammill, D. D. (1994). *Comprehensive Receptive and Expressive Vocabulary Test*. Austin, TX: Pro-ed.

Watkins, K. E., Vargha-Khadem, F., Ashburner, J., Passingham, R. E., Connelly, A., Friston, K. J., et al. (2002). MRI analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain, 125*(3), 465-478.

Webster, R. I., & Shevell, M. I. (2004). Neurobiology of Specific Language Impairment. *Journal of Child Neurology, 19*(7), 471-481.

Wechsler, D. (1967). *Wechsler Preschool and Primary Scale of Intelligence*. New York: Psychological Corporation.

Wechsler, D., & Naglieri, J. A. (2006). *Wechsler Nonverbal Scale of Ability*. San Antonio, TX: Harcourt Assessment, Inc.

Wechsler, D. J. (1974). *Wechsler Intelligence Scale for Children - Revised*. New York: Psychological Corporation.

Weindrich, D., Jennen-Steinmetz, C., Laucht, M., Esser, G., & Schmidt, M. H. (1998). At risk for language disorders? Correlates and course of language disorders in preschool children born at risk. *Acta Paediatrica, 87*(12), 1288-1294.

Weismer, S., & Hesketh, L. (1993). The influences of prosodic and gestural cues on novel word acquisition by children with specific language impairment. *Journal of Speech & Hearing Research, 36*, 1013-1025.

Weismer, S., & Murray-Branch, J. (1989). Modeling versus modeling plus evoked production training: A comparison of two language interventions. *Journal of Speech & Hearing Disorders, 54*, 269-281.

Whitehurst, G. J., Fischel, J. E., Lonigan, C. J., Valdez-Menchaca, M. C., Arnold, D. S., & Smith, M. (1991). Treatment of early expressive language delay: If, when, and how. *Topics in Language Disorders, 11*, 55-68.

Wickstrom, S., Goldstein, H., & Johnson, L. (1985). On the subject of subjects: Suggestions for describing subjects in language intervention studies. *Journal of Speech & Hearing Disorders, 50*(3), 282-286.

Wiederholt, J. L., & Bryant, B. R. (1986). *Gray Oral Reading Tests - Revised*. Austin, TX: Pro-ed.

Wiig, E. H., Secord, W. A., & Semel, E. (2004). *Clinical Evaluation of Language Fundamentals Preschool – Second Edition*. San Antonio: Harcourt Assessment, Inc.

Wilson, B. N., Kaplan, B. J., Fellowes, S., Gruchy, C., & Faris, P. (1992). The efficacy of sensory integration treatment compared to tutoring. *Physical and Occupational Therapy in Pediatrics, 12*(1), 1-36.

Wing, C. (1990). A preliminary investigation of generalization to untrained words following two treatments of children with word-finding problems. *Language, Speech, and Hearing Services in Schools, 21*, 151-156.

Wood, F. B., & Grigorenko, E. L. (2001). Emerging issues in the genetics of dyslexia: A methodological preview. *Journal of Learning Disabilities, 34*(6), 503-511.

Woodcock, R. W., & Johnson, M. B. (1977). *Woodcock-Johnson Psychoeducational Battery*. Hingham, MA: Teaching Resources.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III: Tests of Achievement*. Itasca, IL: Riverside Publishing.

Woods, M., & Moe, A. (1995). *Analytical reading inventory*. Engelwood Cliffs, NJ: Prentice-Hall.

Yoder, P., & McDuffie, A. (2002). Treatment of primary language disorders in early childhood: Evidence of efficacy. In P. Accardo, B. Rogers & A. Capute (Eds.), *Disorders of language development* (pp. 151-177). Baltimore: York Press.

Young, A. R., Beitchman, J. H., Johnson, C., Douglas, L., Atkinson, L., Escobar, M., et al. (2002). Young adult outcomes in a longitudinal sample of early identified language impaired and control children. *Journal of Child Psychology and Psychiatry, 43*(5), 635-645.

Zimmerman, I., Steiner, V., & Pond, R. (1979). *Preschool Language Scale - Revised*. Columbus, OH: Charles E. Merrill.