Second Language Proficiency in Sequential Bilingual Children

with and without Primary Language Impairment

by

Ekaterina Smyk

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2012 by the
Graduate Supervisory Committee:

Maria Adelaida Restrepo, Chair
Shelley Gray
Joanna Gorin

ARIZONA STATE UNIVERSITY

August 2012

ABSTRACT

Identification of primary language impairment (PLI) in sequential bilingual children is challenging because of the interaction between PLI and second language (L2) proficiency. An important step in improving the accurate diagnosis of PLI in bilingual children is to investigate how differences in L2 performance are affected by a length of L2 exposure and how L2 assessment contributes to differentiation between children with and without PLI at different L2 proficiency levels.

Sixty one children with typical language development (TD) ages 5;3-8 years and 12 children with PLI ages 5;5-7;8 years participated. Results revealed that bilingual children with and without PLI, who had between 1 and 3 years of L2 exposure, did not differ in mean length of utterance (MLU), number of different words, percent of maze words, and performance on expressive and receptive grammatical tasks in L2. Performance on a grammaticality judgment task by children with and without PLI demonstrated the largest effect size, indicating that it may potentially contribute to identification of PLI in bilingual populations. In addition, children with PLI did not demonstrate any association between the length of exposure and L2 proficiency, suggesting that they do not develop their L2 proficiency in relation to length of exposure in the same manner as children with TD.

Results also indicated that comprehension of grammatical structures and expressive grammatical task in L2 may contribute to differentiation between the language ability groups at the low and intermediate-high proficiency levels. The discriminant analysis with the entire sample of bilingual children with and without PLI revealed that among L2 measures, only MLU contributed to the discrimination between the language ability groups. However, poor classification accuracy suggested that MLU alone is not a sufficient predictor of PLI.

There were significant differences among L2 proficiency levels in children with TD in MLU, number of different words, and performance on the expressive and receptive grammatical tasks in L2, indicating that L2 proficiency level may potentially impact the differentiation between language difficulties due to typical L2 acquisition processes and PLI.

DEDICATION

This dissertation is dedicated to my husband, Tim, whose love, patience, never-ending support, and encouragement provided me strength and perseverance during this journey; and to a little one on the way, who brought unconditional happiness and joy in our lives. I love you both very much.

ACKNOWLEDGMENTS

I am truly indebted and thankful to my dissertation committee members, Dr. Maria Adelaida Restrepo, Dr. Shelley Gray, and Dr. Joanna Gorin, for their guidance, support, and thoughtful input. Many thanks go to my mentor and dissertation chair, Dr. Restrepo, for her time, enthusiasm, constructive criticism, and encouragement that helped me to shape my thinking and writing as a researcher.

This research would never been completed without help from the members and volunteers in the Bilingual Language and Literacy Lab. I would like to thank Giovanna Moreno and Beatriz Barragan for their invaluable help with data collection and recruitment, Koren Anders and Penelope Thompson for transcribing numerous language samples and testing many children, Hyrum Borrego for scoring the experimental measures and transferring data to teleforms, Maria Kapantzoglou, Carol Mesa, and Yunely Ripley for their help with Spanish transcriptions, and Megan Figueroa for coding English transcriptions.

I would like to express my gratitude to the department of Speech and Hearing Science and Dr. Restrepo for their financial support during my doctoral studies. I would like to thank all the children and their parents who participated in this study.

Last, but not least I would like to thank my family for their love, emotional support, and unending encouragement, which kept me focused during my studies. You taught me the importance of education and inspired me to strive for the best in my work.

The data in this research were collected via support by grant R324A080024 awarded to Maria Adelaida Restrepo from the U. S Department of Education Institute of Education Science and by Graduate Research Grant awarded to Ekaterina Smyk from Graduate Professional Student Association at Arizona State University.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Introduction

Sequential bilingual children are those who learn a first language (L1) from birth and begin acquisition of a second language (L2) later in childhood (Kohnert, 2004). In the context of the U.S., the majority of sequential bilingual children acquire English as a L2. According to the U.S. Department of Education, these sequential bilinguals represented approximately 10.4% of the total public school student enrollment in the 2009-2010 school year (Aud et al., 2012). It is estimated that by 2030, 40% of the school population will speak English as a L2 (U.S. Department of Education & National Institute of Child Health and Human Development, 2003). L2 acquisition and bilingualism in the school population presents a challenge because of over- and under-identification of children with language impairment (Artiles, Rueda, Salazar, & Higareda, 2005; Artiles & Trent, 1994; MacSwan, 2000; MacSwan & Rolstad, 2006; Samson & Lesaux, 2009; Sullivan, 2011).

The diagnosis of primary language impairment (PLI)[1] in sequential bilinguals is challenging because of the limited availability of measures designed to identify PLI in bilingual learners (Bedore & Peña, 2008; Dollaghan & Horner, 2011; Peña & Bedore, 2009) and because of the interaction between PLI and L2 acquisition (Bedore, Peña, Gillam, & Ho, 2010; Bedore, Peña, Joyner, & Macken, 2011). PLI has been defined as impairment in language comprehension and/or language production due to an unidentified etiology that cannot be attributed to any evident hearing, neurological, emotional, cognitive problems, or a diagnosis of autism (Leonard, 1998; Schwartz, 2009). The prevalence of PLI in sequential bilinguals is unknown (Damico, Oller, & Storey, 1983; Kohnert, 2004; 2010); however, it is reasonable to assume that it is consistent with the overall 7% prevalence rate of PLI among monolingual children (Tomblin et al., 1997). In addition, bilingual children do not demonstrate an increased risk for language impairment when compared to monolingual children in each language (Peña, Gillam, Bedore, & Bohman, 2011). Nevertheless, sequential bilinguals are referred to special education services more often than their monolingual English-speaking peers and may be viewed as poor language learners (Artiles et al., 2005; Artiles & Trent, 1994; MacSwan, 2000; MacSwan & Rolstad, 2006; Samson & Lesaux, 2009; Sullivan, 2011). At the same time, sequential bilingual children with PLI can be under-identified when language difficulties due to PLI are mistaken as a developmental L2 acquisition process, resulting in a lack of appropriate special education services (Artiles et al., 2005; Crutchley, 1999; Samson & Lesaux, 2009). Inaccurate identification of PLI in sequential bilinguals creates a health discrepancy in this population and negatively impacts children's academic achievement and social interaction development. In addition, it results in inefficient use of the financial and personnel resources in the public school system.

When working with sequential bilingual children it is crucial to distinguish between language difficulties that are caused by a low level of linguistic attainment due to insufficient exposure to a particular language and language difficulties that are due to PLI. Language limitations in PLI are apparent to some extent in all of the languages that a bilingual person speaks (Kohnert, 2008;

Restrepo & Kruth, 2000; Willig, 1986) and thus, these limitations cannot be explained by patterns of language exposure and L2 processes. Conversely, L2 proficiency levels in children with typical language development (TD) cannot be attributable to any deficits in language knowledge or processing mechanisms due to atypical cognitive or neurological deficiencies (Kohnert, 2008).

Current measures used to identify PLI in monolingual children overlap to some extent with the measures used to identify L2 proficiency levels in bilingual children. For example, morphological development in English as a L2 is frequently measured by the accuracy of individual morphemes or groups of morphemes, such as past tense, third person singular, and auxiliary and copula [BE], in the obligatory contexts in language tasks (e.g., Bland-Stewart & Fitzgerald, 2001; Jia & Fuse, 2007; Marinis & Chondrogianni, 2010; Paradis, 2005, 2010). Measures of the same skill are used in identification of PLI in monolingual English-speaking children (e.g., Bedore & Leonard, 1998; Dunn, Flax, Sliwinski, & Aram, 1996; Leonard, Miller, & Gerber, 1999; Rice & Wexler, 1996) and bilingual children (Blom & Paradis, in press; Jacobson & Livert, 2010; Jacobson & Schwartz, 2005). Thus, when bilingual children are compared to monolingual children with PLI, any similarity in performance on identical measures challenges the diagnosis of PLI in bilingual children.

Currently there is limited evidence regarding differences in L2 proficiency in sequential bilingual children with and without PLI and the effects of PLI on L2 acquisition (Blom & Paradis, in press; Håkansson, Samaeh, & Nettlebladt, 2003; Verhoeven, Steenge, & van Balkom, 2011; Verhoeven, Steenge, van

Weerdenburg, & van Balkom, 2011), which in turn challenges the differentiation between language difficulties due to a typical L2 acquisition process and PLI. Despite the growing number of studies examining PLI in bilingual populations, most studies compare the performance between bilingual children with TD and monolingual children with and without PLI (Crutchley, Conti-Ramsden, & Botting, 1997; Håkansson, 2001; Håkansson & Nettlebladt, 1996; Kohnert, Windsor, & Yim, 2006; Paradis, 2005; Paradis, Crago, Genesee, & Rice, 2003; Paradis, Rice, Crago, & Marquis, 2008; Windsor & Kohnert, 2004). However, these studies can be confounded by the differences in language developmental levels as well as differences in language use and experience patterns between bilingual and monolingual language learners; thus, the effects of PLI on L2 development remain unclear. In addition, Grosjean (1989) argued that bilinguals should not be considered as two monolinguals in one because bilingual learners represent a unique population with their own language features and behaviors that set them apart from monolingual speakers. Comparisons of language performance in sequential bilingual children with and without PLI will help to better understand the interaction between PLI and L2 proficiency.

The purpose of the study was to examine the effects of PLI on L2 proficiency in sequential bilingual children and the contribution of L2 assessment to differentiation between children with and without PLI. The introductory part of the dissertation is structured as follows. First, I define L2 proficiency and language ability and discuss Dynamic System Theory (DST) as a theoretical framework of L2 proficiency and PLI. Second, I present available research on L2 proficiency components in bilingual children with and without PLI. Third, I discuss how the length of L2 exposure and L2 proficiency level affect L2 performance in bilingual children with and without PLI. I conclude the summary with the goals of the study.

Dynamic System Theory: L2 Proficiency and PLI

Definition of L2 proficiency and language ability. The terms *language ability* and *language proficiency* are frequently used interchangeably in different contexts as the umbrella terms for the continuum between proficient and non-proficient speakers of L2 (Bachman, 1990; Bachman & Cohen, 1998; Baker, 2006; Herdina & Jessner, 2002; Kohnert, 2008). In addition, these terms are used interchangeably in relation to children with and without PLI, which complicates the differentiation between the two concepts (Kester Stubbe & Peña, 2002; Leonard, 1998; Tilstra & McMaster, 2007). In the present paper, language proficiency is the phenomenon of L2 acquisition that indicates a level of linguistic attainment in a particular language (Bedore et al., 2011), while language ability is the general language skills that are necessary for using a complex language system effectively. In this study atypical language ability is referred to as PLI, while typical language ability is referred to as typical development (TD).

L2 proficiency is a constructive, input-driven psychological phenomenon of L2 acquisition that relies on and interacts with cognitive skills (Bialystok, 2001; Butler & Hakuta, 2004). Language proficiency is considered language knowledge obtained through experience with that language and the ability to implement the language knowledge to a specified level of performance in a situation defined by cognitive and linguistic demands (Bialystok, 2001). Thus, the definition of language proficiency is based on a functional linguistic approach to L2 acquisition (Canale & Swain, 1980; Bachman, 1990). Specifically, in this study language proficiency is defined as a continuum of oral language skills between non-proficiency and native-like proficiency in L2.

Dynamic systems theory. DST (de Bot, 2008; de Bot, Lowie, & Verspoor, 2007; Fogel & Thelen, 1987; Herdina & Jessner, 2002; Smith & Thelen, 2003; van Geert, 2008) offers a unified framework for describing and explaining variability and non-linear, continuous change in L1 and L2 development in children with and without PLI with rapid accelerations, slow changes, and possible declines or relative plateaus in processing efficiency (de Bot, 2008; de Bot et al., 2007; Evans, 2001, 2002; Larsen-Freeman, 1997; Spoelman & Verspoor, 2010; van Dijk & van Geert, 2005; van Geert, 2004, 2008). Given that DST describes language as a complex dynamic system, language proficiency and the effects of PLI on L1 and L2 can be described under the same theoretical umbrella.

According to DST, L2 proficiency is characterized as a complex dynamic system comprised of multiple interrelated components that interact with each other and the environment (Herdina & Jessner, 2002; Larsen-Freeman, 2006; van Geert, 2008). Furthermore, L2 proficiency components cannot be considered constant along the language proficiency continuum. They represent multifaceted and multicomponential constructs with different roles at different proficiency levels and communication contexts (Norris & Ortega, 2009). It is assumed that a proficient speaker displays advanced language skills by producing complex, accurate, and fluent L2 performance. However, Skehan (1998) argued that only in ideal conditions of communication can all the areas of performance be achieved simultaneously. Actual performance may depend on prioritizing of a particular area and on the characteristics of the tasks and circumstances under which tasks were performed (Skehan & Foster, 1999). For example, research on sequential bilingual children indicates that a story retell task is likely to elicit longer utterances and more complex grammatical structures than spontaneous conversation or story generation (e.g., Gazella & Stockman, 2003; Gutiérrez-Clellen, 2002; Restrepo & Gutiérrez-Clellen, 2001). Therefore, the assessment of L2 proficiency should be based on a range of language components rather than a single task (Larsen-Freeman, 2006).

As a dynamic system, L2 proficiency changes due to the interaction between L1 and the L2 components, environmental factors, and available resources (e.g., individual differences in the ability to learn L2 and motivation; de Bot, 2008; Larsen-Freeman, 1997; Smith & Thelen, 2003; van Geert, 2004, 2008). Different environmental factors, such as parental level of education, quality of L2 input, length and age of L2 exposure, and the quantity of L2 use at home have been found to affect the development of L2 proficiency (Armon-Lotem, Walters, & Gagarina, 2011; Bohman, Bedore, Peña, Mendez-Pérez, & Gillam, 2010; Chondrogianni & Marinis, 2011; Golberg, Paradis, & Crago, 2008; Gutiérrez-Clellen, Simon-Cereijido, & Sweet, 2012; Oller & Eilers, 2002; Paradis, 2011). For example, the age of L2 exposure can be positively associated with the rate of vocabulary development, with later L2 exposure resulting in faster vocabulary gains than early L2 exposure (Chondrogianni & Marinis, 2011; Golberg et al., 2008). Conversely, the age of exposure can be negatively associated with the rate of acquisition of morphosyntax, with early L2 exposure resulting in more accurate production of grammatical morphology than later L2 exposure (Jia & Aaronson, 2003; Jia & Fuse, 2007).

Research demonstrates that sequential bilingual children represent a heterogeneous group with considerable variability in language performance that is driven by various factors described above (de Bot et al., 2007; Bedore et al., 2010; Herdina & Jessner, 2002; Kohnert, Windsor, & Ebert, 2009; Larsen-Freeman, 2006; Verspoor, Lowie, & van Dijk, 2008). These children may not exhibit a consistent pattern of one language being stronger than the other. Further, the impact of learning a L2 may cause a shift of language dominance from a L1 to a L2, which is still being developed (Anderson, 2004; Francis, 2005; Jia, Kohnert, Collado, & Aquino-Garcia, 2006; Kan & Kohnert, 2005; Kohnert, 2004; Kohnert, Bates, & Hernandez, 1999), and slow down growth in L1. Thus, variation in language performance is an inherent characteristic of the dynamic system that emerges from the interactions between L2 components, L1, and the available resources, including linguistic environment and language ability (de Bot et al., 2007; Verspoor et al., 2008).

PLI also demonstrates the properties of a dynamic complex system (Evans, 1996; van Geert, 2004). Children with PLI exhibit a variety of linguistic deficits with the salient features changing across development (see Leonard, 1998; Schwartz, 2009, for a review), which is an expected property of a dynamic system. From a DST perspective, the heterogeneous nature of PLI is the result of developmental processes, such as in language proficiency, determined by the interaction between different language components, available resources, and the environment, rather than the result of a direct manifestation of missing or impaired linguistic structures (Evans, 2001; van Geert, 2004). Thus, the course of L2 development in bilingual children with PLI, and consequently symptoms of PLI over time, are not linear, but dynamic in nature.

There is considerable variability in the manifestation of deficits across languages, although certain symptoms of PLI cluster in the area of morphosyntactic and lexical skills (Kohnert et al., 2009; Restrepo & Gutiérrez-Clellen, 2001). Cross-linguistic studies reveal that the extent to which grammatical morphology is affected varies across languages (e.g., Bedore & Leonard, 2001; Leonard, Bortolini, Caselli, McGregor, & Sabbadini, 1992; Oetting & Rice, 1993; Paradis & Crago, 2001; Rice, Noll, & Grimm, 1997; Rice & Wexler, 1996). Different patterns of impairment have been reported, even among languages that belong to the same language family (e.g., Crago & Paradis, 2004; Le Normand, Leonard, & McGregor, 1993). Further, variability in the expression of PLI in the same language has been found when there are different ages and sociolinguistic contexts (Restrepo & Gutiérrez-Clellen, 2012).

In conclusion, L2 proficiency in children with and without PLI is a complex dynamic system that comprises multiple components that demonstrate non-linear continuous development, depending on the interaction between its components, environmental factors, and available resources. As a result, sequential bilingual children demonstrate considerable variability in L2 performance that is driven by many factors, including but not limited to the quantity of L2 input, age and length of L2 acquisition, language dominance, and context of L2 acquisition, as well as L1 resources, development and language ability. Therefore, according to DST, assessment of L2 proficiency and PLI cannot be based on a single task, but rather it should include assessment of a range of L1 and L2 components. Furthermore, it is necessary to consider whether the differences in L2 performance are driven by the differences in the L2 acquisition patterns and environmental factors or language ability.

L2 Proficiency Components in Children with and without PLI

Although there is agreement that L2 proficiency is a complex system consisting of different linguistic and processing components, there is no agreement on the specific components required for L2 proficiency (e.g., Bachman & Palmer, 1982; Canale & Swain, 1980; McNamara, 1996; Skehan, 1998). Nevertheless, in L2 acquisition research, syntactic complexity, grammatical accuracy, verbal fluency, and lexical diversity are frequently used to describe the continuum of L2 proficiency (e.g., Ellis & Barkhuizen, 2005; Iwashita, 2010; Iwashita, Brown, McNamara, & O'Hagan, 2008; Norris & Ortega, 2003, 2009; Skehan, 2009; Larsen-Freeman, 2006), with development in these language components being important indicators of L2 proficiency level. Furthermore, these components have been found to contribute to differentiation between monolingual children with and without PLI, including syntactic complexity (e.g., Eisenberg, McGovern Fersko, & Lundgren, 2001; Simon-Cereijido & Gutiérrez-Clellen, 2007; Restrepo, 1998), grammatical accuracy (e.g., Bedore & Leonard, 1998; Restrepo, 1998), and lexical diversity (e.g., Hewitt, Hammer, Yont, & Tomblin, 2005; Leonard et al., 1999; Watkins, Kelly, Harbers, & Hollins, 1995). Such similarities often confound the differentiation and diagnosis of PLI in sequential bilingual children.

Syntactic complexity. Syntactic complexity reflects length, elaborateness and diversity of verbal utterances produced by a language learner (Housen & Kuiken, 2009; Norris & Ortega, 2009; Pallotti, 2009). One of the most extensively used measures of syntactic complexity is the length of linguistic units measured by mean length of utterance (MLU) in words or morphemes (e.g., Brown, 1973; Hickey, 1991; Thordardottir & Weismer, 1998) or by the number of words per terminable unit (T-unit; Hunt, 1970). Other measures of syntactic complexity include the subordination index (Givon, 1991) or the total frequency of certain syntactic forms (Ellis & Yuan, 2005; Robinson, 2007) to account for elaborateness of syntactic structures.

Syntactic complexity, as measured by MLU or the number of words per T-unit and the number of clauses per T-unit, is positively associated with oral L2 proficiency in young adults (Halleck, 1995; Iwashita, 2006; Iwashita et al., 2008). Sequential bilingual children with TD ages 4-8 years at lower levels of L2 proficiency have shorter MLUs and fewer subordinate clauses in comparison to children at higher levels of proficiency (Smyk, Restrepo, Gorin, & Gray, in preparation). In addition, sequential bilingual children with TD demonstrate an increase in MLU in English as a L2 from kindergarten to second grade (Miller et al., 2006; Muñoz, Gillam, Peña, & Gulley-Faehnle, 2003; Rojas, 2011; Swasey Washington, 2010).

MLU differentiates children with TD and PLI in monolingual (e.g., Goffman & Leonard, 2000; Rice, Redmond, & Hoffman, 2006; Rice et al., 2010; Hewitt et al., 2005) and bilingual populations (e.g., Simon-Cereijido & Gutiérrez-Clellen; 2009; Restrepo & Kruth, 2000). Further, Bedore at al. (2010) reported that MLU and grammaticality in English contributed to differentiation between bilingual children with and without PLI that were initially diagnosed on the standardized measure of language ability. These results suggest that MLU in English as a L2 may contribute to discrimination between bilingual children with and without PLI.

Grammatical accuracy. Grammatical accuracy is the degree of deviation in the use of grammatical structures from the community norm for that language (Housen & Kuiken, 2009; Pallotti, 2009). Deviations are considered as correct or incorrect, with the latter indicating grammatical errors. Measures of grammatical accuracy include the percentage of grammatical utterances (Dunn et al., 1996; Fiestas & Peña, 2004; Muñoz et al., 2003; Verhoeven, Steenge, & van Balkom, 2011) and number of errors per utterance in a language sample (Restrepo, 1998; Restrepo et al., 2010). Further, grammatical accuracy is measured by the percentage of correct use of individual morphemes (e.g., Jia & Fuse, 2007; Paradis, 2005) or groups of morphemes (i.e., tense vs. non-tense) in obligatory contexts in spontaneous or elicited language production tasks (e.g., Bedore &

Leonard, 1998; Leonard et al., 1999; Polite & Leonard, 2006). In addition, grammatical accuracy can also be assessed by receptive tasks, such as comprehension of grammatical contrasts marked by morphemes, function words, and word order (e.g., Bishop & Adams, 1992; Bishop, Bright, James, Bishop, & van der Lely, 2000; Marinis & Chondrogianni, 2010), and grammaticality judgment tasks, in which a participant is asked to judge whether a sentence was grammatical or ungrammatical (e.g., McDonald & Roussel, 2010; Paradis et al., 2008; Rice, Wexler, & Redmond, 1999; Saxton, Dockrell, Bevan, & van Herwegen, 2007; van der Lely, Jones, & Marshall, 2011).

Grammatical accuracy increases as a function of L2 proficiency (e.g., Halleck, 1995; Smyk et al., in preparation). Smyk et al. (in preparation) reported that bilingual children at low L2 proficiency levels produced only 17% grammatically correct utterances, while children at higher L2 proficiency levels produced about 73% grammatically correct utterances in a story retelling task. Further, Muñoz et al. (2003) reported that predominantly English-speaking bilingual children ages 3;10-4;8 produced on average 59% grammatically correct utterances in a story telling task in comparison with children ages 5-5;6 who produced 80% grammatically correct utterances in the same task. Similarly, Fiestas and Peña (2004) reported that Spanish-English bilingual children ages 4-6;11, who were fluent speakers of English, had approximately 80% grammatical utterances in narratives elicited by a wordless picture book.

Bilingual children with PLI demonstrate deficits in tense morphology in English as a L2 (Blom & Paradis, in press; Jacobson & Livert, 2010; Jacobson &

Schwartz, 2005; Restrepo & Kruth, 2000). Using a verb morphology composite score from Bedore and Leonard (1998), Gutiérrez-Clellen and Simon-Cereijido (2009) reported that a cutoff verb composite score of 80% yielded sensitivity of 91% in identification of English-dominant bilingual children with PLI. Jacobson and Schwartz (2005) investigated the discrimination accuracy of the past tense production task in bilingual children with and without PLI. They reported sensitivity of 92% and specificity of 87% for regular past tense, sensitivity of 92% and specificity of 73% for irregular past tense, and sensitivity of 100% and specificity of 73% for novel verbs that have phonological similarity to regular and irregular past tense verbs in English. Therefore, grammatical morphology, specifically tense markers in English as a L2, should be a sensitive indicator of PLI in bilingual children.

Research on performance of bilingual children with and without PLI on receptive grammatical tasks is limited, although performance of monolingual and bilingual children indicates that it may differentiate between children with and without PLI. For example, Verhoeven, Steenge, van Weerdenburg, et al. (2011) reported that 7-8 year old bilingual children with PLI learning Dutch as a L2 performed significantly lower on a measure of comprehension of function words than bilingual children with TD. Studies of monolingual children with PLI demonstrate that children with PLI perform significantly lower than children with TD on a grammaticality judgment task (e.g., Rice et al., 1999; Rice, Hoffman, & Wexler, 2009; Saxton et al., 2007) and on a comprehension of grammatical structures task (e.g., Bishop & Adams, 1992; Norbury, Bishop, & Briscoe, 2002).

Thus, it is possible that receptive grammatical tasks may contribute to identification of PLI in bilingual children, given that bilingual children catch up to monolingual levels on these tasks faster than on expressive language tasks (Marinis & Chondrogianni, 2010).

Verbal fluency. Verbal fluency is a learners' control over their language knowledge manifested by temporal variables such as speech rate, pauses, false starts, reformulations, and repetitions (e.g., Housen & Kuiken, 2009; Lennon 1990; Riazantseva, 2001; Schmidt, 1992). Fluency is a multicomponential phenomenon in which breakdowns, such as a number, length, and distribution of pauses, and repairs, such as the number of false starts, reformulations, and repetition, and rate of delivery, can be measured separately (Foster & Tavakoli, 2009; Pallotti, 2009; Skehan, 2009; Tavakoli & Skehan, 2005).

Language proficiency has been found to affect verbal fluency in terms of length and frequency of pauses (Riazantseva, 2001) and speech rate measured by the total number of syllables per minute (Ginther, Dimova, & Yang, 2010; Kormos & Denes, 2004). Fiestas, Bedore, Peña, and Nagy (2005) reported that sequential bilingual children ages 4-7 years had less fluent speech as indicated by a higher rate of repetitions in the sound, word, and phrase level in story narratives in English as a L2 in comparison with functionally dominant English-speaking children. They also found that bilingual children produced more lexical than grammatical revisions in English as a L2. Smyk at al. (in preparation) found that children at lower proficiency levels produced a ratio of 40% maze words (i.e., repetitions, revisions, and fillers) to total number of words, while children at higher proficiency levels demonstrated a ratio of 13% maze words in a story retelling task.

Currently, no studies have investigated whether verbal fluency can be a sensitive indicator of PLI in bilingual children. Limited research on monolingual children with PLI demonstrates that the production of maze words differs in children with and without PLI (e.g., Leadholm & Miller; 1995; Navarro-Ruiz & Rallo-Fabra, 2001; Thordardottir & Weismer, 2002). Leadholm and Miller (1995) reported that children with PLI tend to produce more than one maze per utterance and 20-25% of mazes out of the total productions, with mazes longer than 3-4 words. Navarro-Ruiz and Rallo-Fabra (2001) found that children with PLI were more likely to use hesitations and repetitions, and less likely to use revisions that children with TD. In contrast, Thordardottir and Weismer (2002) reported that children with and without PLI produced more revisions than filled pauses; however, the difference between two types of mazes was larger for children with PLI. Overall, verbal fluency may also differ in bilingual children with and without PLI and may contribute to the discrimination between the language ability groups.

Lexical diversity. Lexical diversity is defined as the variety of words produced by a speaker, with a greater number indicating a greater lexical diversity and higher L2 proficiency levels (e.g., Johansson, 2008; McCarthy & Jarvis, 2010; Owen & Leonard, 2002). Measures of lexical diversity are calculated using spontaneous or elicited language samples and may include a number of different words (NDW), type-token ratio (TTR), and *D* (Duran, Malvern, Richards, & Chipere, 2004; Golberg et al., 2008; Klee, 1992; Muñoz et al., 2003; Owen &

Leonard, 2002; Richards, 1987; Watkins et al., 1995; Uccelli & Paez, 2007; Vermeer, 2000). TTR is calculated by comparing NDW to total number of words (word tokens). *D* is based on the comparison of the mathematical curve of TTR with a repeated calculation of TTR over a range of word tokens from a language sample. This comparison yields the *D* score, with higher values indicating greater lexical diversity.

Lexical diversity has been found to predict L2 proficiency (Iwashita et al., 2008; Iwashita, 2010; Zareva, Schwanenflugel, & Nikolova, 2005; Yu, 2010). Research indicates that the NDW in narratives and semistructured interviews is a sensitive developmental measure for Spanish-English bilingual children, with older and more proficient children demonstrating a greater variety of words than younger and less proficient children (Golberg et al., 2008; Miller et al., 2006; Muñoz et al., 2003; Uccelli & Paez, 2007). Smyk et al. (in preparation) found children at lower L2 proficiency levels produced less NDW (mean = 24) than children at higher L2 proficiency levels (mean = 72).

Currently it is unclear whether lexical diversity measures in L2 are potentially sensitive to discriminating bilingual children with and without PLI. Bedore at al. (2010) reported that NDW in L1 and L2 did not predict significantly language ability scores based on a standardized measure developed for identification of PLI in bilingual children. However, other studies on bilingual children with PLI found lexical deficits in L2 (Restrepo & Kruth, 2000; Simon-Cereijido & Gutiérrez-Clellen, 2009; Thordardottir, Weismer, & Smith, 1997; Verhoeven, Steenge, van Weerdenburg, et al., 2011). Similarly, monolingual

English-speaking children with PLI demonstrate lower NDW in spontaneous language samples than children with TD ages 3-6;7 years (Hewitt et al., 2005; Leonard et al., 1999; Watkins, et al., 1995). Similar findings were reported for 4-5 year old Cantonese-speaking children with and without PLI (Wong, Klee, Stokes, Fletcher, & Leonard, 2010). Therefore, lexical diversity in L2 can potentially contribute to the identification of PLI in bilingual children.

Effects of Length of L2 exposure and L2 Proficiency Level on the Differences between Bilingual Children with and without PLI

Research indicates that bilingual children with and without PLI demonstrate differences in L2 performance on language sample measures and morphology tasks in different languages (in English: Blom & Paradis, in press; Gutiérrez-Clellen, Simon-Cereijido, & Wagner, 2008; Jacobson & Livert, 2010; Jacobson & Schwartz, 2005; Restrepo & Kruth, 2000; Paradis, 2008; in Arabic: Salameh, Håkansson, & Nettlebladt, 2004; in Dutch: Orgassa & Weerman, 2008; Verhoeven, Steenge, & van Balkom, 2011; Verhoeven, Steenge, van Weerdenburg, et al., 2011). However, most studies do not examine how differences in L2 proficiency are affected by an increased length of L2 exposure and how L2 performance contributes to discrimination between bilingual children with and without PLI at different language proficiency levels.

Studies on bilingual children with TD and monolingual children with and without PLI suggest that the length of L2 exposure influences L2 performance (e.g., Marinis & Chondrogianni, 2010; Paradis, 2005; Paradis & Crago, 2000). Specifically, bilingual children with TD who had up to two years of L2 exposure demonstrate similarities in performance on grammatical morphology tasks to monolingual children with PLI, indicating that bilingual children struggle with acquisition of grammar at the earlier stages of L2 learning, which can lead to misdiagnosis as PLI (e.g., in English as a L2: Paradis, 2005; in French as a L2: Paradis & Crago, 2000). Further, Marinis and Chondrogianni (2010) found that bilingual children with TD who had at least four years of English as a L2 exposure were as accurate as monolingual children with TD in the production of past tense, but not in 3rd person singular [-s], indicating that past tense can contribute to identification of PLI in bilingual children with longer period of L2 exposure.

Studies that examined the effect of L2 exposure on differences in L2 performance between in bilingual children with and without PLI are sparse. Preliminary research indicates that the effects of PLI on L2 proficiency may not be evident at earlier stages of L2 acquisition (Gutiérrez-Clellen et al., 2012; Paradis, 2008; Smyk, Restrepo, Gorin, & Gray, 2011; Verhoeven, Steenge, & van Balkom, 2011; Verhoeven, Steenge, van Weerderburg, et al., 2011). Specifically, Paradis (2008) found that after two years of L2 exposure bilingual children with PLI scored at least 1.5 standard deviation below of the bilingual children with TD on a tense morpheme task that included assessment of regular and irregular past tense, third person singular [-s], auxiliary and copula [BE], and auxiliary [DO]. Smyk et al. (2011) reported differences between bilingual children with and without PLI in L2 proficiency assessed by a rating scale only after two years of English exposure. Further, Gutiérrez-Clellen et al. (2012) reported that the number of hours of L2 exposure in the classroom may not produce noticeable affects on L2 growth in sequential bilingual preschoolers with PLI at the earlier stages of L2 acquisition. In addition, Verhoeven, Steenge, van Weerdenburg, et al. (2011) reported that the effects of PLI are not constant across ages, with a greater effect of PLI on L2 proficiency in older bilingual children than in younger children.

It is hypothesized that exposure to L2 is positively associated with the L2 proficiency level; however, other factors, such as the quantity and quality of L2 input, language of instruction, and the language environment, significantly influence the L2 proficiency level (e.g., Dixon, 2011; Scheele, Laseman, & Mayo, 2010). Thus, children who are grouped only on the length of L2 exposure may still demonstrate different levels of language proficiency. Therefore, it is necessary to examine how differences in L2 proficiency are affected by the length of L2 exposure in bilingual children with and without PLI and how English language proficiency level impacts differentiation between these two groups of L2 learners.

Summary

Identification of PLI in bilingual children and differentiation between language difficulties due to typical L2 acquisition processes and PLI is challenging because of the interaction between PLI and L2 proficiency. An important step in improving the accurate diagnosis of PLI in bilingual children is to investigate how PLI affects L2 performance in relation to length of L2 exposure and L2 proficiency levels. Research indicates that the effects of PLI on

L2 proficiency may not be evident at earlier stages of L2 acquisition (Paradis, 2008; Smyk et al., 2011); however, most studies on bilingual children with and without PLI do not consistently control for L2 exposure patterns (Crutchley et al., 1997; Verhoeven , Steenge, & van Balkom, 2011; Verhoeven, Steenge, van Weerderburg, et al., 2011), which may influence the differences in overall L2 proficiency levels and limit the generalization of the effects of PLI on L2 proficiency. Further, children grouped only on the length of L2 exposure may still demonstrate different L2 proficiency levels. Therefore, a systematic examination is needed to determine how differences in L2 performance between bilingual children with and without PLI are affected by length of L2 exposure and how L2 proficiency level impacts differentiation between the language ability groups. This will help clinicians understand how L2 assessment contributes to identification of PLI in bilingual populations.

The goals of the present project were two-fold: (1) to investigate, cross-sectionally, if there are differences in L2 proficiency components in sequential bilingual Spanish-English speaking children with and without PLI due to the length of L2 exposure; (2) determine the best set of predictors among L2 proficiency components that differentiate between sequential bilingual Spanish-English speaking children with and without PLI at the low, intermediate, and high L2 proficiency levels.

Second Language Proficiency in Sequential Bilingual Children with and without Primary Language Impairment

Ekaterina Smyk

Arizona State University

Author Note

The data in this research were collected via support by grant R324A080024 from the U. S Department of Education Institute of Education Science and by Graduate Research Award from Graduate Professional Student Association at Arizona State University.

Second Language Proficiency in Sequential Bilingual Children with and without Primary Language Impairment

According to the National Clearinghouse for English Language Acquisition (NCELA, 2010), sequential bilingual children learning English as a second language (L2) represent a fast growing population in U.S. public schools. Currently there is a problem of misdiagnosis of Primary Language Impairment (PLI) in bilingual children (Bedore & Peña, 2008; Dollaghan & Horner, 2011). Children can be over-identified with PLI and unnecessarily referred to special education services or be under-identified, resulting in a lack of appropriate special education services (Artiles, Rueda, Salazar, & Higareda, 2005; MacSwan & Rolstad, 2006; Sullivan, 2011).

One of the challenges in identifying PLI in sequential bilinguals comes from a possible interaction between PLI and L2 acquisition (Bedore, Peña, Gillam, & Ho, 2010; Bedore, Peña, Joyner, & Macken, 2011). This interaction challenges the differentiation between language difficulties that are due to the typical process of L2 acquisition and language difficulties that are due to impaired language system. Despite the growing number of studies examining PLI in bilingual populations, most studies compare the performance of bilingual children with TD and monolingual children with and without PLI (Crutchley, Conti-Ramsden, & Botting, 1997; Håkansson, 2001; Håkansson & Nettlebladt, 1996; Kohnert, Windsor, & Yim, 2006; Paradis, 2005; Paradis, Crago, Genesee, & Rice, 2003; Paradis, Rice, Crago, & Marquis, 2008; Windsor & Kohnert, 2004). However, comparisons between bilingual and monolingual children are confounded by the differences in developmental language levels and in language use and acquisition patterns. One of the important steps in improving the diagnostic accuracy in identification of bilingual children with PLI is to understand the effects of PLI on L2 proficiency, which requires comparing L2 performance in bilingual children with and without PLI.

L2 Performance in Bilingual Children with and without PLI

Language deficits in bilingual children with PLI are somewhat similar to language deficits in monolingual children with PLI, with lexical and grammatical skills being the most affected (Blom & Paradis, in press; Cleave, Girolametto, Chen, & Johnson, 2010; Crutchley et al., 1997; Jacobson & Schwartz, 2005; Peña & Bedore, 2009; Restrepo & Kruth, 2000). Research indicates that there are differences in L2 performance between bilingual children with and without PLI on language sample measures and morphology tasks in different languages (in English: Blom & Paradis, in press; Gutiérrez-Clellen, Simon-Cereijido, & Wagner, 2008; Jacobson & Livert, 2010; Jacobson & Schwartz, 2005; Restrepo & Kruth, 2000; Paradis, 2008; in Arabic: Salameh, Håkansson, & Nettlebladt, 2004; in Dutch: Orgassa & Weerman, 2008; Verhoeven, Steenge, & van Balkom, 2011; Verhoeven, Steenge, van Weerdenburg, & van Balkom, 2011).

In English as a L2 bilingual children with PLI are less accurate in production of tense morphemes, such as past tense, third person singular [-s], auxiliary and copula [BE] in present and past tense, and auxiliary [DO] in present tense, than bilingual children with TD (Blom & Paradis, in press; Gutiérrez-Clellen et al., 2008; Jacobson & Livert, 2010; Jacobson & Schwartz, 2005;

Restrepo & Kruth, 2000). In addition, Restrepo and Kruth (2000) reported that a bilingual child with PLI demonstrated a shorter mean length of sentences, a limited repertoire of sentence types, prepositions, and pronouns, and a reduced number of different verbs and verb type-token ratio in comparison with a bilingual age-matched peer with TD.

Research demonstrates the extent to which bilingual children with and without PLI may differ on a range of language measures as a function of chronological age (Verhoeven, Steenge, & van Balkom, 2011; Verhoeven, Steenge, van Weerderburg, et al., 2011). Verhoeven, Steenge, van Weerderburg, et al. (2011) found that the six-year-old bilinguals with PLI learning Dutch as a L2 performed significantly lower than bilinguals with TD on auditory discrimination, articulation, and sentence imitation tasks. However, the seven-year-old children with PLI differed from bilingual age peers with TD on a wider range of measures, including auditory discrimination, articulation, sentence imitation, word definitions, morphology, comprehension of function words, and sentence comprehension tasks. The eight-year old children with PLI did not differ from bilingual age peers with TD on the auditory discrimination task, but continued to demonstrate poorer performance on all other tasks. In addition, the eight-year-old children with PLI performed significantly worse on the receptive vocabulary task, which did not differ in younger bilinguals with and without PLI. Differences between older bilingual children with and without PLI learning Dutch as a L2 were also reported for mean length of utterance (MLU) in words and the proportion of ungrammatical sentences in narratives elicited by a wordless picture book (Verhoeven, Steenge, & van Balkom, 2011). Overall, these results indicate that the effects of PLI are not constant across ages, with a greater effect of PLI on L2 proficiency in older bilingual children than in younger children. However, these studies were conducted cross-sectionally, and thus, other factors, such as the age of L2 exposure and the quantity and quality of language input, may have impacted the differences in L2 proficiency across ages and ability groups.

Overall, research on bilingual children with and without PLI has been focused on productive grammatical morphology in L2 (Gutiérrez-Clellen et al., 2008; Jacobson & Livert, 2010; Jacobson & Schwartz, 2005; Restrepo & Kruth, 2000), and thus, the extent to which bilingual children with PLI differ from their bilingual peers with TD on other L2 proficiency components, such as lexical diversity, verbal fluency, syntactic complexity, and receptive grammatical tasks, remains unclear. In addition, the heterogeneous nature of PLI is amplified by the variability in bilingual children's language performance that is driven by the quantity of L2 input, L2 proficiency, language dominance, and context of L2 acquisition. Therefore, a systematic examination of various L2 components is necessary to capture differences in language performance between bilingual children with and without PLI.

PLI and L2 Exposure

Research on the effects of PLI on L2 proficiency in relation to the length of L2 exposure suggests that it is difficult to determine the differences between bilingual children with and without PLI before two years of L2 exposure (Paradis, 2008; Smyk, Restrepo, Gorin, & Gray, 2011). In a cross-sectional study, Smyk et al. (2011) found significant differences between bilingual children with and without PLI on L2 proficiency scores after two years of English exposure. They examined the impact of PLI in three age groups (5-6, 6-7, and 7-8 years) who had English language exposure at the age of 5, according to parent report. Children's story retells were scored on sentence length, verbal fluency, grammaticality, and vocabulary, and then were given a global rating score. They reported significant group differences in L2 performance between bilingual children with and without PLI only at ages 7-8 years, indicating that the differences in L2 proficiency between the language ability groups were manifested after two years of exposure to English. Thus, based on a global rating scale, it is difficult to determine the effects of PLI on L2 proficiency after a short period of L2 exposure.

In a longitudinal study, Paradis (2008) examined the effects of PLI on production of tense morphemes in L2 after one, two, and three years of L2 exposure when comparing nine bilingual children with TD and two bilingual children with language impairment (LI). Specifically, she examined the performance on a composite tense score in an elicited language task of English tense morphemes including regular and irregular past tense, third person singular [-s], auxiliary and copula [BE], and auxiliary [DO]. At the beginning of data collection the mean age of participants was 5;4 (years; months) and at the end the mean age was 7;1. After one year of exposure to English, bilingual children with LI performed within one standard deviation of the bilingual group with TD. However, after two years of exposure to English, the discrepancy between bilingual children with and without LI became more pronounced, with the children with LI scoring at least 1.5 standard deviation below children with TD. After three years of exposure to English, one bilingual child with LI demonstrated an increase in performance, scoring within one standard deviation of children with TD, while the other child still lagged behind children with TD, scoring 1.3 standard deviations below the mean. These mixed results on performance of bilingual children with LI may have been due to differences in the degree of severity of LI and to large variability in performance reported for bilingual children with LI (e.g., Gutiérrez-Clellen et al., 2008).

In a longitudinal study on sequential bilingual preschoolers with PLI, Gutiérrez-Clellen, Simon-Cereijido, and Sweet (2012) found that the frequency of L2 use by children in the classroom predicted the growth rate in L2, while L2 proficiency and the number of hours of L2 exposure in the classroom did not. Although the authors examined the quantity of L2 exposure rather than its length, the majority of the children were at the earlier stages of L2 acquisition given that they mostly spoke Spanish at home and had not been exposed to English instruction before preschool, with the exception of 13 children who attended preschool at three years of age. Thus, the study indicated that L2 exposure at the earlier stages of L2 acquisition may not produce noticeable affects on L2 growth in bilingual children with PLI. However, these results should be interpreted with caution because all children attended bilingual classrooms in which the quantity of exposure in L1 and L2 was almost equal during instructional time. In addition, L2 proficiency was assessed only through parent report and it is unclear whether parents were proficient in English to judge their children's L2 proficiency

(Gutiérrez-Clellen & Kreiter, 2003). Moreover, L2 proficiency, use, and exposure were quantified using Likert-type measures that may not be sensitive to differences in L2 acquisition.

Overall, it is unclear how the differences in L2 language proficiency between bilingual children with and without PLI change with increasing length of L2 exposure. The few available studies on the effects of PLI on L2 proficiency indicate that differences in L2 performance between the language ability groups may not be evident at the earlier stages of L2 acquisition. Understanding how PLI affects L2 proficiency in relation to the length of L2 exposure will improve the accuracy in identification of PLI in bilingual children.

PLI and L2 Proficiency Level

Increased exposure to English as a L2 is associated with increases in L2 proficiency levels, with greater exposure indicating greater proficiency (e.g., Bohman, Bedore, Peña, Mendez-Pérez, & Gillam, 2010; Oller & Eilers, 2002). Therefore, it is hypothesized that it is difficult to determine the effects of PLI on L2 performance when children are at the lower L2 proficiency levels. However, the quantity and quality of L2 input, language of instruction, and language environment, can significantly influence L2 proficiency (De Houwer, 2007; Dixon, 2011; Duursma, Romero-Contreras, Szuber, Proctor, & Snow, 2007; Mancilla-Martinez & Lesaux, 2011; Scheele, Laseman, & Mayo, 2010; Umbel, Pearson, Fernandez, & Oller, 1992). Therefore, children grouped only on the length of exposure to the L2 may demonstrate different levels of L2 proficiency. Given the dynamic nature of L2 proficiency and variability in L2 acquisition outcomes that change with age and language experiences (e.g., de Bot, Lowie, & Verspoor, 2007; Freman-Larseen, 2006; Herdina & Jessner, 2002; Jessner, 2008; Kohnert, 2010; Kohnert, Windsor, & Ebert, 2009; Plaza-Pust, 2008; Schiff-Myers, 1992), it is important to examine whether the effects of PLI on L2 vary as a function of the overall proficiency level, impacting the differentiation between bilingual children with and without PLI.

Currently no studies on bilingual children with PLI have examined the role of L2 proficiency levels on the diagnostic accuracy in differentiating between children with and without PLI. Research on L2 proficiency in learners with TD suggests that the contribution of L2 components in identifying an overall proficiency level is not constant (Adams, 1980; de Jong & van Ginkel, 1992; Iwashita, 2010; Iwashita, Brown, McNamara, & O'Hagan, 2008). Specifically, Adams (1980) reported that vocabulary, grammar, and comprehension were the best discriminators between L2 proficiency levels. Comprehension discriminated best between adjacent proficiency levels in the low-intermediate range; grammar and vocabulary discriminated best across all but two adjacent proficiency levels. Vocabulary was the only factor that discriminated at the lower proficiency levels. Further, fluency was a significant discriminator between adjacent levels in the low-intermediate range.

Recent studies have examined the relationship between overall L2 proficiency levels and a range of L2 proficiency components using language sample analysis (Iwashita, 2010; Iwashita et al., 2008). Iwashita et al. (2008) examined the relationship between L2 grammatical accuracy and complexity, vocabulary, pronunciation, and fluency, and holistic L2 proficiency scores. They found that all proficiency components contributed to the differentiation among holistic proficiency scores; however, vocabulary and fluency were particularly important at lower proficiency levels. Accuracy in tense markings, plural markings, and a percentage of grammatical T-units contributed most to differentiation among higher proficiency levels. Similarly, Iwashita (2010) reported that grammatical accuracy measured by a percentage of grammatical T-units, lexical diversity measured by *D*, and fluency measured by speech rate and a number of unfilled pauses per 60 second segments, differentiated between high and low L2 proficiency levels.

Given that various L2 components contribute differently to discrimination among proficiency levels in learners with TD, it can be hypothesized that at different L2 proficiency levels sequential bilingual children with and without PLI would demonstrate distinct L2 characteristics. Consequently, the contribution of different L2 components to discrimination between bilingual children with and without PLI may not be constant at different proficiency levels. Therefore, there is a need for a systematic study examining the contribution of L2 components to discrimination between children with and without PLI at different L2 proficiency levels.

Summary

Research investigating differences in L2 performance in bilingual children with and without PLI is primarily based on investigations of productive grammatical morphology tasks (Blom & Paradis, in press; Jacobson & Livert, 2010; Jacobson & Schwartz, 2005); thus, examination of additional L2 components, including syntactic complexity, verbal fluency, and lexical diversity, will lead to better understanding of the impact of PLI on L2 proficiency in bilingual children. Further, preliminary research indicates that the effects of PLI on L2 proficiency may not be evident at earlier stages of L2 acquisition (Paradis, 2008; Smyk et al., 2011; Verhoeven, Steenge, & van Balkom, 2011; Verhoeven, Steenge, van Weerderburg, et al., 2011); however, the majority of studies on bilingual children with and without PLI do not control for L2 exposure patterns (Crutchley et al., 1997; Verhoeven, Steenge, & van Balkom, 2011; Verhoeven, Steenge, van Weerderburg, et al., 2011), which may influence the differences in L2 proficiency levels and limit the generalizations of the effects of PLI on L2 proficiency.

The question of how L2 assessment can contribute to differentiation between children with and without PLI at different proficiency levels remains open. It is plausible that different L2 measures contribute to the discrimination between the language ability groups at different L2 proficiency levels. Investigating of how PLI impacts L2 performance in relation to length of L2 exposure and across L2 proficiency levels in bilingual children does not negate the necessity to evaluate two languages for identification of PLI, but rather adds the unique contribution to understanding the effects of PLI on L2 proficiency, the role of L2 in differentiating between the language ability groups, and the future development of bilingual language measures and norms.

Purpose

The present study aimed to (a) investigate cross-sectionally the differences in L2 proficiency components, including syntactic complexity, grammatical accuracy, lexical diversity, verbal fluency, and comprehension of grammatical structures in sequential bilingual Spanish-English speaking children with and without PLI, given different length of L2 exposure, (b) determine the best set of predictors among L2 proficiency components that differentiates between sequential bilingual Spanish-English speaking children with and without PLI at the low, intermediate, and high proficiency levels.

Research Questions and Hypotheses

This study addressed the following research questions:

(1) Are there differences between bilingual children with and without PLI in the following L2 proficiency components: syntactic complexity, grammaticality accuracy, lexical diversity, verbal fluency, and comprehension of grammatical structures, after one, two, and three years of English language exposure?

It is hypothesized that L2 proficiency scores would be significantly lower for bilingual children with PLI than for bilingual children with TD across all L2 components after at least two years of L2 exposure.

(2) What set of L2 proficiency measures best discriminates between children with and without PLI at each of following L2 proficiency levels: low, intermediate, and high?

It is hypothesized that discrimination accuracy for syntactic complexity, grammatical accuracy, lexical diversity, verbal fluency, and comprehension of grammatical structures will vary with different L2 proficiency levels, with syntactic and grammatical measures providing the best discrimination between sequential bilingual children with and without PLI at the intermediate and high proficiency levels, while verbal fluency, lexical diversity and comprehension providing the best discrimination at the low proficiency level.

Method

The study had a group experimental design in which qualified sequential bilingual children with TD and PLI were selected for the project.

Participants

Recruitment. Children were recruited from a larger study of predominantly Spanish-speaking children learning English as a L2, ages 5-7 years old conducted in the Phoenix metropolitan area. Selection measures were administered as a part of the larger study, and thus only children who completed the selection measures were invited to participate in the study. Following approval from the university's institutional review board (Appendix A), parents were sent home a separate consent form to allow their children's participation in the experimental tasks. Families of the children, who completed all the experimental tasks, received a five dollar gift card for their participation.

Sample characteristics. One hundred seventy children were recruited for the study. Sixty-three children were disqualified prior to data collection because they spoke Spanish and English from birth or were older than 8 years of age. One-hundred-seven sequential bilingual children ages 5-7 years who spoke Spanish as their primary language and English as a L2 were consented and tested. Among them, 34 children (31.8%) did not qualify because they did not meet the selection criteria.

Seventy-three children met all selection criteria. Sixty one participants (35 boys and 26 girls) were classified as children with TD and 12 participants (11 boys and 1 girl) were classified as children with PLI. The mean age for children with TD was 6 years 8 months (*SD* = 10 months), ranging from 5 years 3 months to 8 years. The mean age for children with PLI was 6 years 9 months (*SD* = 9 months), ranging from ages 5 years 5 months to 7 years 8 months.

Ninety percent of the parents reported that their children spoke a Mexican dialect of Spanish, 2.7% reported that their children spoke a Guatemalan dialect of Spanish, and 6.8% did not report that information. School lunch program status was used as an indirect measure of the income level. Ninety-six percent of the participants qualified for free lunch, 1.4% qualified for regular lunch, and 2.7% did not report that information. Twelve percent of mothers had a college degree, 46.6% had a high school diploma, 37 % completed primary school, and 3% did not report that information.

Selection criteria. All children met the following criteria: (a) attended English development classrooms and were identified as L2 learners based on teacher questionnaire; (b) had no record of significant sensory, neurological, or cognitive problems as indicated by teacher and parent questionnaires; (c) passed a bilateral hearing screening at 20 dB HL in the frequency region from 1000 through 4000 Hz (American National Standards Institute, 1969); (d) scored 75 or above on the Matrices and Recognition subtests from The Wechsler Nonverbal Scale of Ability (WNV; Wechsler & Naglieri, 2006); (e) spoke Spanish more than 50% of the time at home based on parent questionnaire; (f) spoke Spanish from birth and were exposed to English after at least two years of age based on parent questionnaire; (g) scored 4 out of 5 on a phonological probe examining the production of final consonants in single words that are critical for production of English grammatical morphemes.

Children with TD met the following criteria: (a) had no history of speech and language problems based on parent questionnaire; (b) scored 66 or higher on the Clinical Evaluation of Language Fundamentals-4 Spanish (CELF-4; Semel, Wiig, & Secord, 2006); (c) had .20 or less grammatical errors per T-unit in a language sample in Spanish (Restrepo, 1998); (d) scored 4.5 or higher in Spanish on the Spanish-English Language Proficiency scales (SELPS; Smyk, Restrepo, Gorin, & Gray, in preparation).

Children with PLI met the following criteria: (a) scored less than 66 on the CELF-4 Spanish (Semel, Wiig, & Secord, 2006); (b) had more than .20 of grammatical errors per T-unit in a language sample in Spanish (Restrepo, 1998); (c) were judged as native Spanish speakers by trained bilingual research assistants; (d) did not score more than .5 point higher in English than in Spanish on the SELPS (Smyk et al., in preparation) to exclude children with TD who are in the process of switching their language dominance from Spanish to English; (e) did not score higher than 81 on the Structured Photographic Expressive Language

Test- 3 (SPELT-3; Dawson, Stout, & Eyer, 2003) to exclude children with TD who are in the process of switching their language dominance from Spanish to English.

Any discrepancy in the language ability classification based on the standard score on the CELF-4 and number of grammatical errors per T-unit in a language sample in Spanish were solved by a clinical judgment of a bilingual certified speech-language pathologist. The agreement rate between the CELF-4 and number of errors was 88.67%.

The CELF-4 cut off score in identifying language ability was set to standard score of 66, which is below a score of 85 recommended by a test manual. Administration of the CELF-4 to 606 bilingual children ages 5-7 years old in Phoenix metropolitan area yielded a mean score of 82.22 and a standard deviation of 15.6, which is more than one standard deviation below the test mean of 100. All participants in the current study came from the sample of 606 children. Thus, the cut off score was adjusted to 1 standard deviation below the mean of 82.22 to avoid the over-identification of children with PLI due to the average lower performance on the CELF-4. If the cut off score of 85 was used, it would have resulted in classifying 54% of participants as children with PLI, which would have been much higher the 7% prevalence rate reported for monolingual children.

Selection Measures

Parent and teacher questionnaire. Parent and teacher questionnaires were used to select children for participation in the project and to establish group membership. Parent questionnaires have been used to identify the age of L2 exposure, the quantity and quality of L2 input outside school, and history of speech and language development in L1 and L2 (e.g., Restrepo, 1998; Gutiérrez-Clellen & Kreiter, 2003; Paradis, Emmerzael, & Sorenson Duncan, 2010). The parent questionnaire included questions pertaining to the children's past and current language background and degree and length of exposure to the two languages, parents' concerns about speech and language development, and a history of speech, language, hearing, neurological, attention, or emotional deficits (Appendix B). Parent report of speech and language problems in combination with more than 20% of grammatical errors per T-unit has good sensitivity of 87.5% and excellent specificity of 100% in identifying PLI in predominantly Spanish-speaking children (Restrepo, 1998).

Teacher questionnaires have been used in the assessment of L2 proficiency in bilingual children (Bedore et al., 2011; Gutiérrez-Clellen & Kreiter, 2003; Peña, Bedore, & Rappazzo, 2003; Restrepo & Gutiérrez-Clellen, 2001). The teacher questionnaire included questions pertaining to concerns about speech and language development and L2 proficiency (Appendix C). Gutiérrez-Clellen and Kreiter (2003) found that teachers' ratings of L2 proficiency on a five-point scale in bilingual children with various levels of proficiency were moderately correlated with the proportion of grammatical utterances in language samples ($r = .44$, $p < .01$). Similarly, Bedore et al. (2011) reported that teachers' ratings on a five-point Likert-type scale were significantly correlated with the language scores on English semantics and morphosyntax subtests in Spanish-English bilingual children ages 4-5;11 years (Kendall's $\boldsymbol{\tau}$ coefficient of 0.18, $p = .006$ and .23, $p <$

.001, respectively). In contrast, Smyk et al. (in preparation) reported that despite moderate correlations between L2 proficiency scores on the SELPS and teachers' ratings, the agreement between the two ratings was poor (39.47%), with approximately 50% of the children receiving lower scores on the teachers' ratings. Teachers were asked to rate L2 proficiency on a three-point scale in which a score of 1 indicated low proficiency, a score of 2 indicated intermediate proficiency, and a score of 3 indicated high proficiency.

The Wechsler Nonverbal Scale of Ability. The WNV (Wechsler & Naglieri, 2006) is a test of general cognitive ability for ages 4 to 22 years. According to the manual, the test minimizes verbal content and can be used with bilingual populations with various levels of L2 proficiency (Wechsler & Naglieri, 2006). Test instructions can be provided in Spanish, English, or nonverbally. The norming sample included Hispanic children, minimizing the effects of the ethnic background on the score interpretations. The standard measurement of error was reported as 4.60 for the Full Scale Score of the 2-subtest battery in the U.S. normative sample. The interrater reliability ranged from .95 to 1.00. Full Scale Score Fisher's z reliability coefficient ranged from .64 (ages 4;0 years to 7;11; two subtests) to .85 (ages 13;0 to 21;11years; four-subtests). The test manual reported that the criterion-related validity was assessed through predominantly moderate relationship with other measures of cognitive abilities. In addition, the authors stated that the WNV is a good measure of general ability for children with language disorders and English language learners due to small effect sizes or negligible effect sizes.

The Clinical Evaluation of Language Fundamentals-4, Spanish. The CELF-4 (Semel, Wiig, & Secord, 2006) is a standardized measure that is used for identification of language impairment in primary Spanish-speaking children ages 5 to 21. The standardization sample included bilingual Spanish-English students; the test manual reports that Spanish was the first language of all participants. The core subtests administered to all participants included Concepts and Directions, Word Structures, Sentence Repetition, and Formulating Sentences. The test-retest reliability across subtests ranged from .80 to .95. The test mean is 100 and the standard deviation is 15.

The Structured Photographic Expressive Language Test-3. The SPELT-3 (Dawson, Stout, & Eyer, 2003) is a norm-referenced measure that is used to identify monolingual English-speaking children who perform significantly below their age peers in the production of morphosyntactic structures and to identify children with language impairment. Perona, Plante, and Vance (2005) reported that the standard score of 95 yielded a sensitivity of 91 % and a specificity of 100% in identification of language impairment in children ages 4;0 to 5;11 years. All children spoke English as their L1 and 12 participants spoke Spanish as L2. Currently, there are no studies that report sensitivity and specificity of the SPELT-3 with older children. Although the test is not intended to diagnose sequential bilingual children with language impairment, the cut off score of 81 was used with participants identified with PLI to rule out possible children with TD undergoing a language dominance switch from Spanish to English and scoring below the TD range in Spanish. Bilingual children with PLI are likely to have lower levels of language development in both languages (Kohnert, 2008; Restrepo & Kruth, 2000; Willig, 1986), and thus, participants that were identified with PLI based on the CELF-4, but scored higher than 81 on the SPELT-3 were excluded from the study due to a possible misdiagnosis of PLI.

Number of grammatical errors per T-unit in Spanish. The number of grammatical errors per T-unit in Spanish was obtained through a story retelling task of one of the wordless storybooks, *Frog on His Own* (Mayer, 1973) or *A Boy, a Dog, a Frog, and a Friend* (Mayer & Mayer, 1971) in Spanish. Story scripts were adapted from the publicly available scripts on the website of the Systematic Analysis of Language Transcripts (SALT) to control the number of T-units, mean length of T-units, and the number of pictures presented to participants. Children's retells were audio recorded, then orthographically transcribed, segmented into T-units, and coded for grammatical errors. A T-unit is defined as an independent clause with all its dependent clauses (Hunt, 1970). The number of grammatical errors per T-unit was calculated by dividing the total number of grammatical errors by a total number of T-units. Criterion of more than .20 of grammatical errors per T-unit in language samples was used for identification of children with PLI (Restrepo, 1998).

Spanish-English Language Proficiency Scales. The SELPS (Smyk et al., in preparation) is a rating scale that assesses several domains of oral language production, such as sentence length, grammaticality, verbal fluency, and vocabulary. The score inferences based on the scale were validated for identifying English language proficiency levels in sequential Spanish-English bilingual children (Smyk et al., in preparation; Smyk et al., 2011). The domain scores and composite score were found to be significantly correlated (from $r = .34$ to $r = .66$) with the language sample metrics, such as MLU, the index of grammaticality, total number of different words, and the proportion of maze words to the total number of words. In addition, the composite scores were found to be moderately correlated with teachers' ratings of L2 proficiency ($\rho = .39$). A composite score between 1 and 5 with .5 intervals were assigned, with a score of 1 indicating silent or nonverbal period and a score of 5 indicating native-like proficiency. In addition, a preliminary study indicated that the composite scores in Spanish were moderately and significantly correlated with MLU and the number of different words in Spanish ($r = .33$ and $r = .49$, $p < .05$, respectively).

The SELPS was completed in conjunction with a story retelling task in English and in Spanish. A composite score in Spanish was used to ensure that children spoke Spanish with native-like proficiency in the group with TD and better or as good as English in the group with PLI. In addition, a composite score in English was used to group children based on their L2 proficiency level. The story retelling task in English and Spanish were based on one of the wordless storybooks, *Frog on His Own* (Mayer, 1973) or *A Boy, a Dog, a Frog, and a Friend* (Mayer & Mayer, 1971). Children were administered one of the stories in English and another one in Spanish. Story scripts in English were adapted from the SALT website to control the number of T-units, mean length of T-units, and the number of pictures presented to participants.

Experimental Measures

Test of Early Grammatical Impairment. The Test of Early Grammatical Impairment (TEGI; Rice & Wexler, 2001) is a criterion-referenced language test used for the identification, diagnosis, screening, and follow-up assessment of grammatical deficits in children ages 3-8 years. It assesses the finite verb morphology in English that is found to be difficult for children with language impairments and that includes the production of the following morphemes: third person singular [-s], regular and irregular past tense, copula and auxiliary [BE], and auxiliary [DO]. In addition, the test includes a grammaticality judgment probe that evaluates the child's understanding of the use of grammatical morphemes. The test takes about 45 minutes to administer. The test scores were based on the percentage of correct responses. The test-retest reliability across probes ranged from .65 to .95. The correlations between the TEGI and the Word Structure subtest from CELF-3, English ranged from .37 to .53 for children with TD and from .41 to .54 for children with language impairments.

The phonological probe from the TEGI (Rice & Wexler, 2001) is a screener of children's ability to produce English phonemes related to marking grammatical morphemes, specifically the phonemes /s/, /z/, /t/, and /d/ in the final position. The probe consists of 20 items that require less than five minutes to administer. The minimal passing score of four for each phoneme groups was required to exclude children who omit the final consonants that are critical for assessment of English tense morphemes. Children received credit for marking the final consonant spontaneously or in response to a model. Sound distortions and substitutions were not penalized because only final consonant omissions could influence children's scores in the production of English grammatical morphemes.

Test for Reception of Grammar-Second Edition. The Test for Reception of Grammar-Second Edition (TROG-2; Bishop, 2003) is a receptive language test that investigates comprehension of grammatical forms in English such as inflections, function words, and word order. The test takes 10 to 20 minutes to administer and consists of 80 items that are organized into blocks of four items each. Each block tests a single grammatical form. The TROG-2 has split-half reliability of .88. The vocabulary screener consisting of 48 items was used as a qualitative measure of children's knowledge of vocabulary items that are used in the test. Although this test was designed for a range of clinical populations, such as children with PLI, hearing loss, and learning difficulties, it allows assessing receptive grammatical skills in L2. In addition, this measure has been previously used with bilingual children with TD in assessment of receptive grammatical skills (e.g., Marinis & Chondrogianni, 2010).

Language sample analysis. Language samples were collected through a story retelling task based on the wordless picture book, *Frog Goes to Dinner* (Mayer, 1974). Language sample analysis is a valuable source of information about children's level of language skills (Botting, 2002; Dunn, Flax, Sliwinski, & Aram, 1996; Heilmann et al., 2008; Heilman, Miller, & Nockerts, 2010; Miller et al., 2006). It is frequently recommended as an alternative to standardized assessments for bilingual children (Stockman, 1996; Gutiérrez-Clellen & Simon-Cereijido, 2009; Simon-Cereijido & Gutiérrez-Clellen, 2009) due to its relative universality of sampling context (Oller & Damico, 1991; Fiestas & Peña, 2004; Muñoz, Gillam, Peña, & Gulley-Faehnle, 2003; Uccelli & Paez, 2007) and the similarity of the general structural organization of narratives across different cultures (Mandler, Scribner, Cole, & DeForest, 1980; Price, Roberts, & Jackson, 2006).

The following experimental measures in English derived from the language sample analysis were used: (a) syntactic complexity based on MLU; (b) grammatical accuracy based on the number of grammatical errors per T-unit (a total number of grammatical errors divided by a total number of T-units; Restrepo, 1998); (c) lexical diversity based on the number of different words (NDW); (d) verbal fluency based on percent of maze words (PMW), such as false starts, repetitions, and reformulations, to total number of words produced in the language sample.

Language sample analysis has been found to be a reliable form of assessment. Heilmann et al. (2008) reported the stability of language sample scores over a period of two months with the correlation coefficients ranging from.65 for the MLU to .79 for the NDW. The coding reliability was 98-100%. Crosslinguistic studies demonstrated the usefulness of language sample measures in identification of monolingual children with PLI. Specifically, the number of grammatical errors per utterance alone has sensitivity of 69.57% and specificity of 100%, while MLU has sensitivity from 58% to 69.57% and specificity from 74 to 86.96% in identifying PLI in predominantly Spanish-speaking children ages 4-7 years (Simon-Cereijido & Gutiérrez-Clellen, 2007; Restrepo, 1998). Eisenberg, McGovern Fersko, and Lundgren (2001) reported that a cutoff MLU score of 1.5 standard deviation below the mean yielded an overall accuracy of 79% with 96% specificity and 54% sensitivity for monolingual English-speaking children ages 1;6-5 years.

Procedures

Parent and teacher questionnaires were collected prior to data collection to determine whether children should be selected to participate in the study based on linguistic and demographic background. All children participated in two to three testing sessions of 45-60 minutes each as a part of the larger study. During these sessions the selection measures, including the WNV, the SELPS, the hearing screening, the SPELT-3, and CELF-4, were administered by trained bilingual graduate and undergraduate students.

The consented and qualified children participated in an additional testing session of about 45-60 minutes in a quiet place at the child's school. During this session, the experimental measures in English were administered individually in a random order. Prior to the administration, an examiner established rapport with each child and collected the child's verbal assent to participate in the testing session. The TEGI and language samples in English were audio recorded by an Olympus WS-600S digital voice recorder for subsequent scoring in the Bilingual Language and Literacy Laboratory. In order to avoid subjective bias, examiners were blind to the language ability of children (PLI or TD) and different examiners administered experimental and selection measures.

The TEGI and TROG-2 were administered according to the testing procedures specified by each test manual. Administration of the story retelling task was based on the standardized procedures (see Appendix D for the story script and the testing instructions). First, the examiner read the story script to each child while the child looked at the pictures in the book. Then the examiner asked the child to look at the pictures and to retell the story. The book was available for children to minimize memory load by providing the contextual support. The examiner asked general questions if the child struggled with the retell (e.g., *What happened here? Can you tell me more about it? Keep going! What else?*) and briefly affirmed and encouraged children, especially in the beginning of their retells. If children started to switch to their L1 or to talk about unrelated topics, the examiner re-directed children multiple times if necessary (e.g., *Can you say that in English? So what happened next in the story?*).

Procedural fidelity. Four examiners participated in the administration of the experimental tasks. There was an additional examiner who administered only one story retelling task. To ensure the proper administration of the experimental measures, all the examiners were trained to reliability of 90% or better prior to data collection. The training included an online session which presented an overview of the experimental tasks, goals, and testing instructions and procedures for each measure. After the completion of the session, the examiners were required to pass two online quizzes with a minimum score of 90% each. Additionally, the examiners practiced the administration of each task with the author who gave a detailed feedback regarding the testing procedures. During the data collection, the author was on site and supervised each examiner during the majority of the testing sessions. The procedural fidelity checklist (Appendix E) was completed by the author on at least 20% of the testing sessions for each examiner. Each of the examiners received a score of 90% or higher for every observed testing session.

Scoring and inter-rater reliability of experimental measures. Language samples in English were orthographically transcribed using SALT (Miller & Iglesias, 2010) and segmented into T-units. A T-unit was defined as an independent clause with all its dependent clauses (Hunt, 1970). Grammatical errors were coded in each T-unit.

Reliability of language transcriptions was established through the process of consensus, which was found as accurate as comparing an initial transcription to a transcription that was completed by a second independent transcriber (Heilmann et al., 2008). In this process, an initial transcription was completed by a trained transcriber, and then verified by a second transcriber who listened to the audio file and reviewed the written transcript. All the disagreements were recorded, and then solved by the consensus between two transcribers. Consensus was established on 100% of the transcriptions.

Reliability of coding was established by coding of 19 (24%) randomly selected transcriptions by a trained research assistant. The agreement for the coding of T-units was 99.48% and for coding of grammatical errors was 94.17%.

The TROG-2 was scored according to the test manual. There were 20 test blocks with four items per block and one practice block with two practice items.

If the child responded incorrectly to at least one item in the block, the block was scored as failed. The test administration was discontinued if a child failed 5 consecutive blocks of items. A number of passed blocks represented the total score. Inter-rater reliability of the TROG-2 was established by double scoring 16 (22%) randomly selected protocols by an independent scorer. The agreement between the scorers was 100%.

The TEGI was scored according to the test manual. Participants received a separate score for the third person singular [-s], past tense, [BE] and [DO] production probes. The scores were calculated as percent of correct items. Then the elicited grammar composite score was computed by summing the probe scores and dividing them by the number of probes attempted. A probe was considered not attempted if a child did not use the target structure correctly or incorrectly. For example, the target structure in the third person singular [-s] probe is a subject in the third person singular form + verb with the [-s] morpheme (e.g., a teacher teaches). The response was considered correct if the target was used and incorrect if [-s] morpheme was omitted. If the child used any other grammatical structures, such as a subject in the third person singular form + auxiliary + verb with the [-ing] morpheme (e.g., a teacher is teaching), the test item was considered unscorable. If a child did not have any scorable responses in any given probe, the probe score was not assigned and the probe was considered as not attempted. Three separate scores were calculated for the grammaticality judgment task based on the item type, such as dropped marker (e.g., *he furry), agreement (e.g., *he are mad), and dropped –ing form (*he is smile). The grammaticality judgment score was computed by summing the scores for three item types.

Inter-rater reliability of the TEGI was established by double scoring 16 (22%) randomly selected protocols by an independent scorer. The agreement for the third person singular [-s] probe was 97%, for the past tense probe was 96.6%, for the [BE] probe was 91.7%, for the [DO] probe was 92.2%, and for the grammaticality judgment probe was 100% for every item type.

Grouping participants on level of L2 proficiency. Children with and without PLI were divided into low, intermediate, and high L2 proficiency groups based on the SELPS composite scores in English and teachers' ratings of L2 proficiency. The range of SELPS scores from 1 to 2.5 indicated low L2 proficiency level, scores from 3 to 3.5 indicated intermediate proficiency level, and scores equal or above 4 indicated high proficiency level. The disagreement between the SELPS scores and teachers' ratings was resolved by an independent trained rater who identified a proficiency level on a 3-point scale. The agreement of two out of three proficiency scores was required to identify a proficiency level.

The SELPS scoring in English was done by two independent raters. The rater 1 scored 48 files (65.8%), and the rater 2 scored 25 files (34.2%). The rater 1 double scored 5 (20%) of the randomly selected files that were previously scored by the rater 2. The point-to-point agreement was 90%. The rater 2 double scored 11 (23%) of the randomly selected files that were previously scored by the rater 1. The point-to-point agreement was 82%.

Analysis

Research question 1. To examine the differences between the bilingual children with and without PLI on the L2 proficiency components, including the TROG-2 score, the grammaticality judgment score, and scores based on the language sample analysis, six one-way analyses of variance (ANOVAs) were conducted.

Multiple ANOVAs were selected over a multivariate analysis of variance (MANOVA) due to the focus on testing univariate group contrasts for each experimental measure rather than testing differences on a linear combination of the dependent variables (Huberty & Morris, 1989; Jaccard & Guillamo-Ramos, 2002). Although it is hypothesized that there are group differences in performance on each of the measure, it remains an empirical question whether PLI has similar impact on the L2 proficiency components. In addition, the use of MANOVA as a first step before conducting multiple ANOVAs to control for familywise Type I error has been criticized because only when the null hypothesis in MANOVA is true, the alpha value for each ANOVA would be less than or equal to the alpha used for the MANOVA (Huberty & Morris, 1989).

Language sample measures, such as MLU, the number of errors per T-unit, NDW, PMW, and the TROG-2 and the grammaticality judgment scores were the dependent variables, while the language ability (PLI and TD) was the factor. To control for Type I error across multiple ANOVAs, the Holm's sequential procedure (Holm, 1979) was used. The procedure orders the *p* values obtained in the multiple analyses from smallest to largest and then compares the smallest *p*-

value to α/k, in which k is the number of the null hypotheses. If that p-value is less than α/k, then that null hypothesis is rejected and the remaining $k - 1$ hypotheses are tested by comparing the smallest remaining p value to $\alpha / (k - 1)$. The process is continued until the hypothesis with the smallest p-value cannot be rejected. After that none of the remaining null hypotheses can be rejected.

Each ANOVA had the unequal *n*-factorial design, in which *n* is the sample size for each group. According to Tabachnick and Fidell (2007), it is recommended to use Type III sum of squares with unequal sample sizes when the interaction between factors is nonzero in the population. Type III sum of squares tests differences in unweighted marginal means and represents an easier interpretable solution than Type I or Type II sum of squares (Tabachnick & Fidell, 2007).

Power analysis. A power analysis was conducted using G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007) to determine a priory a sufficient sample size in identification of the meaningful group differences between bilingual children with and without PLI at an alpha level of .05 with a power value of .80. The effect sizes based on partial eta squared reported in the literature vary from .06 for MLU in words to .38 for grammaticality score in language samples of 7-9 year old children learning Dutch as a second language (Verhoeven, Steenge, & van Balkom, 2011). Currently, there are no studies that reported the effect sizes on language sample measures and grammatical production and comprehension measures for 5-7 year old sequential bilingual children learning English as a L2. Thus, a total sample size was estimated based on Ferguson's (2009) recommendations of interpreting eta squared in social science data as small (.04), medium (.25) and large (.64). There were six groups based on the language ability (TD and PLI) and the length of exposure to L2 (one, two, and three years). A total sample size of 11 children is required to determine group differences with a large effect size, a sample size of 33 children is required with a medium effect size, and a sample size of 235 children is required with a small effect size. Thus, a sample size of 73 children would be sufficient to identify group differences for large and medium effect sizes, but would be under power for small effect size. However, a priori power analysis assumes that there will be an equal number of participants in each group, and thus, the unequal sample size results in reduced power.

Research question 2. To examine the extent to which the language sample measures and scores on TEGI and TROG-2 differentiate between sequential bilingual Spanish-English speaking children with and without PLI at low, intermediate, and high L2 proficiency levels of L2, the stepwise discriminant function analyses were planned, one for each proficiency level. Language sample measures, including MLU, the number of grammatical errors per T-unit, NDW, PMW, and the scores on TROG-2 and TEGI were the predictors, while the group membership (with two levels, PLI or TD) was a dependent variable. In addition, the discriminant analysis for the entire sample of bilingual children with and without PLI was conducted.

The discriminant functions were assessed based on their level of significance and effect size. Wilks's lambda examined the level of significance, while eigenvalues and canonical correlation coefficients examined the effect size.

To estimate classification accuracy, the percentage of participants correctly classified was estimated using the leave-one-out procedure. The correlation between each measure and the discriminant function (structure correlation coefficients) and the unique contribution of each measure with respect to the discriminant function in differentiating groups (standardized discriminant function coefficients) were provided.

Results

Descriptive Statistics for the Selection Measures

Descriptive statistics (the mean, the median, standard deviation, skew and kurtosis) for the selection measures are reported in Table 1 for each language ability group. Pearson correlations among the selection measures, chronological age, age and the length of L2 exposure for each language ability group are reported in Table 2.

Based on the parent questionnaire, two children with PLI (16.7%) received speech and language therapy services, nine children (75%) did not receive services, and one child (8.3%) had missing information. Thirty-six children with TD (59%) did not receive speech and language therapy services and 25 children (41%) had missing information because parents did not respond to that question in the parent questionnaire. Parents' concerns about the participants' language development for each language ability group are summarized in Table 3. Overall, parents of children with PLI appeared to be more concerned about their child's language development than parents of children with TD.

Based on the teacher questionnaire, two children with PLI (16.7%) received speech and language therapy services, which confirmed the information reported by parents, and 10 children (83.3%) did not receive services. Fifty-five children with TD (90.2%) did not receive speech and language therapy services and six children (9.8%) had missing information. Teachers' concerns about the participants' language development for each language ability group are summarized in Table 3. Teachers appeared to be almost equally concerned about language development in bilingual children with and without PLI.

Research Question 1 - the differences in L2 proficiency components in children with and without PLI who were grouped on a length of L2 exposure

Grouping children on length of L2 exposure. Based on the parent questionnaire, bilingual children with and without PLI were grouped on a length of L2 exposure within one-year intervals (the length of exposure was calculated based on the chronological age at the time of testing and the age of L2 exposure). Mean chronological age, standard deviation, and the number of children for each ability and L2 exposure group within each one-year interval are presented in Table 4. Five children with TD who had up to one year of L2 exposure were excluded from the analysis because there were no children with PLI with the same length of exposure for comparison. Given that only two children with PLI had between one and two years of exposure, they were combined into one group for the analysis with children with PLI who had more than two and up to three years of exposure. Overall, in the group with PLI, there were eight children with more than one and up to three years of L2 exposure (M_{age} = 6;8 years, *SD* = 9 months) and three children with more than three years of exposure (M_{age} = 7;3 years, *SD* = 2 months). In the group with TD, there were 34 children with more than one and up to three years of L2 exposure (M_{age} = 6;6 years, *SD* = 10 months) and 21 children with more than three years of exposure (M_{age} = 7;2 years, *SD* = 8 months).

To investigate age differences between children with PLI and TD at each L2 length of exposure group, two independent sample t-tests allowing for unequal variances between groups were conducted. There were no significant age differences between children with and without PLI who had between one and three years of exposure, $t(12.69) = .62$, $p = .55$ and children who had more than three years of exposure, $t(11.45) = .55$, $p = .60$.

Descriptive statistics and correlations among the experimental measures. Descriptive statistics for MLU, the number of grammatical errors per T-unit, the NDW, PMW, grammaticality judgment, the TEGI, and the TROG-2 scores for each language ability group as a function of the L2 exposure are presented in Table 5. Correlations among the experimental measures, the length of L2 exposure, and L2 proficiency level for each language ability group are presented in Table 6. The descriptive statistics and correlations among the measures for children with PLI should be interpreted with caution given the small sample size.

In the group with PLI, the number of grammatical errors per T-unit was negatively correlated with the NDW ($r = -.59$, $p < .05$) indicating that an increase in lexical diversity was associated with a decrease in the number of grammatical errors per T-unit. In addition, the number of grammatical errors per T-unit was negatively correlated with the TEGI score ($r = -.81$, $p < .01$) indicating that an increase in the grammar composite score was associated with a decrease in the number of grammatical errors in the language sample. The length of L2 exposure was not associated with the children's performance on any experimental measures.

In the group with TD, all the experimental measures were correlated with each other, with the exception of MLU and PMW, the number of grammatical errors per T-unit and PMW, and grammaticality judgment and PMW. The magnitude of correlations ranged from weak for the TEGI and PMW ($r = -.26$, $p < .05$) and the NDW and PMW ($r = -.27$, $p < .05$) to strong for grammaticality judgment and the TEGI ($r = .77$, $p < .01$), the number of grammatical errors per T-unit and the TEGI ($r = -.74$, $p < .01$), and MLU and the NDW ($r = .74$, $p < .01$). In addition, the length of L2 exposure was significantly correlated with the children's performance on all experimental measures, ranging from $r = -.27$ for the PMW to $r = .53$ for the NDW. This finding suggests that the differences in the length of exposure among children with TD were related with their performance in L2, with a longer period of L2 exposure associated with better performance.

Given strong significant correlations between the number of grammatical errors per T-unit and the TEGI scores for the children with and without PLI and a small sample size of children with PLI who completed the TEGI, the TEGI composite scores were not analyzed. Both measures assess the same construct of grammatical accuracy, which is represented with the more complete data in the number of grammatical errors per T-unit.

Group differences in L2 proficiency components. Given that there were only three children with PLI who had more than three years of L2 exposure and only two of them completed all experimental measures (Table 4), six one-way ANOVAs were conducted only with children who had between one and three years of exposure. Although these analyses do not address the proposed hypothesis of the differences on L2 proficiency measures between bilingual children with and without PLI after at least two years of L2 exposure, they do examine whether the differences between children with and without PLI are apparent between one and three years of L2 exposure.

After controlling for Type I error, results indicated non-significant group differences between children with and without PLI across all L2 proficiency components including MLU, $F(1,39) = 1.18$, $p = .28$, partial $\eta^2 = .03$; NDW, $F(1,39) =.05$, $p = .83$, partial $\eta^2 = .001$; number of grammatical errors per T-unit, $F(1,39) = .33$, $p = .57$, partial $\eta^2 = .007$; percent of maze words to total number of words, $F(1,39) = .22$, $p = .64$, partial $\eta^2 = .006$; TROG-2 scores, $F(1.37) = 1.48$, $p = .23$, partial $\eta^2 = .04$, and grammaticality judgment scores, $F(1,37) = 4.06$, $p = .05$, partial $\eta^2 = .10$. This last effect size indicates that the grammatical judgment measure maybe sensitive to differences in L2 performance between children with and without PLI given a larger sample size.

Research Question 2 – the discrimination between children with and without PLI at each L2 proficiency level

Grouping children by L2 proficiency level. Descriptive statistics and correlations between teachers' ratings of language ability and the SELPS English scores are presented in Table 1 and Table 2 for each language ability group. Based on a skewness threshold of ± 1.00 (George & Mallery, 2003), the distribution of SELPS English scores for children with TD was considered negatively skewed, indicating that a large number of children received high scores. Figure 1 illustrates the distribution of the SELPS English scores for each language ability group.

In the group with PLI, the SELPS English scores were positively correlated with the CELF Spanish scores ($r = .62$, $p < .05$) and the chronological age ($r = .59$, $p < .05$). Correlation with the CELF scores may be related to the severity of PLI: children who demonstrated low Spanish language skills were more likely to have low English language skills. These results should be interpreted with caution due to the small sample size of children with PLI. In the group with TD, the SELPS scores were positively correlated with the teachers' ratings of proficiency ($r = .42$, $p < .01$), chronological age ($r = .67$, $p < .01$), and the length of L2 exposure ($r = .56$, $p < .01$).

Children were classified into low, intermediate, or high L2 proficiency levels according to the SELPS English scores and teachers' ratings. The agreement between the SELPS scores and teachers' ratings was 50% for children with PLI and 37.7% for children with TD. Discrepancies between proficiency levels were solved by a third rater assigning a separate proficiency score. After resolving the discrepancies between teachers' ratings and SELPS scores, in the group with PLI, there were two children (16.7%) at low proficiency level, nine children (75%) at intermediate proficiency level, and one child at high proficiency level (8.3%). In the group with TD, there were eight children (13.1%) at low proficiency level, 31 children (50.8%) at intermediate proficiency level, and 22 children (36.1%) at high proficiency level.

Descriptive statistics for MLU, the number of grammatical errors per T-unit, the NDW, PMW, grammaticality judgment, the TEGI, and the TROG-2 scores for each language ability group as a function of L2 proficiency are presented in Table 7. To investigate whether L2 proficiency level is related to children's performance on the experimental measures in each language ability group, Pearson correlations were examined (Table 6). In the group with PLI, L2 proficiency was correlated with MLU (r = .66, p < .05), number of grammatical errors (r = -.82, p < .01), and the TEGI scores (r = .83, p < .05). In the group with TD, L2 proficiency was correlated with all experimental measures with the exception of PMW. The correlations ranged from r = .47 for the TROG-2 to r = .64 for the TEGI. In addition, L2 proficiency was moderately correlated with the length of L2 exposure in children with TD (r = .41, p < .01), but not in children with PLI (r = .00, p = 1.00). This finding supports evaluation of the discriminant accuracy of L2 measures for each L2 proficiency level.

Discriminant accuracy at each L2 proficiency level. To examine the extent to which the experimental measures differentiate between bilingual children with and without PLI at different L2 proficiency levels, stepwise discriminant function analyses, one for each L2 proficiency levels, were proposed. It was hypothesized that syntactic and grammatical measures would provide the best discrimination between children with and without PLI at the intermediate and high proficiency levels, while verbal fluency, lexical diversity and comprehension would provide the best discrimination at the low proficiency level.

Given that there was one child with PLI in the high proficiency and two children with PLI in the low proficiency group, the discriminant analyses for these proficiency levels were not conducted. The discriminant analysis for the intermediate proficiency group indicated that none of the predictor variables met the statistical test for inclusion in the analysis (the probability of F to include the predictor in the analysis was .05 and to remove a predictor from the analysis was .10). The results were similar when children with high and intermediate proficiency levels were combined in one group and when children with low and intermediate proficiency were combined in one group.

To examine whether the experimental measures have the potential to differentiate between children with and without PLI at different L2 proficiency levels, a series of independent sample t-tests allowing for unequal variances between groups were conducted. In the low proficiency group, children with and without PLI differed only in the TEGI scores, $t(7.81) = -6.18$, $p < .001$, but not in performance on any other measures. In the intermediate proficiency group, children with and without PLI differed only in PMW, $t(19.68) = -2.18$, $p = .042$, but not in performance on any other measures. When children from the intermediate and high proficiency groups were combined in one group, the results indicated differences between children with and without PLI in the TEGI, $t(9.53) = -2.49$, $p = .03$ and the TROG-2, $t(13.11) = -3.47$, $p = .004$, but not in performance on any other measures.

Given that the discriminant analysis was not conducted for each L2 proficiency level, the discrimination accuracy of L2 measures was evaluated for the entire sample of bilingual children with and without PLI. The discriminant function was statistically significant, $\Lambda = .93$, $\chi^2(1, N = 68) = 4.86$, $p = .027$, indicating that there are significant differences between children with and without PLI among the predictors variables. The discriminant function had an eigenvalue of .077 and a canonical correlation of .27 indicating that 7% of the variability of the scores was accounted for by the differences between children with and without PLI. The within-group correlations between the predictors and the discriminant function and the standardized discriminant function coefficients are presented in Table 8. Based on the standardized discriminant function coefficients, the MLU was the only measure that contributed to the discrimination between children with and without PLI after controlling for other predictors. Children with TD had the mean MLU of 6.68 and the standard deviation of 1.32, while children with PLI had the mean MLU of 5.98 and the standard deviation of 1.3.

The discriminant function classified correctly classified 83.3 % of the sample as with TD or PLI. The leave-one-out procedure yielded cross-validation classification accuracy of 81.9%. Sensitivity, specificity, and positive and negative likelihood ratios were not calculated due to the small sample size of children with PLI.

Differences in performance on the experimental measures in children with TD across L2 proficiency levels. Given the small sample size of children with PLI, additional one-way ANOVAs evaluating differences in performance on the experimental measures in children with TD across different L2 proficiency levels were conducted, with L2 proficiency as a factor with three levels (a low, intermediate, and high level) and each of the experimental measure as dependent variables. To control for Type I error across multiple ANOVAs, the Holm's sequential procedure (Holm, 1979) was used.

There was a significant effect of L2 proficiency on the TEGI scores, $F(2,58) = 20.86$, $p < .001$, partial $\eta^2 = .42$. According to the Holm's procedure, this comparison was significant because p was less than the required alpha level of .007 for the first comparison ($p < \alpha/k$, where k is the number of comparisons). Follow up analyses revealed significant group differences between children with low and intermediate proficiency levels (mean difference of -20.93), between children with intermediate and high proficiency levels (mean difference of -29), and between children with low and high proficiency levels (mean difference of -49.93). Figure 2 presents the mean TEGI score for each L2 proficiency level.

There was a significant effect of L2 proficiency on the grammaticality judgment scores, $F(2,58) = 18.45$, $p < .001$, partial $\eta^2 = .39$. This comparison was significant at an alpha level of .008 for the second comparison ($p < \alpha/(k - 1)$). Follow up analyses revealed significant group differences between children with low proficiency and intermediate proficiency levels (mean difference of -.60), between children with intermediate and high proficiency levels (mean difference of -.57), and between children with low and high proficiency levels (mean difference of -1.17). Figure 3 presents the mean grammaticality judgment score for each L2 proficiency level.

There was a significant effect of L2 proficiency on NDW, $F(2,57) = 13.76$, $p < .001$, partial $\eta^2 = .33$. This comparison was significant at an alpha level of .01 for the third comparison ($p < \alpha/(k - 1)$). Follow up analyses revealed significant differences in the NDW between children with low and intermediate proficiency levels (mean difference of -23.70), between children with intermediate and high proficiency levels (mean difference of -18.41), and between children with low and high proficiency levels (mean difference of -42.11). Figure 2 presents the mean NDW for each L2 proficiency level.

There was a significant effect of L2 proficiency on number of grammatical errors per T-unit, $F(2,57) = 13.62$, $p < .001$, partial $\eta^2 = .32$. This comparison was significant at an alpha level of .0125 for the forth comparison ($p < \alpha/(k - 1)$). Follow up analyses revealed significant group differences between children with low proficiency and intermediate proficiency levels (mean difference of .38) and between children with low and high proficiency levels (mean difference of .56). Figure 4 presents the mean number of grammatical errors for each L2 proficiency level.

There was a significant effect of L2 proficiency on MLU, $F(2,57) = 8.63$, $p = .001$, partial $\eta^2 = .23$. This comparison was significant at an alpha level of

.016 for the firth comparison ($p < \alpha/(k - 1)$). Follow up analyses revealed that children with low proficiency level had significantly shorter MLU than children with intermediate and high proficiency levels (mean difference of -1.34 and -2.03, respectively). Figure 3 presents the mean MLU for each L2 proficiency level.

There was a significant effect of L2 proficiency on the TROG-2 scores, $F(2,58) = 8.33$, $p = .001$, partial $\eta^2 = .22$. This comparison was significant at an alpha level .025 for the sixth comparison ($p < \alpha/(k - 1)$). The follow up analyses revealed significant group differences between children with low and high proficiency levels (mean difference of -4.99) and between children with intermediate and high proficiency levels (mean difference of -2.59). Figure 3 presents the mean TROG-2 score for each L2 proficiency level. Results also indicated a non-significant effect of L2 proficiency level for the PMW, $F(2,57) = .41$, $p = .66$, partial $\eta^2 = .01$.

Discussion

The purpose of the study was to examine whether there were differences in L2 proficiency components in bilingual children with and without PLI after at least two years of L2 exposure and whether the L2 proficiency measures discriminated between the language ability groups at low, intermediate, and high proficiency levels. The present study represents an initial step in improving the accurate diagnosis of PLI in bilingual children by comparing L2 performance in sequential bilingual children with and without PLI on a range of language tasks, which can potentially disentangle the effects of PLI from the effects of typical L2 acquisition processes. The study is the first to document how PLI affects L2 proficiency with an increasing length of L2 exposure and how L2 assessment contributes to differentiation between children with and without PLI at different L2 proficiency levels.

Differences between Bilingual Children with and without PLI on L2 Proficiency Components

Cross-linguistic research indicates differences in L2 performance on different language tasks between bilingual children with and without PLI (in English: Gutiérrez-Clellen et al., 2008; Jacobson & Livert, 2010; Jacobson & Schwartz, 2005; Restrepo & Kruth, 2000; in Arabic: Salameh et al., 2004; in Dutch: Orgassa & Weerman, 2008; Verhoeven, Steenge, & van Balkom, 2011; Verhoeven, Steenge, van Weerderburg, et al., 2011); however, these studies did not examined whether these differences are affected by an increased length of L2 exposure. Furthermore, most studies investigated only productive grammatical morphology (Blom & Paradis, in press; Gutiérrez-Clellen et al., 2008; Jacobson & Livert, 2010; Jacobson & Schwartz, 2005), and therefore, the effects of PLI on other L2 proficiency components, such as syntactic complexity, receptive grammatical morphology, verbal fluency, and lexical diversity, remained unclear. Based on preliminary research (Paradis, 2008; Smyk et al., 2011), it was hypothesized that L2 proficiency scores would be significantly lower for bilingual children with PLI than for bilingual children with TD across syntactic complexity, grammaticality accuracy, lexical diversity, verbal fluency, and comprehension of grammatical structures after at least two years of L2 exposure.

Language ability differences across the length of exposure groups. Given the small sample size of children with PLI across the length of exposure groups, the analysis was conducted only for children who had between one and three years of L2 exposure. Therefore, the proposed hypothesis was not addressed as planned. Nonsignificant language ability group differences on all L2 proficiency measures suggest that the differences in L2 performance on language sample measures and productive and receptive grammatical tasks are not apparent when bilingual children have between one and three years of L2 exposure. However, these results should be interpreted with caution because of the small sample size of children with PLI and the unequal factorial design that reduced the statistical power of the analyses. Thus, it is possible that given a larger sample size and a balanced design, the significant differences between the language ability groups as a function of a length of exposure could be identified.

When the effect sizes were evaluated to assess the clinical relevance of the group differences, the grammaticality judgment task demonstrated the largest effect size (the partial $\eta^2 = .10$), suggesting that bilingual children with and without PLI demonstrate clinically relevant differences in performance on this task. Up to date, there are no studies that compared the performance of bilingual children with and without PLI on a grammaticality judgment task. Although bilingual children with TD who had up to 18 months of L2 exposure were found to be as accurate in detecting ungrammaticality as MLU-matched monolingual children with PLI (Paradis et al., 2008), research on monolingual children indicates that there are language ability group differences in grammaticality judgment (e.g., Rice, Hoffman, & Wexler, 2009; Rice, Wexler, & Redmond, 1999; Saxton, Dockrell, Bevan, & van Herwegen, 2007). Therefore, this task may potentially contribute to diagnosis of PLI in bilingual populations, when the performance of bilingual children with and without PLI is compared.

The effect sizes for all other L2 measures were negligible, indicating that the differences between the language ability groups with more than one and up to three years of L2 exposure may not be clinically relevant. This conclusion contrasts with Paradis (2008) and Smyk et al. (2011) who reported language ability group differences in performance on productive grammatical morphology and a rating scale of L2 proficiency after two years of exposure. However, in both studies children were grouped on the length of L2 exposure within one year intervals. Therefore, it is possible that in addition to the reduced power, the effect of L2 exposure was not evident because of the grouping children with a wide range of the length of L2 exposure, which resulted in greater variance in performance on L2 measures in each language ability group.

In line with previous research on bilingual children with TD, results also indicated that the length of L2 exposure was related to children's performance in L2, with a longer period of exposure associated with better performance (e.g., Armon-Lotem, Walters, & Gagarina, 2011; Chondrogianni & Marinis, 2011; Golberg, Paradis, & Crago, 2008; Jia & Fuse, 2007; Marinis & Chondrogianni, 2011). At the same time, this finding contrasts Paradis (2005) and Paradis et al. (2008) who found no associations between the length of L2 exposure and L2 performance in children with TD. However, children in Paradis and collegues' studies were within their first and a half year of L2 exposure, and, thus such limited range of the length of exposure could have affected the results.

The strongest associations between the length of exposure and L2 performance were found for lexical diversity and grammatical accuracy measured by the grammaticality judgment task and the TEGI. This finding could have been influenced by the strong associations between lexical diversity and grammatical accuracy themselves, which were previously reported in bilingual children with TD (e.g., Castilla, Restrepo, & Perez-Leroux, 2009; Kohnert, Kan, & Conboy, 2010; Simon-Cereijido & Gutiérrez-Clellen, 2009). In addition, it is possible that the length of L2 exposure has the stronger relationships with lexical diversity and grammatical accuracy in comparison with other L2 proficiency components.

In contrast to children with TD, the length of L2 exposure was not associated with L2 performance in children with PLI, suggesting that the relationship between the length of exposure and L2 performance is marginal in children with PLI. However, this finding should be interpreted with caution given a small sample size of children with PLI, and thus, future studies with bigger sample sizes are necessary to examine how length of L2 exposure affects L2 performance in bilingual children with PLI.

Discrimination between Children with and without PLI at Different L2 Proficiency Levels

The second goal of the study was to examine what set of L2 proficiency measures discriminated best between children with and without PLI separately at low, intermediate, and high proficiency level. Because the study is the first to examine the role of L2 proficiency level in differentiating between children with and without PLI, the results could potentially provide a comprehensive picture of the relationship between PLI and L2 proficiency and the contribution of L2 assessment to diagnosis of PLI in bilingual children.

Research on L2 learners with TD indicates that the contribution of L2 components in identifying an overall L2 proficiency level is not constant (e.g., Adams, 1980; de Jong & van Ginkel, 1992; Iwashita, 2010; Iwashita et al., 2008). Therefore, it was hypothesized that at each L2 proficiency level, bilingual children with and without PLI would demonstrate distinct L2 differences, and thus, different L2 measures may contribute to discrimination between the language ability groups. Specifically, it was hypothesized that syntactic and grammatical measures provide the best discrimination between bilingual children with and without PLI at the intermediate and high proficiency levels, while verbal fluency, lexical diversity and comprehension of grammatical structures provide the best discrimination at the low proficiency level.

First, the relationships between the length of L2 exposure and L2 proficiency levels for each language ability group were examined to confirm whether it was necessary to control for L2 proficiency level. Results indicated moderate associations between the length of L2 exposure and L2 proficiency level only for children with TD. This conclusion is consistent with previous studies on bilingual children with TD indicating that length of L2 exposure is not the only factor that affects the L2 proficiency level (Armon-Lotem et al., 2011; Bohman, et al., 2010; Chondrogianni & Marinis, 2011; Golberg et al., 2008; Gutiérrez-

Clellen et al., 2012; Oller & Eilers, 2002; Paradis, 2011). However, children with PLI did not demonstrate the same pattern of association between the length of exposure and L2 proficiency, suggesting that they do not develop their L2 proficiency skills in relation to length of L2 exposure in the same manner as children with TD. Further, no associations between the length of L2 exposure and L2 proficiency levels and between the length of exposure and L2 performance on the experimental measures in children with PLI support the conclusion that effects of PLI on L2 performance cannot be explained by the patterns of L2 exposure (Kohnert, 2008; Restrepo & Kruth, 2000; Willig, 1986). Overall, these results suggest that bilingual children with and without PLI who are grouped only on a length of L2 exposure may still demonstrate different L2 proficiency levels, and thus, the length of L2 exposure is not a sufficient control for language proficiency.

Discriminant analyses for the low and high proficiency groups were not conducted given the small sample size of children with PLI. In the intermediate proficiency group, none of the predictor variables met statistical requirements necessary for inclusion in the analysis. Similar result was found when children with the low and intermediate proficiency levels and when children with intermediate and high proficiency levels were combined in one group.

To examine whether the experimental measures have the potential to differentiate between children with and without PLI at different proficiency levels, t-tests were conducted. At the low proficiency level, grammatical accuracy assessed by the TEGI may potentially contribute to differentiation between the language ability groups, which contrasts the proposed hypothesis. Although previous studies on bilingual children with and without PLI did not control for L2 proficiency level, a large body of research indicates that L2 performance on expressive grammatical tasks contributes to discrimination between the language ability groups (Blom & Pardise, in press; Gutiérrez-Clellen & Simon-Cereijido, 2009; Jacobson & Schwartz, 2005; Jacobson & Livert, 2010). Specifically, Jacobson and Schwartz (2005) reported that performance on a past tense task in English as a L2 has sensitivity of 92% and specificity of 87% in identification of PLI for regular verbs, sensitivity of 92% and specificity of 73% for irregular verbs, and sensitivity of 100% and specificity of 73% for novel verbs that have phonological similarity to regular and irregular past tense verbs.

At the intermediate proficiency level, verbal fluency assessed by the PMW may contribute to differentiation between the language ability groups. This finding was also unexpected especially because children with PLI demonstrated on average lower PMW than children with TD. Although research on monolingual children suggests that children with PLI tend to produce more maze words than children with TD (Leadholm & Miller; 1995), it is possible that children with TD attempted to produce more complex utterances, which put greater demands on verbal fluency and resulted in larger number of maze words (Evans, 1985; Leadhold & Miller, 1995). Research on monolingual children also suggests that language ability group differences may be apparent in terms of the types of disfluencies, such as revisions, repetitions, and hesitations, rather than in the percent of maze words (Navarro-Ruiz & Rallo-Fabra, 2001; Thordardottir & Weismer, 2002).

At the intermediate-high proficiency level, comprehension of grammatical structures assessed by the TROG-2 and grammatical accuracy assessed by the TEGI may have the potential to differentiate between the language ability groups. These findings are consistent with those of Gutiérrez-Clellen and Simon-Cereijido (2009) and Verhoeven, Steenge, van Weerdenburg, et al. (2011). Specifically, Gutiérrez-Clellen and Simon-Cereijido (2009) reported that a productive grammar composite score yielded sensitivity of 91% in identification of bilingual children with PLI. Although the authors did not control for L2 proficiency level, they reported that bilingual children were more dominant in English than in Spanish based on parent and teacher interviews. Verhoeven, Steenge, van Weerdenburg, et al. (2011) reported that 7-8 year old bilingual children with PLI learning Dutch as L2 performed significantly worse on a measure of comprehension of grammatical structures than bilingual children with TD. These findings suggest that receptive and expressive grammatical tasks contribute to identification of PLI, especially when children reach higher L2 proficiency levels.

The language ability group differences at each L2 proficiency level could have been influenced by variability in performance, which was evident from the substantial standard deviations and ranges of scores for L2 measures. For example, children with PLI at the intermediate proficiency level demonstrated a range of MLU from 3.97 to 7.27, while children with TD at the same L2 proficiency level demonstrated a range from 4.35 to 8.66. This finding is consistent with a large body of research indicating heterogeneous nature of L2 performance in bilingual children with and without PLI (e.g., Goldberg, et al., 2008; Jacobson & Schwartz, 2005; Kan & Kohnert, 2005). Overall, the results of the study should be replicated with a bigger sample size that can potentially reduce substantial variance in L2 performance, especially after controlling for L2 proficiency level.

Discrimination between the language ability groups without controlling for L2 proficiency level. Based on the findings from the discriminant analysis for the entire sample of bilingual children with and without PLI, MLU was the only predictor of children's diagnosis as TD or PLI. However, the effect size was quite small, with only 7% of the variability of the scores accounting for the differences between children with and without PLI. Further, MLU classified correctly 83.3 % of the sample as with TD or PLI, which was lower than the 90% classification accuracy suggested for measures of PLI (Plante & Vance, 1994). Previous studies on monolingual children with and without PLI suggest that MLU alone does not demonstrate good classification accuracy. Specifically, Restrepo (1998) and Simon-Cereijido and Gutiérrez-Clellen (2007) reported that MLU had sensitivity between 58% and 69.57% and specificity between 74% and 86.96% in identifying PLI in predominantly Spanish-speaking children ages 4-7 years. In contract with Spanish, Eisenberg et al. (2001) reported higher specificity of 96%, but lower sensitivity of 54% for MLU in monolingual English-speaking children ages 1;6-5 years. It is also possible that MLU discriminates better between the language ability groups in combination with other measures. Restrepo (1998) reported that when MLU was combined with parent report of speech and language difficulties, number of grammatical errors, and family history of speech and language difficulties, the sensitivity and specificity increased to 91.3% and 100%, respectively. Overall, the results of this study should be replicated with a bigger sample size, given a small sample size of children with PLI.

Differences in performance on the experimental measures across L2 proficiency levels in children with TD. Additional analyses were conducted to evaluate the differences in performance on L2 measures in children with TD across L2 proficiency levels. Results indicated an effect of language proficiency on all L2 measures, with the exception of PMW. This finding indirectly supported the identification of L2 proficiency level using a combination of the SELPS scores in English and teacher questionnaire and resolving any discrepancies between the two measures by a third rater. Further, the effect sizes for the L2 proficiency group differences were large, ranging from .22 for the TROG-2 scores to .42 for the TEGI scores. Therefore, children with TD across L2 proficiency levels demonstrate clinically relevant differences in L2 performance, which is consistent with previous research indicating that MLU, NDW, and performance on the expressive and receptive grammatical tasks are strongly associated with L2 proficiency level (e.g., Ellis & Barkhuizen, 2005; Iwashita, 2010; Iwashita et al., 2008; Norris & Ortega, 2003; 2009; Skehan, 2009; Larsen-Freeman, 2006).

Results indicated that grammatical accuracy assessed by the TEGI and grammaticality judgment task differed across all L2 proficiency levels, which is consistent with Adams (1980) and Iwashita et al. (2008) who reported the contribution of grammatical measures to identification of L2 proficiency level. In addition, performance on other grammatical measures also differed across proficiency groups. Specifically, children with the low and intermediate proficiency levels and between the low and high proficiency levels demonstrated differences in number of grammatical errors, suggesting that grammatical accuracy based on the language sample analysis may be more sensitive to differences among children at the lower L2 proficiency levels. These findings contrast with those of Iwashita et al., who suggested that grammatical accuracy contributes most to differentiation among higher proficiency levels. However, in the study by Iwashita and colleagues participants were older L2 English language learners who were given a pilot oral test tasks developed by Educational Testing Service, and thus, it is possible differences in L2 proficiency assessment and the participants' demographic characteristics resulted in the ambiguous findings. Further, performance on the TROG-2 differed between the low and high proficiency levels and between the intermediate and high proficiency levels, suggesting that comprehension of grammatical structures may be more sensitive to differences among children with TD at the higher L2 proficiency levels.

Lexical diversity assessed by NDW differed across all language proficiency levels. Adams (1980) and Iwashita et al. (2008) reported that lexical component of L2 proficiency was one of the best discriminators of L2 proficiency, especially at the lower proficiency levels. Further, syntactic complexity assessed by MLU differed only between the low and intermediate and between the low and high proficiency levels, suggesting that this measure may be more sensitive to difference in performance at the lower proficiency levels. Although Iwashita et al. (2008) reported differences in MLU across all L2 proficiency levels, the methodological differences discussed above could have affected the results. Overall, the results suggest that differences in L2 performance across L2 proficiency levels in children with TD may potentially impact the differentiation between language difficulties due to typical L2 acquisition processes and PLI, and thus, future research on bilingual children with and without PLI should control for L2 proficiency level.

Limitations and Future Directions

The major limitation of the study was the small and unequal sample size of children with PLI across length of exposure and L2 proficiency groups. As a result, the statistical power was significantly reduced and did not allow for all planned analyses. Future studies with a bigger sample size and adequate power are necessary to confirm the preliminary conclusions of the study and address the proposed hypotheses.

The study aimed to improve the diagnosis of PLI in bilingual children by investigating the contribution of L2 assessment to differentiation between the language ability groups. A priori classification of PLI and TD was based on a combination of different measures, including a standardized measure of language ability, parents and teachers' concerns about children's language development, and evaluation of grammatical accuracy in a language sample. However, the participants in this study performed lower than expected on the standardized measure of language ability for predominantly Spanish-speaking children, which resulted in the change of the cut off score in identification of PLI after the data were collected. If the cut off score was not changed, the sample could potentially over-identify more than 50% of the participants as with PLI. The adjusted cut off score in combination with other measures resulted in the classification of 16.4% of participants with PLI, which was more consistent with the 7% prevalence rate reported for monolingual children (Tomblin et al., 1997). Nevertheless, further research is necessary to investigate the validity of the interpretations that were based on a lower cut off score. It is possible that a combination of low socioeconomic status and English-only instruction, in which development of L1 is not supported, resulted in this pattern of L1 performance in this study. Overall, identification of PLI in bilingual children is challenging given that there is no gold standard.

The classification of bilingual children into L2 proficiency levels was based on agreement between a global rating scale and teachers' ratings of L2 proficiency. Previous research suggested that teachers can be influenced by children's level of academic achievement and use more stringent criteria in assigning the proficiency levels (Bedore et al., 2011; Gutiérrez-Clellen & Kreiter, 2003; Smyk et al., in preparation). Although all disagreements between the rating scale and teachers' ratings were solved by a third rater and the differences on L2 measures among L2 proficiency levels in children with TD indirectly supported the classification of proficiency levels, the agreement between the rating scale and teachers' ratings was poor. Future research should examine the validity of inferences about L2 proficiency based on teachers' ratings. Overall, identification of L2 proficiency levels is a challenging task because there are few reliable measures that provide valid score inferences (e.g., Bedore & Peña, 2008; Pray, 2005).

Another limitation of the study was calculating the length of L2 exposure based on an ordinal rather than a continuous scale from a parent questionnaire, which affected how participants were grouped on length of L2 exposure. It is also possible that parent interviews can provide more accurate information regarding the length of L2 exposure than a parent questionnaire because both an interviewer and parents can ask clarification questions. Further, the study did not include any children with PLI with less than one year of L2 exposure, which in combination with the small sample size of children with PLI, made it challenging to address the proposed hypothesis and conduct the planned analyses.

Conclusions

The study examined whether there are differences in L2 proficiency components measured by a range of language tasks in sequential bilingual children with and without PLI who were grouped on the length of L2 exposure and whether L2 proficiency measures contribute to discrimination between language ability groups at low, intermediate, and high proficiency levels. No differences between the language ability groups were found when children had between one and three years of L2 exposure. However, the grammaticality judgment task demonstrated the largest effect size among all L2 proficiency measures, suggesting that given a larger sample size and the adequate statistical power, the effect of length of exposure in performance on this task in children with and without PLI may be evident.

Among L2 measures only MLU contributed to the discrimination between the language ability groups. However, poor classification accuracy suggested that MLU alone is not a sufficient predictor of language ability in bilingual children, which is consistent with cross-linguistic research on monolingual children (Eisenberg et al., 2001; Simon-Cereijido & Gutiérrez-Clellen, 2007; Restrepo, 1998). Results also suggested that comprehension of grammatical structures and expressive grammatical task may contribute to differentiation between the language ability groups at the low and intermediate-high proficiency levels.

There were significant differences among L2 proficiency levels in children with TD on all L2 proficiency measures with the exception of verbal fluency, suggesting that L2 proficiency level may potentially impact the differentiation between language difficulties due to typical L2 acquisition processes and PLI. Future research on identification of PLI in bilingual children should include a range of measures for different L2 proficiency levels, controlling for chronological age. Overall, given a small sample size of children with PLI, further investigations of the effects of the length of L2 exposure and L2 proficiency level in identification of PLI in sequential bilingual children are necessary.

Note

[1] The term PLI is used instead of the term *specific language impairment* (SLI). The term SLI has been used in the literature to describe children whose difficulties lie only within the area of language in the absence of any other possible disorders, and thus, these difficulties are specific to the language system and cannot be explained by any other disorders (de Villiers, 2003; Leonard, 1998). However, during the last two decades the concept of SLI as a purely linguistic deficit has been questioned by a large body of research indicating that children with SLI demonstrate lower nonverbal test scores and subtle weakness in non-linguistic processing skills in comparison with typically developing peers (e.g., Hill, 2001; Kohnert & Windsor, 2004; Leonard, 1998; Restrepo, Swisher, Plante, & Vance, 1992; Swisher, Plante, & Lowell, 1994). Thus, in the present study the term PLI is preferred to account for both the subtle nonlinguistic processing weaknesses and the language deficits (Kohnert, 2008; 2010; Kohnert et al., 2009).

References

Adams, M. L. (1980). Five co-occurring factors in speaking proficiency. In J. Firth (Ed.), *Measuring spoken proficiency* (pp. 1-6) .Washington, DC: Georgetown University Press.

American National Standards Institute. (1969). *Specifications for audiometers*. New York: Author.

Anderson, R. T. (2004). First language loss in Spanish-speaking children: patterns of loss and implications for clinical practice. In B.Goldstein (Ed.), *Bilingual language development and disorders in Spanish-English speakers* (pp. 187-212). Brooks Publishing.

Armon-Lotem, S., Walters, J., & Gagarina, N. (2011). The impact of internal and external factors on linguistic performance in the home language and in L2 among Russian-Hebrew and Russian-German preschool children. *Linguistic Approaches to Bilingualism, 1*(3), 291 317.

Artiles, A.J., Rueda, R., Salazar, J.J., & Higareda, I. (2005). Within-group diversity in minority disproportional representation: English language learners in urban school districts. *Exceptional Children, 71*(3), 283-300.

Artiles, A., & Trent, S. (1994). Overrepresentation of minority students in special education: a continuing debate. *The Journal of Special Education, 27(4),* 410-437.

Aud, S., Hussar, W., Johnson, F., Kena, G., Roth, E., Manning, E., Wang, X., & Zhang, J. (2012). *The condition of education 2012* (NCES 2012-045). U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved from http://nces.ed.gov/pubsearch

Bachman, L. F. (1990). *Fundamental considerations in language testing.* Oxford: Oxford University Press.

Bachman L., & Cohen, A. D. (1998). Language testing – SLA interfaces: an update. In L. Bachman & A. D. Cohen (Eds.), *Interfaces between second language acquisition and language testing research* (pp. 1-31). Cambridge: Cambridge University Press.

Bachman, L. F, & Palmer, A. S. (1982). The construct validation of some components of communicative proficiency. *TESOL Quarterly, 16*(4), 449-465.

Baker, C. (2006). *Foundations of bilingual education and bilingualism.* Multilingual Matters Ltd.

Bedore, L. M., & Leonard, L. (1998). Specific language impairment and grammatical morphology: a discriminant function analysis. *Journal of Speech, Language, and Hearing Research, 41*(5), 1185-1192.

Bedore, L., & Leonard, L. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 44,* 905-924.

Bedore, L. M., & Peña, E. D. (2008). Assessment of bilingual children for identification of language impairment: current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism, 1*7(1), 1-29.

Bedore, L.M., Peña, E.D., Gillam, R.B., & Ho, T. (2010). Language sample measures and language ability in Spanish-English bilingual kindergarteners. *Journal of Communication Disorders*, *43*(6), 498-510.

Bedore, L. M., Peña, E.D., Joyner, D., & Macken, C. (2011). Parent and teacher rating of bilingual language proficiency and language development concerns. *International Journal of Bilingual Education and Bilingualism, 15*(5), 489-511.

Bialystok, E. (2001). *Bilingualism in development: language, literacy, and cognition.* Cambridge, UK: Cambridge University Press.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second language acquisition. *Journal of Memory and Language, 44,* 235-249.

Bishop, D.V.M. (2003). *The Test for Reception of Grammar, version 2 (TROG-2)*. London, Psychological Corporation.

Bishop, D. V. M., & Adams, C. (1992). Comprehension problems in children with specific language impairment: Literal and inferential meaning. *Journal of Speech and Hearing Research, 35*, 119-129.

Bishop, D. V. M., Bright, P., James, C., Bishop, S.J., & van der Lely, H.K.J. (2000). Grammatical SLI: A distinct subtype of developmental language impairment? *Applied Psycholinguistics, 21,* 159-181.

Bland-Stewart, L. M., & Fitzgerald, S.M. (2001). Use of Brown's 14 grammatical markers by bilingual Hispanic preschoolers: a pilot study. *Communication Disorders Quarterly, 22*(4), 171-186.

Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis, 20*(1-2), 3-49.

Blom, E., & Paradis, J. (in press). Past tense production by English second language learners with and without language impairment. *Journal of Speech, Language, and Hearing Research.*

Bohman, T., Bedore, L. M., Peña, E. D., Mendez-Pérez, A., & Gillam, R. B. (2010). What they hear and what they say: language performance in young Spanish-English bilinguals. *International Journal of Bilingualism and Bilingual Education, 13,* 325-344.

Botting, N. (2002). Narrative as a tool for the assessment of linguistic and pragmatic impairments. *Child Language Teaching and Therapy, 18,* 1-21.

Botting, N., & Conti-Ramsden, G. (2004). Characteristics of children with specific language impairment. In L. Verhoeven & H. van Balkom (Eds.), *Classification of developmental language disorders* (pp.23-38). Lawrence Erlbaum Associates. Mahwah, NJ.

Brown, R. (1973). *A first language: the early stages*. London: George Allen & Unwin Ltd.

Butler, Y. G., & Hakuta, K. (2004). Bilingualism and second language acquisition. In T. K. Bhatia & W. C. Ritchie (Eds.), *The handbook of bilingualism* (pp. 114-145). Blackwell Publishing.

Canale, M., & Swain. M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1-47.

Campbell, T., Dollaghan, C, Needleman, H., & Janosky, J. (1997). Reducing bias in language assessment: processing-dependent measures. *Journal of Speech, Language, and Hearing Research, 40*(3), 519-525.

Carrol, J.B. (1961). *Fundamental considerations in testing for English language proficiency of foreign students.* Testing. Washington, DC: Center for Applied Linguistics.

Castilla, A., Restrepo, M.A., & Perez-Leroux, A. (2009). Individual differences and language interdependence: A study of sequential bilingual development in Spanish English preschool children. *International Journal of Bilingual Education and Bilingualism, 12*(5), 565-580.

Chondrogianni, V. V., & Marinis, T. (2011). Differential effects of internal and external factors on the development of vocabulary, tense morphology and morpho-syntax in successive bilingual children. *Linguistic Approaches to Bilingualism, 1*(3), 318-345.

Cleave, P., Girolametto, L., Chen, X., & Johnson, C. (2010). Narrative abilities in bilingualchildren with specific language impairment. *Journal of Communication Disorders, 43*(6), 511-522.

Crago, M., & Paradis, J. (2004). Language impairment in children: cross-linguistic studies. In R. Kent (Ed.), *The MIT encyclopedia of communication sciences and disorders* (pp. 331-333). Cambridge, MA: MIT Press.

Crutchley, A. (1999). Bilingual children with SLI attending language units: getting the biggerpicture. *Child Language Teaching and Therapy*, *15*(3), 201-217.

Crutchley, A., Conti-Ramsden, G., & Botting, N. (1997) Bilingual children with specific language impairment and standardized assessments: preliminary findings from a study of children in language units. *International Journal of Bilingualism, 1*(2), 117-134.

Culatta, B., Page, J., & Ellis, J. (1983). Story telling as a communication performance screening tool. *Language, Speech, and Hearing Services in Schools, 14,* 66-74.

Damico, J.S., Oller, J.W., & Storey, M.E. (1983). The diagnosis of language disorders in bilingual children: pragmatic and surface-oriented criteria. *Journal of Speech and Hearing Disorders*, *48*, 285-294.

Dawson, J. I., Stout, C. E., & Eyer, J. A. (2003). *Structured Photographic Expressive Language Test: Third Edition.* DeKalb, IL: Janelle Publications.

De Bot, K. (2008). Introduction: second language development as a dynamic process. *The Modern Language Journal, 92,* 166-178.

De Bot, K., Verspoor, M., & Lowie, W. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism, Language, and Cognition, 10,* 7-21.

De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics, 28,* 411-424.

De Jong, J. (2008). Bilingualism and language impairment. In M. J. Ball, M. R. Perkins, N. Müller, & S. Howard (Eds.), *The handbook of clinical linguistics* (pp. 261-274). Blackwell Publishing.

De Jong, J. H. A. L., & van Ginkel, L. W. (1992). Dimensions in oral foreign language proficiency. In J. H. A. L. De Jong (Ed.), *The construct of language proficiency* (pp. 112-40). Philadelphia, John Benjamin.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*(4), 499–533.

De Villiers, J. G. (2003). Defining SLI: a linguistic perspective. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations. Towards a definition of specific language impairment* (pp. 425-447). Lawrence Erlbaum Associates, Mahwah, NJ.

Dixon, L. Q. (2011). The role of home and school factors in predicting English vocabulary among bilingual kindergarten children in Singapore. *Applied Psycholinguistics, 32,* 141-168.

Dollaghan, C.A., & Horner, E.A. (2011). Bilingual language assessment: a meta-analysis of diagnostic accuracy. *Journal of Speech, Language, and Hearing Research, 54,* 1077-1088.

Dunn, M., Flax, J., Sliwinski, M., & Aram, D. (1996). The use of spontaneous language measures as criteria for identifying children with specific language impairment: an attempt to reconcile clinical and research incongruence. *Journal of Speech and Hearing Research, 39*, 643- 655.

Duran, P., Malvern, D., Richards, B., & Chipere, N. (2004). Developmental trends in lexical diversity. *Applied Linguistics, 25*, 220-242.

Duursma, E., Romero-Contreras, S., Szuber, A., Proctor, P., & Snow, C. E. (2007). The role of home literacy and language environment on bilinguals' English and Spanish vocabulary development. *Applied Psycholinguistics, 28,* 171-190.

Eisenberg, S. L., McGovern Fersko, T. M., & Lundgren, C. (2001). The use of MLU for identifying language impairment in preschool children: a review. *American Journal of Speech-Language Pathology, 10,* 323-342.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Ellis, R., & Barkhuizen, G. (2005). *Analyzing learner language*. Oxford University Press.

Ellis, R., & Yuan, F. (2005).The effects of careful within-task planning on oral and written task performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 167- 193). Amsterdam: Benjamins.

Evans. J. L. (1996). Plotting the complexities of language sample analysis: linear and nonlinear dynamical models of assessment. In K. N. Cole, P. S. Dale, & D. J. Thal (Eds.), *Assessment of communication and language* (pp. 207-256). Baltimore: P.H. Brookes.

Evans, J. L. (2001). An emergent account of language impairments in children with SLI: implications for assessment and intervention. *Journal of Communication Disorders, 34,* 39-54.

Evans, J. L. (2002). Variability in comprehension strategy use in children with specific language impairments: a dynamical systems account. *International Journal of Language and Communication Disorders*, *37*(2), 95-116.

Evans, M.A. (1985). Self-initiated speech repairs: A reflection of communicative monitoring in young children. *Developmental Psychology, 21*(2), 365-371.

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology Research and Practice, 40,* 532-538.

Fiestas, C. F., Bedore, L. M., Peña, E. D., & Nagy, V. J. (2005). Use of mazes in the narrative language samples of bilingual and monolingual 4- to 7-year old children. In J. Cohen, K. T. McAlister, K. Rolstad, & J. MacSwan (Eds.), *ISB4: Proceedings of the 4th international symposium on bilingualism* (pp. 730-740). Somerville, MA: Cascadilla Press.

Fiestas, C.F., & Peña E.D. (2004). Narrative discourse in bilingual children. *Language, Speech, and Hearing Services in Schools, 35*, 155-168.

Fogel, A., & Thelen, E. (1987). The development of expressive and communicative action in the first year: reinterpreting the evidence from a dynamic systems perspective. *Developmental Psychology, 23,* 747-761.

Foster, P., & Tavakoli, P. (2009). Native speakers and task performance: comparing effects on complexity, fluency, and lexical diversity. *Language Learning, 59*(4), 866-896.

Francis, N. (2005). Research findings on early first language attrition: implications for the discussion on critical periods in language acquisition. *Language Learning, 55,* 491-531.

Gazella, J., & Stockman, I. J. (2003). Children's story retelling under different modality and task conditions: implications for standardizing language sampling procedures. *American Journal of Speech-Language Pathology, 12,* 61-72.

George, D., & Mallery, P. (2003). *SPSS for windows step by step a simple guide and reference 11.0 update*. Allyn and Bacon, Boston.

Ginther, A., Dimova, S., & Yang, R. (2010). Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. *Language Testing, 27,* 379-399.

Givon, T. (1991). Markedness in grammar: distributional, communicative and cognitive correlates of syntactic structure. *Studies in Language*, *15*(2), 335-370.

Goffman, L., & Leonard, J. (2000). Growth of language skills in preschool children with specific language impairment: implications for assessment and intervention. *American Journal of Speech-Language Pathology, 9,* 151-161.

Golberg, H., Paradis, J., & Crago, M. (2008). Lexical acquisition over time in minority L1 children learning English as a L2. *Applied Psycholinguistic*s, *29*(1), 41-65.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*, 3-15.

Grüter, T. (2005). Comprehension and production of French object clitics by child second language learners and children with specific language impairment. *Applied Psycholinguistics, 26*(3), 363-391.

Gutiérrez-Clellen, V. F. (1996). Language diversity: implications for assessment. In K. Cole, P. Dale, & D. Thal (Eds.), *Advances in assessment of communication and language* (pp. 29–56). Baltimore MD: Paul H. Brookes Publishing.

Gutiérrez -Clellen, V.F. (2002). Narratives in two languages: assessing performance of bilingual children. *Linguistics and Education, 13,* 175-197.

Gutierrez-Clellen, V.F., & Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics, 24*, 267-288.

Gutiérrez-Clellen, V.F., & Simon-Cereijido, G. (2009). Using language sampling in clinical assessments with bilingual children: challenges and future directions. *Seminars in Speech and Language, 30,* 234-245.

Gutiérrez-Clellen, V.F., Simon-Cereijido, G., & Sweet, M.(2012). Predictors of language acquisition in Latino children with specific language impairment. *American Journal of Speech-Language Pathology, 21*(1), 64-77.

Gutiérrez-Clellen, V.F., Simon-Cereijido, G., & Wagner, C. (2008). Bilingual children with language impairment: a comparison with monolinguals and second language learners. *Applied Psycholinguistics, 29,* 3-19.

Håkansson, G. (2001). Tense morphology and verb-second in Swedish L1 children, L2 children and children with SLI. *Bilingualism: Language and Cognition, 4*, 85-99.

Håkansson, G., & Nettlebladt, U. (1996). Similarities between SLI and L2 children: evidence from the acquisition of Swedish word order. In C.E. Johnson & J.H.V. Gilbert (Eds.), *Children's language* (pp. 135-151). Mahwah, NJ: Lawrence Erlbaum Associates.

Håkansson, G., Salameh, E-K., & Nettelbladt, U. (2003). Measuring language development in bilingual children: Swedish-Arabic children with and without language impairment. *Linguistics, 41*(2), 255-288.

Halleck, G. B. (1995). Assessing oral language proficiency: comparison of holistic and objective measures. *The Modern Language Journal*, *79*, 223-234.

Heilmann, J., Miller, J., Iglesias, A., Fabiano-Smith, I., Nockerts, A., & Andriacchi, K. (2008). Narrative transcription accuracy and agreement in two languages. *Topics in Language Disorders, 28,* 178-187.

Heilmann, J., Nockerts, A., & Miller, J. (2010). Language sampling: does the length of the transcript matter? *Language, Speech, and Hearing Services in Schools, 41,* 393-404.

Herdina, P., & Jessner, U. (2002). *A dynamic model of multilingualism*. Clevedon: Multilingual Matters LTD.

Hewitt, L., Hammer, C.S., Yont, C., & Tomblin, J.B. (2005). Language sampling for kindergarten children with and without SLI: mean length of utterance, IPSYN, and NDW. *Journal of Communication Disorders, 38,* 197-213.

Hickey, T. (1991). Mean length of utterance and the acquisition of Irish. *Journal of Child Language, 3,* 553-569.

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 and Communication Disorders, 36,* 149-171.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65-70.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics, 30*(4), 461-473.

Huberty, C. J, & Morris, J. D. (1989). Multivariate analysis versus multiple univariate analyses. *Psychological Bulletin*, *105,* 302-308.

Hunt, K. W. (1970). *Syntactic maturity in school children and adults.* Monographs of the Society for Research in Child Development (35/1, Serial No. 134). Chicago: University of Chicago Press.

Iwashita, N. (2006). Syntactic complexity measures and their relation to oral proficiency in Japanese as a foreign language. *Language Assessment Quarterly, 2*(4), 151-170.

Iwashita, N. (2010). Features of oral proficiency in task performance by EFL and JFL learners. In M. T. Prior, Y. Watanabe, & S.-K. Lee (Eds.), *Selected proceedings of the 2008 second language research forum* (pp. 32-47). Somerville, MA: Cascadilla Proceedings Project.

Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: how distinct? *Applied Linguistics, 29*(1), 24-49.

Jaccard, J., & Guillamo-Ramos, V. (2002). Analysis of variance frameworks in clinical child and adolescent psychology: issues and recommendations. *Journal of Clinical Child and Adolescent Psychology, 31*, 130-146.

Jacobson, P., & Livert, D. (2010). English past tense use as a clinical marker in older bilingual children with language impairment. *Clinical Linguistics and Phonetics, 24*(2), 101-121.

Jacobson, P. F., & Schwartz, R. G. (2005). English past tense use in bilingual children with language impairment. *American Journal of Speech-Language Pathology, 14,* 313-323.

Jessner, U. (2008). A DST model of multilingualism and the role of metalinguistic awareness. *The Modern Language Journal, 92,* 270-283.

Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics, 24*, 131-161.

Jia, G., & Fuse, A. (2007). Acquisition of English grammatical morphology by native Mandarin speaking children and adolescents: age-related differences. *Journal of Speech, Language, and Hearing Research, 50,* 1280-1299.

Jia, G., Kohnert, K., Collado, J., & Aquino-Garcia, F. (2006). Action naming in Spanish and English by sequential bilingual children and adolescents. *Journal of Speech, Language, and Hearing Research, 49,* 588-602.

Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: a developmental perspective. *Working Papers, 53*, 61-79. Lund University, Department of Linguistics and Phonetics.

Johnson, J.S., & Newport, E.L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21,* 60-99.

Kan, P. F., & Kohnert, K. (2005). Preschoolers learning Hmong and English: lexical–semantic skills in L1 and L2. *Journal of Speech, Language, and Hearing Research, 48*, 372-378.

Kester Stubbe E., & Peña, E.D. (2002). Language ability assessment of Spanish-English bilinguals: future directions. *Practical Assessment, Research and Evaluation*, *8*(4). Retrieved from http://PAREonline.net/getvn.asp?v=8&n=4

Klee, T. (1992). Developmental and diagnostic characteristics of quantitative measures of children's language production. *Topics in Language Disorders, 12*(2), 28-41.

Klee, T., Schafer, M., May, S., Membrino, I., & Mougey, K. (1989). A comparison of the age MLU relations in normal and specifically language impaired preschool children. *Journal of Speech and Hearing Disorders, 54,* 226-233.

Kohnert, K. (2004). Processing skills in early sequential bilinguals. In B.A. Goldstein (Ed.), *Bilingual language development and disorders in Spanish-English speakers* (pp. 53-76). Paul H. Brooks Publishing Co, Baltimore.

Kohnert, K. (2008). *Language disorders in bilingual children and adults*. San Diego: Plural Publishing.

Kohnert, K. (2010). Bilingual children with primary language impairment: issues, evidence and implications for clinical actions. *Journal of Communication Disorders, 43*(6), 456-73.

Kohnert, K., Bates, E., & Hernandez, A. E. (1999). Balancing bilinguals: Lexical-semantic production and cognitive processing in children learning Spanish and English. *Journal of Speech, Language, and Hearing Research, 42,* 1400-1413.

Kohnert, K., & Danahy, K. (2007). Young L2 learners' performance on a novel morpheme task. *Clinical Linguistics and Phonetics, 21*, 557-569.

Kohnert, K., Kan, P.F., & Conboy, B. (2010). Lexical and grammatical associations in sequential bilingual preschoolers. *Journal of Speech, Language, and Hearing Research, 53,* 684-698.

Kohnert, K., & Pham, G. (2010). The process of acquiring a first and second language. In M. Shatz & L.C. Wilkinson (Eds.), *The education of English language learners. Research to practice* (pp. 48-66).The Guilford Press, NY.

Kohnert, K., & Windsor, J. (2004). The search for common ground part II: nonlinguistic performance by linguistically diverse learners. *Journal of Speech, Language, and Hearing Research, 47,* 891-903.

Kohnert, K., Windsor, J., & Ebert, K. D. (2009). Primary of "specific" language impairment and children learning a second language. *Brain and Language*, *109,* 101-111.

Kohnert, K., Windsor, J., & Yim, D. (2006). Do language-based processing tasks separate children with language impairment from typical bilinguals? *Learning Disabilities Research and Practice, 21,* 19-29.

Kormos, J., & Denes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32,* 145-164.

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics, 18*(2), 141-165.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics, 27*(4), 590-619.

Leadholm, B.J., & Miller, J. (1995). *Language sample analysis: The Wisconsin guide.* Madison, WI: Wisconsin Department of Public Health.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, *40,* 387-417.

Le Normand, M. T., Leonard, L., & McGregor, K. (1993). A cross-linguistic study of article use by children with specific language impairment. *European Journal of Disorders of Communication, 28,* 153-163.

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

Leonard, L., Bortolini, U., Caselli, C., McGregor, K., & Sabbadini, L. (1992). Morphological deficits in children with specific language impairment: the status of features in the underlying grammar. *Language Acquisition, 2,* 151-179.

Leonard, L., Miller, C., & Gerber, E. (1999). Grammatical morphology and the lexicon in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 678-689.

MacSwan, J. (2000). The threshold hypothesis, semilingualism, and other contributions to a deficit view of linguistic minorities. *Hispanic Journal of Behavioral Sciences, 22,* 3-45.

MacSwan, J., & Rolstad, K. (2006). How language tests mislead us about children's abilities: implications for special education placements. *Teachers College Record*, *108*(11), 2304-2328.

Mancilla-Martinez, J. J., & Lesaux, N. K. (2011). Early home language use and later vocabulary development. *Journal of Educational Psychology, 103*(3), 535-546.

Mandler, J. M., Scribner, S., Cole, M., & DeForest, M. (1980). Cross-cultural invariance in story recall. *Child Development, 51*, 19-26.

Marinis, T., & Chondrogianni, V. (2010). Production of tense marking in successive bilingual children: when do they converge with their monolingual peers? *International Journal of Speech-Language Pathology*, *12*(1), 19-28.

Mayer, M. (1973). *Frog on his own*. NewYork, NY: Dial Books for Young Readers.

Mayer, M. (1974). *Frog goes to dinner*. NewYork, NY: Dial Books for Young Readers.

Mayer, M., & Mayer, M. (1971). *A boy, a dog, a frog, and a friend.* New York: Dial Books for Young Readers.

McCarthy, P., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: a validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods, 42,* 381-392.

McDonald, J.L., & Roussel, J. L., C.C. (2010). Past tense grammaticality judgment and production in non-native and stressed native English speakers. *Bilingualism: Language and Cognition, 13,* 429-448.

McNamara, T. (1996). *Measuring second language performance*. London: Longman.

Miller, J. F., Heilmann, J., Nockerts, A., Iglesias, A., Fabiano, L., & Francis, D. J. (2006). Oral language and reading in bilingual children. *Learning Disabilities Research and Practice*, *21*, 30-43.

Miller, J. F., & Iglesias, A. (2008). Systematic analysis of language transcripts (SALT), English & Spanish (Version 9) [Computer software]. Madison: University of Wisconsin-Madison, Waisman Center, Language Analysis Laboratory.

Muñoz, M.L., Gillam, R.B., Peña, E.D., & Gulley-Faehnle, A. (2003). Measures of language development in fictional narratives of Latino children. *Language, Speech, and Hearing Services in Schools, 34*, 332-342.

National Clearinghouse for English Language Acquisition (NCELA; 2010). *The growing numbers of limited English proficient students*. Retrieved from http://www.ncela.gwu.edu/files/uploads/9/growingLEP_0708.pdf

Navarro-Ruiz, M. I., & Rallo-Fabra, L. (2001). Characteristics of mazes produced by SLI children. *Clinical Linguistics and Phonetics, 15*(1-2), 63-66.

Norbury, C.F., Bishop, D.V.M., & Briscoe, J. (2002). Does impaired grammatical comprehension provide evidence for an innate grammar module? *Applied Psycholinguistics, 23 ,* 247-268.

Norris, J., & Ortega, L. (2003). Defining and measuring SLA. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 717-761). Blackwell Publishing.

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: the case of complexity. *Applied Linguistics, 30*(4), 555-578.

Oetting, J, B., & Rice, M. L. (1993). Plural acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 36,* 1236-1248.

Oller, J.W., Jr., & Damico, J.S. (1991).Theoretical considerations in the assessment of LEP students. In E. Hamayan & J.S. Damico (Eds.), *Limiting bias in the assessment of bilingual students* (pp. 77-110). Austin, TX: Pro-Ed Publications.

Oller, D. K., & Eilers, R. E. (2002). *Language and literacy in bilingual children*. Clevedon, UK: Multilingual Matters.

Owen, A. J., & Leonard, L. B. (2002). Lexical diversity in the spontaneous speech of children with specific language impairment: application of D. *Journal of Speech, Language, and Hearing Research, 45,* 927-937.

Orgassa, A. & Weerman, F. (2008) Dutch gender in specific language impairment and second language acquisition. Second Language Research, 24, 333-364.

Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics, 30*(4), 590-601.

Paradis, J. (2005). Grammatical morphology in children learning English as a second language: implications of similarities with specific language impairment. *Language, Speech, and Hearing Services in Schools, 36,* 172-187.

Paradis, J. (2008). Tense as a clinical marker in English L2 acquisition with language delay/impairment. In E. Gavruseva & B. Haznedar (Eds.), *Current trends in child second language acquisition: a generative perspective* (pp. 337-356). Amsterdam: John Benjamins.

Paradis, J. (2010). The interface between bilingual development and specific language impairment. *Applied Psycholinguistics, 31,* 3-28.

Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-Internal and child-external factors. *Linguistic Approaches to Bilingualism, 1*(3), 213-237.

Paradis, J., & Crago, M. (2000). Tense and temporality: Similarities and differences between language-impaired and second-language children. *Journal of Speech, Language, and Hearing Research*, *43*(4), 834-848.

Paradis, J., & Crago, M. (2001). The morphosyntax of specific language impairment in French: evidence for an Extended Optional Default account. *Language Acquisition, 9,* 269-300.

Paradis, J., Crago, M., Genesee, F. & Rice, M. (2003). Bilingual children with specific language impairment: How do they compare with their monolingual peers? *Journal of Speech, Language, and Hearing Research, 46,* 1-15.

Paradis, J., Emmerzael, K., & Sorenson Duncan, T. (2010). Assessment of English language learners: using parent report on first language development. *Journal of Communication Disorders, 43,* 474-497.

Paradis, J., Rice, M., Crago, M., & Marquis, J. (2008). The acquisition of tense in English: distinguishing child L2 from L1 and SLI. *Applied Psycholinguistics*, *29,* 1-34.

Peña, E. D., & Bedore, L. M. (2009). Bilingualism in child language disorders. In R. G. Schwartz (Ed.), *Handbook of child language disorders* (pp. 281-307). Psychology Press: New York.

Peña, E., Bedore, L., & Rappazzo, C. (2003). Comparison of Spanish, English and bilingual children's performance across semantic types. *Speech-Language-Hearing Services in Schools, 34*, 5-16.

Peña, E. D., Gillam, R. B., Bedore, L., & Bohman, T. M. (2011). Risk for poor performance on a language screening measure for bilingual preschoolers and kindergarteners. *American Journal of Speech-Language Pathology, 20,* 302-314.

Perona, K., Plante, E., & Vance, R. (2005). Diagnostic accuracy of the Structured Photographic Expressive Language Test: Third edition (SPELT-3). *Language, Speech, & Hearing Services in Schools, 36,* 103-115.

Plante, E., & Vance, R. (1994). Selection of preschool language tests: a data based approach. *Language, Speech, and Hearing Services in Schools*, *25*, 15-23.

Plaza-Pust, C. (2008). Dynamic systems theory and universal grammar: holding up a turbulent mirror to development in grammars. *The Modern Language Journal, 92,* 250-269.

Polite, E. J., & Leonard, L. (2006). Finite verb morphology and phonological length in the speech of children with specific language impairment. *Clinical Linguistics and Phonetics, 20,* 751-760.

Pray, L. (2005). How well do commonly used language instruments measure English oral-language proficiency*? Bilingual Research Journal, 29*(2), 387-409.

Price, J.R., Roberts, J.E., & Jackson, S.C. (2006). Structural development of the fictional narratives of African American preschoolers. *Language, Speech, and Hearing Services in School*, 37*,* 178-190.

Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research, 41*(6), 1398-1411.

Restrepo, M. A., Castilla, A. P., Schwanenflugel, P., Neuharth-Pritchett, S., Hamilton, C., & Arboleda, A. (2010). Sentence length, complexity and grammaticality growth in Spanish speaking children attending English only and bilingual preschool programs. *Language Speech and Hearing Services in Schools, 41,* 3-13.

Restrepo, M. A., & Gutiérrez -Clellen, V. F. (2001). Article use in Spanish speaking children with specific language impairment. *Journal of Child Language, 28*(2), 433-452.

Restrepo, M.A., & Gutiérrez-Clellen, V.F. (2012). Grammatical impairments in Spanish-English bilingual children. In B. A. Goldstein (Ed.), *Bilingual language development and disorders in Spanish–English speakers* (pp. 213-232). Brookes Publishing Company.

Restrepo, M. A., & Kruth, K. (2000). Grammatical characteristics of a Spanish-English bilingual child with specific language impairment. *Communication Disorders Quarterly, 21,* 66-76.

Restrepo, M. A., Swisher, L., Plante. E., & Vance, R. (1992). Relations among verbal and nonverbal cognitive skills in normal language and specifically language impaired children. *Journal of Communication Disorders, 25,* 205-219.

Riazantseva, A. (2001). Second language proficiency and pausing: a study of Russian speakers of English. *Studies in Second Language Acquisition, 23,* 497-526.

Rice, M., Noll, R. K., & Grimm, H. (1997). An extended optional infinitive stage in German speaking children with SLI. *Language Acquisition, 6*(4), 255-296.

Rice, M.L., Hoffman, L., & Wexler, K. (2009). Judgments of omitted BE and DO in questions as extended finiteness clinical markers of SLI to fifteen years: A study of growth and asymptote. *Journal of Speech, Language, and Hearing Research, 52,* 1417-1433.

Rice, M., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English speaking children. *Journal of Speech and Hearing Research, 39,* 1239-1257.

Rice, M. L., & Wexler, K. (2001). *Rice/ Wexler Test of Early Grammatical Impairment.* San Antonio, TX: The Psychological Corporation.

Rice, M. L., Wexler, K., & Redmond, S. M. (1999). Grammaticality judgments of an extended optional infinitive grammar: evidence from English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*, 943-961.

Rice, M.L., Redmond, S.M., Hoffman, L. (2006). MLU in children with SLI and younger control children shows concurrent validity, stable and parallel growth trajectories. *Journal of Speech, Language, and Hearing Research, 49,* 793-808.

Rice, M. L., Smolik, F., Perpech, D., Thompson, T., Rytting, N., & Blossom, M.(2010). Mean Length of Utterance Levels in 6-Month Intervals for Children. *Journal of Speech, Language, and Hearing Research, 53,* 333-349.

Richards, B. J. (1987). Type/token ratios: what do they really tell us? *Journal of Child Language, 14,* 201-209.

Robinson, P. (2007). Task complexity, task difficulty, and task production: exploring interactions in a componential framework. *Applied Linguistics, 22*, 27-57.

Rojas, P. (2011). *Waves of Change: longitudinal growth profiling of bilingual (Spanish-English) language development*. (Unpublished doctoral dissertation). Temple University, Philadelphia, PA.

Salameh, E.-K., Håkansson, G., & Nettelbladt, U. (2004). Developmental perspectives on bilingual Swedish- Arabic children with and without language impairment: A longitudinal study. *International Journal of Language and Communication Disorders, 39,* 65-91.

Samson, J., & Lesaux, N.K. (2009). Language minority learners in special education: rates and predictors of identification for services. *Journal of Learning Disabilities, 42*(2), 148-162.

Saxton, M., Dockrell, J., Bevan, E., & Van Herwegen, J. (2007) Grammaticality judgements of children with and without language delay. *30th Child Language Seminar,* 136-150. Reading, U.K.

Scheele, A.H., Laseman, P.P.M., & Mayo, A.Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics, 31*, 117-140.

Schiff-Myers, N. (1992). Considering arrested language development and language loss in the assessment of second language learners. *Language, Speech, and Hearing Services in Schools*, *23,* 28-33.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition, 14*, 357-385.

Schwartz, R. G. (2009). Specific language impairment. In R. G. Schwartz (Ed.), *Handbook of child language disorders* (pp. 3-43). New York: Taylor and Francis.

Semel, E., Wiig, E. H., & Secord, W. H. (2006). *Clinical evaluation of language fundamentals Spanish edition (CELF-4).* San Antonio, TX: Psychological Corporation.

Simon-Cereijido, G., & Gutiérrez-Clellen, V.F. (2007) Spontaneous language markers of Spanish language impairment. *Applied Psycholinguistics*, *28*(3), 317-339.

Simon-Cereijido, G., & Gutiérrez-Clellen, V.F. (2009) A cross-linguistic and bilingual evaluation of the interdependence between lexical and grammatical domains. *Applied Psycholinguistics, 30,* 315-338.

Skehan, P. (1998). *A cognitive approach to language testing*. Oxford: Oxford University Press.

Skehan, P. (2009). Modeling second language performance: integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510-532.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning, 49,* 93-120.

Smith, L.B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences, 7*(8), 343-348.

Smyk, E., Restrepo, M.A., Gorin, J., & Gray, S. (in preparation). *The development of Spanish-English Language Proficiency Scale for children.*

Smyk, E., Restrepo, M.A., Gorin, G., & Gray, S. (June 2011). *Developing a language proficiency measure: the potential confound of language ability.* Poster presented at the annual Symposium on Research in Child Language Disorders, Madison, WI.

Spoelman, M., & Verspoor, M. (2010). Dynamic patterns in development of accuracy and complexity: a longitudinal case study in the acquisition of Finnish. *Applied Linguistics*, *31*(4), 532-553.

Stockman, I. J. (1996).The promises and pitfalls of language sample analysis as an assessment tool for linguistic minority children. *Language, Speech, and Hearing Services in Schools, 27,* 355-366.

Sullivan, A.L. (2011). Disproportionality in special education identification and placement of English language learners. *Exceptional Children, 77*(3), 317-334.

Swasey Washington, P. (2010). *Factors affecting norming: a developmental study of oral language measures in Spanish-speaking English language learners.* (Unpublished doctoral dissertation). Temple University, Philadelphia, PA.

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

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*, *5th ed.* Boston: Allyn and Bacon.

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239-274). John Benjamins.

Thordardottir, E. T., & Weismer, S. E. (1998). Mean length of utterance and other language sample measures in early Icelandic. *First Language, 18,* 1-32.

Thordardottir, E.T., & Weismer, S.E. (2002). Content mazes and filled pauses on narrative language samples of children with specific language impairment. *Brain and Cognition, 48*(1-2), 587-592.

Thordardottir, E.T., Weismer, S.E., & Smith, M.E. (1997) Vocabulary learning in bilingual and monolingual clinical intervention. *Child Language Teaching and Therapy, 13*(3), 215-227.

Tilstra, J., & McMaster, K. (2007). Productivity, fluency, and grammaticality measures from narratives: potential indicators of language proficiency? *Communication Disorders Quarterly, 29,* 43-53.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research, 40*(6), 1245-1260.

Uccelli, P., & Paez, M. M. (2007). Narrative and vocabulary development of bilingual children from kindergarten to first grade: developmental changes and associations among English and Spanish skills. *Language, Speech, and Hearing Services in Schools, 38*(3), 225-236.

Umbel, V. M., Pearson, B. Z., Fernández, M. C., & Oller, D. K. (1992). Measuring bilingual children's receptive vocabularies. *Child Development, 63,* 1012-1020.

U.S. Department of Education and National Institute of Child Health and Human Development. (2003, October). *National symposium on learning disabilities in English language learners.* Washington, DC. Retrieved from http://www.nichd.nih.gov/publications/pubs/upload/ELL_summary.pdf

Van der Lely, H. K., Jones M., & Marshall C. R. (2011). Who did Buzz see someone? Grammaticality judgement of wh-questions in typically developing children and children with Grammatical-SLI. *Lingua, 121*(3), 408-422.

Van Geert, P. (2004). A dynamic systems approach to diagnostic measurement of SLI. In L. Verhoeven & H. Van Balkom (Eds.), *Classification of developmental language disorders* (pp.327-348). Mawah (NJ): Erlbaum.

Van Geert, P. (2008). The dynamic systems approach in the study of L1 and L2 acquisition: An introduction. *Modern Language Journal, 92*(2), 179-199.

Van Dijk, M., & van Geert, P. (2005). Disentangling behavior in early child development: Interpretability of early child language and its effect on utterance length measures. *Infant Behavior and Development*, *28*, 99-117.

Verhoeven, L., Steenge, J., & van Balkom, H. (2011). Verb morphology as clinical marker of specific language impairment: evidence from first and second language learners. *Research in Developmental Disabilities, 32*(3), 1186-1193.

Verhoeven, L., Steenge, J., van Weerdenburg, M., & van Balkom, H. (2011). Assessment of second language proficiency in bilingual children with specific language impairment: a clinical perspective. *Research in Developmental Disabilities, 32*(5), 1798-1807.

Vermeer, A. A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing, 17*(1), 65-83.

Verspoor, M., Lowie, W., & Van Dijk, M. (2008) Variability in second language development from a dynamic systems perspective. *Modern Language Journal, 92*(2), 214-231.

Watkins, R. V., Kelly, D. J., Harbers, H. M., & Hollis, W. (1995). Measuring children's lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech and Hearing Research, 38,* 1349-1355.

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

Willig, A. (1986). Special education and the culturally and linguistically different child: an overview of issues and challenges. *Reading, Writing, and Learning Disabilities, 2,* 161-173.

Windsor, J., & Kohnert, K. (2004). The search for common ground – Part I: lexical performance of linguistically diverse learners. *Journal of Speech, Language, and Hearing Research, 47,* 977-890.

Wong, A.M,, Klee, T., Stokes, S.F., Fletcher, P., & Leonard, L.B. (2010). Differentiating Cantonese-speaking preschool children with and without SLI using MLU and lexical diversity (D). *Journal of Speech, Language, and Hearing Research, 53*(3), 794-799.

Yu, G. G. (2010). Lexical diversity in writing and speaking task performances. *Applied Linguistics, 31*(2), 236-259.

Zareva, A., Schwanenflugel, P., Nikolova, Y. (2005). Relationship between lexical competence and language proficiency: variable sensitivity. *Studies in Second Language Acquisition, 27*(4), 567-595.

Table 1
Descriptive Statistics for the CELF[a], SPELT[b], WNV[c], SELPS[d], and Teachers' Ratings of L2 proficiency

	PLI						TD					
	M (*SD*)	Median	Skews (SE)	Kurtosis (SE)	Min	Max	M (*SD*)	Median	Skew (SE)	Kurtosis (SE)	Min	Max
CELF	58.33 (6.05)	61	-1.13 (0.64)	-0.29 (1.23)	47	65	89.92 (12.67)	88	0.44 (0.31)	-.44 (0.60)	68	122
SPELT	47.17 (14.94)	51	-0.25 (0.64)	-1.28 (1.23)	24	70	63.79 (19.55)	65.5	-0.58 (0.31)	0.88 (0.62)	7	112
WNV	100.5 (11.53)	102.5	-0.06 (0.64)	-0.91 (1.23)	82	118	100.46 (11.33)	101	0.28 (0.31)	-.28 (0.61)	76	128
SELPS English	3.5 (0.75)	3.75	-0.60 (0.64)	-0.54 (1.23)	2	4.5	3.79 (0.82)	4	-1.18 (0.31)	1.20 (0.60)	1	5
SELPS Spanish	4.08 (0.70)	4.25	-2.72 (0.64)	8.33 (1.23)	2	4.5	4.72 (0.25)	4.5	0.24 (0.30)	-2.01 (0.60)	4.5	5
TR[e]	1.63 (.52)	2	-.64 (.75)	-2.24 (1.48)	1	2	1.87 (.67)	2	.16 (.31)	-.72 (.60)	1	3

[a] Clinical Evaluation of Language Fundumentals - Spanish. [b] Structured Photographic Expressive Language Test. [c] Wechsler Nonverbal Scale of Ability. [d] Spanish-English Language Proficiency Scales. [e] Teachers' ratings of L2 proficiency.

Table 2
Correlations among the Selection Measures, Teachers' Ratings of L2 Proficiency, Chronological Age, Age and Length of L2 Exposure for the Language Ability Groups

		CELF	SPELT	WNV	SELPS English	SELPS Spanish	TR	Age	Age of exposure	Length of exposure
PLI	CELF[a]	1								
	SPELT[b]	.08	1							
	WNV[c]	.19	-.45	1						
	SELPS[d] English	.62*	.34	-.14	1					
	SELPS Spanish	.52	-.33	.43	.57	1				
	TR[e]	-.47	.22	-.003	-.25	-.49	1			
	Age	.68*	.30	.11	.59*	.37	.05	1		
	Age of exposure	.60	.32	-.08	.60	-.02	.26	.49	1	
	Length of exposure	-.04	.08	-.15	.01	-.05	-.04	.44	-.56	1
TD	CELF	1								
	SPELT	.14	1							
	WNV	.23	.27*	1						
	SELPS English	-.02	.25	.19	1					
	SELPS Spanish	.16	-.25	-.14	-.21	1				
	TR	.06	.31*	.02	.42**	-.32*	1			
	Age	-.24	-.06	.09	.67**	-.22	.15	1		
	Age of exposure	-.03	-.26	.08	-.14	.12	-.17	.09	1	
	Length of exposure	-.12	.21	.03	.56**	-.23	.25	.60**	-.74**	1

Note.* Correlation is significant at the .05 level (2-tailed). ** Correlation is significant at the .01 level (2-tailed).
[a] Clinical Evaluation of Language Fundamentals – Spanish. [b] Structured Photographic Expressive Language Test. [c] Wechsler Nonverbal Scale of Ability. [d] Spanish - English Language Proficiency Scales. [e] Teachers' Ratings of L2 proficiency.

Table 3
Teachers and Parents' Concerns about Participants' Language Development for each Language Ability Group

		Percent of children			
		PLI		TD	
		Yes	No	Yes	No
PQ[1]	Concerns about child's speaking[3]	33.3	66.7	9.8	88.5
	Teacher/relative's concerns about child's speaking	25	75	8.2	91.8
	Other people's difficulties understanding the child	25	75	1.6	98.4
	Child's difficulties explaining/describing things	41.7	58.3	13.1	86.9
	Child speaks differently comparing to other children	33.3	66.7	8.2	91.8
	Child's vocabulary difficulties	41.7	58.3	4.9	95.1
	Child's use of correct sentences	66.7	33.3	82	18
	Child's speech difficulties	50	50	11.5	88.5
	Child's difficulties understanding questions	33.3	66.7	4.9	95.1
	Child's learning difficulties[3]	33.3	66.7	6.6	93.3
TQ[2]	Speech development[4]	25	75	19.6	75.4
	Oral language[4]	25	75	31.2	63.9
	Literacy development[5]	41.7	58.3	39.3	54.1
	Social skills[6]	8.3	58.3	19.7	75.4

Note. [1] PQ = parent questionnaire.
[2] TQ= teacher questionnaire.
[3] 1.6% of parents in the group with TD did not report that information.
[4] 4.9% of teachers did not report that information in the group with TD.
[5] 6.6% of teachers did not report that information in the group with TD.
[6] 4.9% of teachers did not report that information in the group with TD and 33.3% did not report that information in the group with PLI.

Table 4

Descriptive Statistics for Chronological Age for each Language Ability and L2 Exposure Group

Length of exposure	PLI		TD	
	N	M_{age} years; months (*SD* months)	N	M_{age} years; months (*SD* months)
0-1 years	-	-	5	5;8 (2.88)
> 1 and up to 2 years	2	6;6 (7.78)	10	6 (6.18)
> 2 and up to 3 years	6	6;9 (9.91)	24	6;8 (10.37)
> than 3 years	3	7;3 (2.08)	21	7;2 (7.72)
Missing data	1	-	1	-

Table 5
Descriptive Statistics for Experimental Measures for each Language Ability and L2 Exposure Group

	Measure	PLI				TD			
		N	M (SD)	Min	Max	N	M (SD)	Min	Max
> 1-3 years	MLU[a]	8	6.05 (1.17)	3.97	7.27	33	6.61 (1.34)	4.06	10.52
	Gram. errors[b]	8	.48 (.32)	.19	1.14	33	.56 (.32)	.03	1.27
	NDW[c]	8	71 (16.43)	52	92	33	72.76 (21.42)	39	126
	PMW[d]	8	14.88 (12.71)	3	42	33	16.79 (9.87)	1	40
	Gram. judg.[e]	5	1.37 (.95)	0	2.63	34	2 (.61)	.83	3
	TEGI[f]	5	36.8 (20.39)	2	50	34	47.97 (27.55)	10	98
	TROG-2[g]	5	3.4 (1.52)	1	5	34	5.18 (3.19)	0	12
> 3 years	MLU	3	6.63 (.96)	5.52	7.27	21	7.21 (.99)	5.43	9
	Gram. errors	3	.54 (.22)	.34	.78	21	.32 (.21)	.05	.94
	NDW	3	73.67 (11.02)	61	81	21	88.86 (24.41)	51	148
	PMW	3	8.33 (4.16)	5	13	21	11.33 (5.72)	3	23
	Gram. judg.	2	2.22 (.93)	1.56	2.88	21	2.55 (.42)	1.49	3
	TEGI	2	34.5 (26.16)	16	53	21	70.86 (20.18)	34	100
	TROG-2	2	3 (1.41)	2	4	21	6.71 (3.86)	0	16

Note. [a]Mean length of utterance. [b] Number of grammatical errors per T-unit. [c] Number of different words. [d] Percent of maze words to total number of words. [e] Grammaticality judgment. [f] Test of Early Grammatical Impairment. [g] Test for Reception of Grammar, second edition.

Table 6

Correlations among Experimental Measures, Length of L2 Exposure, and L2 Proficiency for each Language Ability Group

		MLU	Gram. errors	NDW	PMW	Gram. judgment	TEGI	TROG-2	Length of L2 exposure	L2 proficiency
	MLU[a]	1								
	Gram. Errors[b]	-.48	1							
	NDW[c]	.54	-.59*	1						
PLI	PMW[d]	-.48	.36	-.19	1					
	Gram. judgment[e]	.22	-.40	.51	-.43	1				
	TEGI[f]	.52	-.81**	.44	-.49	.47	1			
	TROG-2[g]	.66	-.25	.59	.17	-.02	.41	1		
	Length of L2 exposure	.08	-.11	-.001	-.13	.67	.40	-.04	1	
	L2 proficiency	.66*	-.82**	.48	-.50	.40	.83*	.31	.00	1
	MLU	1								
	Gram. errors	-.38**	1							
	NDW	.74**	-.50**	1						
TD	PMW	-.12	.17	-.27*	1					
	Gram. judgment	.64**	-.63**	.66**	-.20	1				
	TEGI	.62**	-.74**	.68**	-.26*	.77**	1			
	TROG-2	.56**	-.47**	.59**	-.35*	.61**	.60**	1		
	Length of L2 exposure	.42**	-.39**	.53**	-.27*	.50**	.49**	.32*	1	
	L2 proficiency	.47**	-.55**	.57**	-.11	.62**	.64**	.47**	.41**	1

Note. * Correlation is significant at the .05 level (2-tailed). ** Correlation is significant at the .01 level (2-tailed). [a]Mean length of utterance. [b] Number of grammatical errors per T-unit. [c]Number of different words. [d] Percent of maze words to total number of words. [e] Grammaticality judgment. [f]Test of Early Grammatical Impairment. [g] Test for Reception of Grammar, second edition.

Table 7

Descriptive Statistics for Experimental Measures for each Language Ability Group as a Function of L2 Proficiency Level

Proficiency level	Measure	PLI				TD			
		N	M (*SD*)	Min	Max	N	M (*SD*)	Min	Max
Low	MLU[a]	2	4.19 (.93)	3.53	4.85	8	5.28 (1.29)	3.09	6.33
	Gram. errors[b]	2	1.18 (.06)	1.14	1.22	8	.88 (.36)	.49	1.33
	NDW[c]	2	52.50 (2.12)	51	54	8	49.75 (7.55)	33	57
	PMW[d]	2	29.50 (17.68)	17	42	8	16 (10.21)	1	34
	Gram. judgment[e]	2	1.09 (1.53)	0	2.17	8	1.42 (.20)	1.05	1.67
	TEGI[f]	2	3.00 (1.41)	2	4	8	25.75 (10.03)	11	43
	TROG-2[g]	2	2.00 (2.83)	0	4	8	2.38 (2.07)	0	5
Intermediate	MLU	9	6.24 (1.07)	3.97	7.27	31	6.62 (1.21)	4.35	8.66
	Gram. errors	9	.46 (.20)	.19	.78	31	.50 (.21)	.11	.94
	NDW	9	73.22 (14.98)	52	92	31	73.45 (19.6)	39	118
	PMW	9	9.67 (6.26)	3	21	31	15.52 (9.42)	3	40
	Gram. judgment	6	1.88 (.71)	1.08	2.88	31	2.02 (.58)	.83	3
	TEGI	6	41.83 (14.11)	16	53	31	46.68 (22.37)	10	100
	TROG-2	6	3.17 (1.47)	1	5	31	4.77 (3.08)	0	12
High	MLU	1	7.27*			21	7.31 (1.1)	5.93	10.52
	Gram. errors	1	.21*			21	.32 (.28)	.03	1.07
	NDW	1	76*			21	91.86 (23.22)	62	148
	PMW	1	15*			21	13.43 (7.79)	3.00	35
	Gram. judgment	0	-			22	2.59 (.44)	1.64	3
	TEGI	0	-			22	75.68 (21.64)	26	100
	TROG-2	0	-			22	7.36 (3.63)	2	16

Note. [*] This is a raw score because there was only one child with PLI in the high proficiency group.
[a]Mean length of utterance. [b] Number of grammatical errors per T-unit. [c] Number of different words. [d] Percent of maze words to total number of words. [e] Grammaticality judgment. [f] Test of Early Grammatical Impairment. [g] Test for Reception of Grammar, second edition.

Table 8
Coefficients from Discriminant Analysis for Entire Sample of Children with and without PLI

Predictors	Structure coefficients	Standardized discriminant function coefficients
	Function 1	Function 1
MLU[a]	1.00	1.00
NDW*[b]	.74	-
TEGI*[c]	.61	-
Grammaticality judgment*	.57	-
TROG-2*[d]	.56	-
Grammatical errors per T-unit*	-.40	-
PMW*[e]	-.16	-

Note. * This variable not used in the analysis
[a] Mean length of utterance.
[b] Number of different words.
[c] Test of Early Grammatical Impairment.
[d] Test for Reception of Grammar, second edition.
[e] Percent of maze words to total number of words.

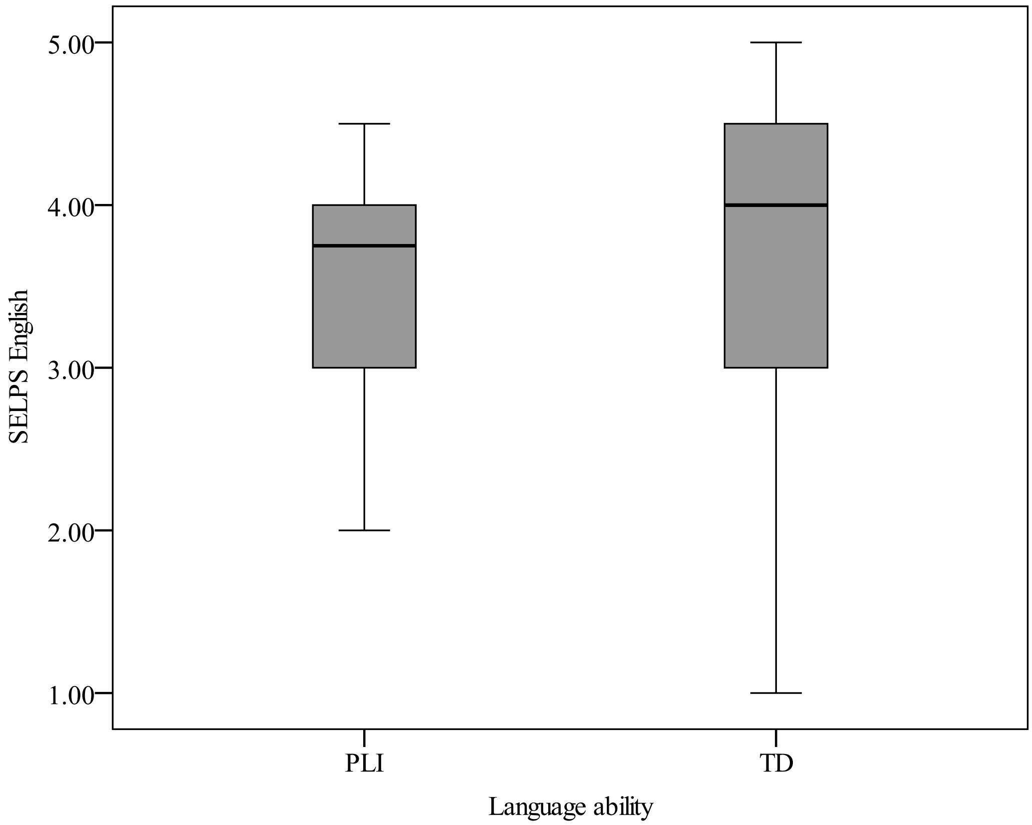

Figure 1. The SELPS English scores for children with PLI and TD.

Note. The boxes represent the middle 50% of the data. The remaining 50% is contained between the box and the whiskers. The bottom of the box represents the 25th percentile and top of the box represents the75th percentile. The band near the middle of the box represents the 50th percentile (the median).

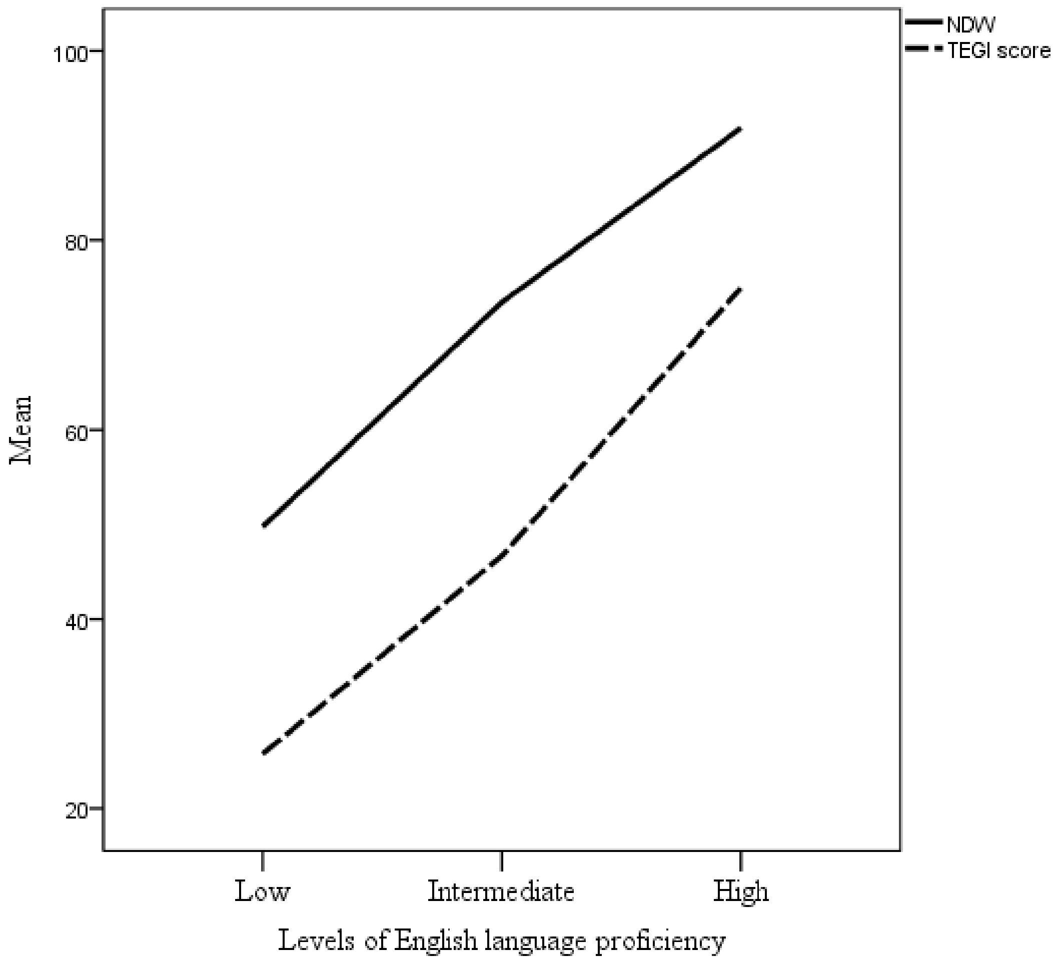

Figure 2. Means for NDW and TEGI in children with TD across L2 proficiency levels.

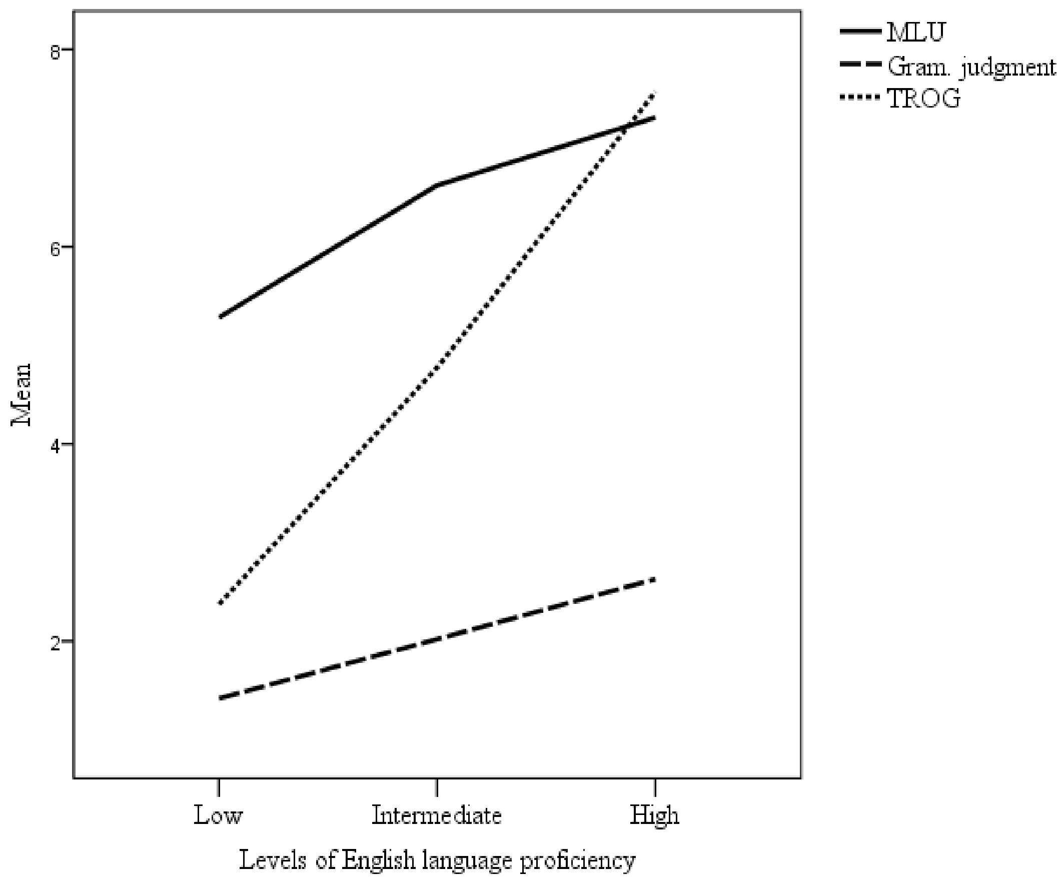

Figure 3. Means for MLU, grammaticality judgment, and TROG-2 in children with TD across L2 proficiency levels.

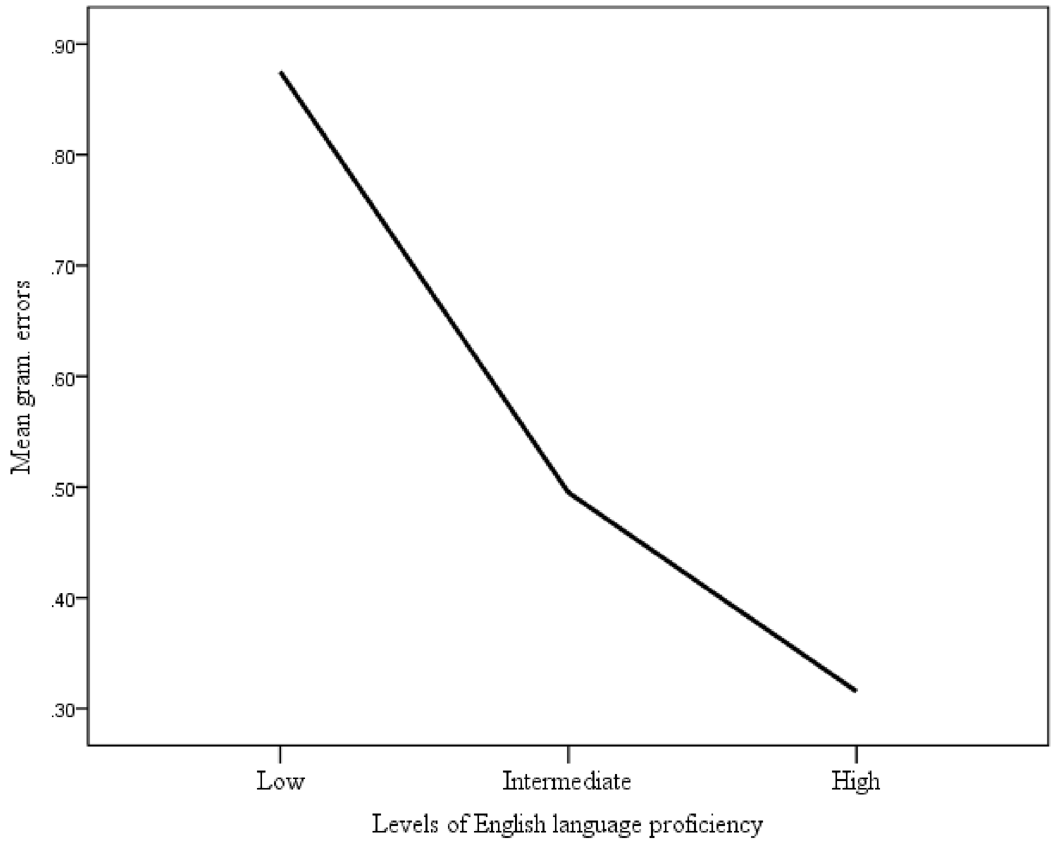

Figure 4. Mean number of grammatical errors in children with TD across L2 proficiency levels.

APPENDIX A

HUMAN SUBJECT DOCUMENTATION

ASU Knowledge Enterprise Development

Office of Research Integrity and Assurance

To:	Maria Restrepo COOR
From:	Mark Roosa, Chair Soc Beh IRB
Date:	10/21/2011
Committee Action:	**Expedited Approval**
Approval Date:	10/21/2011
Review Type:	Expedited F7
IRB Protocol #:	1110006983
Study Title:	Second Language Proficiency in Sequential Bilingual Children with and without Primary Language Impairment
Expiration Date:	10/20/2012

The above-referenced protocol was approved following expedited review by the Institutional Review Board.

It is the Principal Investigator's responsibility to obtain review and continued approval before the expiration date. You may not continue any research activity beyond the expiration date without approval by the Institutional Review Board.

Adverse Reactions: If any untoward incidents or severe reactions should develop as a result of this study, you are required to notify the Soc Beh IRB immediately. If necessary a member of the IRB will be assigned to look into the matter. If the problem is serious, approval may be withdrawn pending IRB review.

Amendments: If you wish to change any aspect of this study, such as the procedures, the consent forms, or the investigators, please communicate your requested changes to the Soc Beh IRB. The new procedure is not to be initiated until the IRB approval has been given.

Please retain a copy of this letter with your approved protocol.

APPENDIX B

PARENT QUESTIONNAIRE

4665

SSLIC Habilidades y Prácticas de Lenguaje en el Hogar

Por favor responda todas las páginas lo mejor que pueda. Todos los datos serán confidenciales y la información no será compartida con ninguna otra institución. Si tiene preguntas sobre el cuestionario por favor contacte a Giovanna Moreno al 480-727-8796 o a Laida Restrepo al 480-727-8596

Por favor no escriba en este cuadro. El número de identificación se mantiene confidencial para cuidar la identidad de su hijo/a.	→	ID

III. Información Social

1. Programa de almuerzo: (Rellene el círculo) ○ Gratis ○ Rebajado ○ Regular

2. ¿Cuál es el nivel de educación de la madre y el padre del niño?

Madre	Nivel de Educación: ○ Primaria ○ High School/Prepa ○ College/Universidad	Padre	Nivel de educación: ○ Primaria ○ High School/Prepa ○ College/Universidad

3. En qué idioma le hablan al niño las siguiente personas en la casa:

Madre	Padre	Hermanos	Otros (Tíos, abuelos, primos, etc.)
○ Solo español	○ Solo español	○ Solo español	○ Solo español
○ Solo inglés	○ Solo Inglés	○ Solo Inglés	○ Solo inglés
○ Inglés y español	○ Inglés y español	○ Inglés y español	○ Inglés y español

4. En qué idioma responde el niño cuando habla con las siguientes personas en la casa:

Madre	Padre	Hermanos	Otros (Tíos, abuelos, primos, etc.)
○ Solo español	○ Solo español	○ Solo español	○ Solo español
○ Solo inglés	○ Solo inglés	○ Solo inglés	○ Solo inglés
○ Inglés y español	○ Inglés y español	○ Inglés y español	○ Inglés y español

IV. Información Escolar

5. Por favor complete el cuadro a continuacion con la informacion de las escuelas a las que su hijo ha asistido:

Edad	Marque si el niño fue a la escuela o se quedó en casa en cada edad.	Marque el idioma principal que se utilizó con el niño durante esa edad en la escuela/guardería * o casa (si no fue a la guardería)
0-1 años	O Escuela O Guarderia/Preescolar O Casa/familia	O Solo español O Solo inglés O Inglés y español
1-2 años	O Escuela O Guarderia/Preescolar O Casa/familia	O Solo español O Solo inglés O Inglés y español
2-3 años	O Escuela O Guarderia/Preescolar O Casa/familia	O Solo español O Solo inglés O Inglés y español
3-4 años	O Escuela O Guarderia/Preescolar O Casa/familia	O Solo español O Solo inglés O Inglés y español
4-5 años	O Escuela O Guarderia/Preescolar O Casa/familia	O Solo español O Solo inglés O Inglés y español
5-6 años	O Escuela O Guarderia/Preescolar O Casa/familia	O Solo español O Solo inglés O Inglés y español
6-7 años	O Escuela O Guarderia/Preescolar O Casa/familia	O Solo español O Solo inglés O Inglés y español

6. ¿Alguna vez su hijo ha repetido grados o años en la escuela? ○ Si ○ No

7. ¿Ha asistido su hijo/a a educación especial? ○ Si ○ No
(Si su respuesta es sí, por favor complete las preguntas a continuación: 7A, 7B)

7A. ¿En qué idioma ha recibido educación especial? ○ Español ○ Inglés ○ Inglés y Español

7B. ¿Qué tipo de servicio ha recibido? (Marque todos los que apliquen)

- ○ Terapia de Lenguaje
- ○ Terapia Ocupacional
- ○ Terapia Física
- ○ Intervención Temprana
- ○ Terapia Auditiva/Audífonos

V. Opinión de la familia acerca del lenguaje del niño/a

8. ¿Le preocupa la manera en que su niño(a) habla? ○ Si ○ No
9. ¿Tienen otras personas dificultad para entender lo que dice el/la niño(a)? ○ Si ○ No
10. ¿Habla su niño(a) tan bien como otros niños (as) de su misma edad? ○ Si ○ No
11. ¿Habla su niño(a) "chistoso" o "raro" comparado a otros niños(as)? ○ Si ○ No
12. ¿Algún familiar/maestro le ha comentado que su niño(a) habla poco o habla mal? ○ Si ○ No
13. ¿Dice su niño(a) frases correctas casi todo el tiempo? ○ Si ○ No
14. ¿Para su edad, tiene su niño(a) dificultad explicando o describiendo cosas? ○ Si ○ No
15. ¿Para su edad su niño(a) sabe muy pocas palabras? ○ Si ○ No
16. ¿En comparación con otros niños(as), tiene su niño(a) dificultades de aprendizaje? ○ Si ○ No
17. ¿En comparación con otros niños(as) de su misma edad, tiene su niño(a) dificultad poniendo atención por mucho tiempo? ○ Si ○ No
18. ¿Para su edad, tiene su niño(a) dificultad pronunciando palabras? ○ Si ○ No
19. ¿Su niño(a) entiende la mayoría de lo que usted dice? ○ Si ○ No
20. ¿Usted tiene que repetir lo que le dice a su niño(a) mas que a otros niños? ○ Si ○ No
21. ¿Comparado con otros niños, su niño(a) tiene dificultad entendiendo preguntas? ○ Si ○ No

VI. Historia de la Comunicación Familiar

22. Para nosotros, es muy importante saber qué lenguaje escucha y habla su hijo/a fuera de la escuela. Por favor complete la tabla siguiente con el lenguaje en que su hijo/a escucha y habla durante las diferentes horas del día, la semana, y los fines de semana.

Idioma que escucha y habla su hijo en la semana y los fines de semana
(Rellene el círculo para cada tiempo en el día)

Durante la última semana qué idioma **ESCUCHÓ** el niño en la casa en los siguientes horarios:

	Lunes	Martes	Miércoles	Jueves	Viernes
6-9am (antes de la escuela)	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español
2-5pm (después de la escuela)	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español
5-8 pm	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español
8-10 pm	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español

	Sábado	Domingo
En la mañana	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español
En la tarde	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español
En la noche	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español

Idioma que escucha y habla su hijo en la semana y los fines de semana
(Rellene el círculo con su respuesta para cada tiempo en el día)

Durante la última semana qué idioma **HABLÓ** el niño en la casa en los siguientes horarios:

	Lunes	Martes	Miércoles	Jueves	Viernes
6-9am (antes de la escuela)	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español
2-5pm (después de la escuela)	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español
5-8 pm	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español
8-10 pm	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español

	Sábado	Domingo
En la mañana	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español
En la tarde	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español
En la noche	O Solo español O Solo inglés O Inglés y español	O Solo español O Solo inglés O Inglés y español

4665

VII. Característica del dialecto

23. Frecuentemente la gente habla varios tipos de español. Marque el que usted habla en su casa:

- O Español Mexicano
- O Español Guatemalteco
- O Español Cubano
- O Español Puerto Riqueño
- O Otro(especifique por favor):

24. Aquí hay una lista de actividades.Por favor indique el idioma principal que usa durante estas actividades, la frecuencia, y el tiempo que dura si hace estas actividades.

Actividad	Veces por semana	Horas por día	Lenguaje (por favor seleccione solamente uno)
Leer libros, revistas, cuentos, Biblia, etc.	O 1 vez O 2-3 veces O más de 3 veces	O menos de 1 hora O 1-3 horas O más de 3 horas	O Solo español O Solo inglés O Inglés y español
Ver la televisión	O 1 vez O 2-3 veces O más de 3 veces	O menos de 1 hora O 1-3 horas O más de 3 horas	O Solo español O Solo inglés O Inglés y español
Jugar con sus compañeros/hermanos(as)	O 1 vez O 2-3 veces O más de 3 veces	O menos de 1 hora O 1-3 horas O más de 3 horas	O Solo español O Solo inglés O Inglés y español
Jugar en la computadora	O 1 vez O 2-3 veces O más de 3 veces	O menos de 1 hora O 1-3 horas O más de 3 horas	O Solo español O Solo inglés O Inglés y español

APPENDIX C

TEACHER QUESTIONNAIRE

SSLIC Teacher Questionnaire (11-09-08)

Spanish Screener for Language Impairment in Spanish-Speaking Children

Identifying information will be removed and replaced with a code

ID

Instructions: Please print in capital letters using black or blue ink only and avoid contact with the edge of the box.

Date (MM/DD/YY) ___ / ___ / ___

Shade Circles Like This--> ●
Not Like This--> ⊗ ✓

School

Teacher's last name

Grade

P - Preschool
K - Kinder
1 - 1st grade
2 - 2nd grade
3 - 3rd grade

Please rate this child's use of Spanish:

Speaks to you in Spanish	◯ Never	◯ Ocassionally	◯ Frequently	◯ Always
Speaks to other adults in the classroom in Spanish	◯ Never	◯ Ocassionally	◯ Frequently	◯ Always
Speaks with classmates in the classroom in Spanish	◯ Never	◯ Ocassionally	◯ Frequently	◯ Always
Speaks with classmates outside the classroom in Spanish	◯ Never	◯ Ocassionally	◯ Frequently	◯ Always

Please rate this child's use of English:

Speaks to you in English	◯ Never	◯ Ocassionally	◯ Frequently	◯ Always
Speaks to other adults in the classroom in English	◯ Never	◯ Ocassionally	◯ Frequently	◯ Always
Speaks with classmates in the classroom in English	◯ Never	◯ Ocassionally	◯ Frequently	◯ Always
Speaks with classmates outside the classroom in English	◯ Never	◯ Ocassionally	◯ Frequently	◯ Always

(Please shade only one circle)
How would you rate this child's understanding of English?

◯ Doesn't understand anything

◯ Understands a little

◯ Undesrstands only the main ideas

◯ Understands most of what is said

◯ Understands as well as a native speaker of English

(Please shade only one circle)
How would you rate this child's ability to speak English?

◯ Cannot speak any English

◯ Speaks a little English

◯ Speaks limited English with errors

◯ Speaks fluent English with errors

◯ Speaks like a native speaker of English

Please rate your concern about this child's development or skills:

Hearing	◯ Not concerned	◯ Somewhat concerned	◯ Very concerned
Speech Development	◯ Not concerned	◯ Somewhat concerned	◯ Very concerned
Oral language development	◯ Not concerned	◯ Somewhat concerned	◯ Very concerned
Literacy development	◯ Not concerned	◯ Somewhat concerned	◯ Very concerned
Motor skills	◯ Not concerned	◯ Somewhat concerned	◯ Very concerned
Thinking skills	◯ Not concerned	◯ Somewhat concerned	◯ Very concerned
Social skills	◯ Not concerned	◯ Somewhat concerned	◯ Very concerned

To your knowledge, has this child ever received any special education services?	◯ YES	◯ NO
To your knowledge, has this child ever participatedin speech or language therapy?	◯ YES	◯ NO
To your knowledge, has this child ever repeated a grade?	◯ YES	◯ NO

36769

APPENDIX D

STORY SCRIPT AND INSTRUCTIONS

Script: (1) A boy was getting dressed in his bedroom. His pet dog, frog and turtle watched as he put on his best clothes. (2) While the boy was petting the dog, the frog jumped into his coat pocket. The boy didn't know he was there. (3) As the boy left with his family, he waved and said "Goodbye" to his pets. The frog waved goodbye too. (4-5) When the boy and his family arrived at a fancy restaurant, the doorman helped them out of the car. The frog peaked out of the boy's pocket but no one noticed him. (6-7) The boy and his family sat down at a table in the restaurant. While they were looking at the menus, the frog jumped out of the boy's pocket towards the band. (8) The frog landed right in the man's saxophone! "Squeak" went the saxophone. (9) The man looked inside the saxophone to see why it made that awful noise. (10) Then the frog fell out of the horn and landed right on the saxophone player's face! (11)The saxophone player was so surprised that he fell backwards into the drum. (12-13) The drummer yelled at the saxophone player, "Look what you did to my drum- it's broken!" While they were arguing, the frog jumped away on a plate of lettuce salad. (14) The waiter didn't notice the frog. He served the salad to a woman. (15) Just as she was about to take a bite, the frog popped out of the lettuce. The woman was shocked to see the frog. (16) She screamed and fell back on her chair. The frog was frightened and he jumped away. (17) There was a man at the next table who was having a glass of wine with his wife. The frog landed right in his glass. (18) The woman complained to the waiter about getting a salad with a frog in it. She was very angry! (19) Meanwhile, when the man went to take a sip of his drink, the frog kissed him right on the nose. (20-21) The angry waiter was about to grab the frog who was waving goodbye to the man and his wife. (22-23) The waiter, who had caught the frog, was going to throw him out of the restaurant. But the boy saw the waiter carrying his frog and shouted, "Hey, that's my frog!" The boy's mother told him to be quiet. (24) The boy asked the waiter to give him back his frog. (25) The angry waiter told the boy and his family, "Take your frog and get out of this restaurant at once. Don't you ever bring that frog in here again!" (26-27) On the way home the boy's family was angry with him. The frog had ruined their dinner! (28-29) when they got home the boy's father scolded him, "You go to your room and stay there!" The dog and the turtle peaked around the corner to see what was going on. (30) When they got in his room, the boy and the frog laughed about everything that had happened at the restaurant. The more they thought about it, the more they laughed.

Instructions: Read the story to the child turning pages as you read so the child can look at the pictures. The page numbers correspond to the numbers in the story the script. When you are done reading the story, say, "*Now it is your turn to tell me the story. You may look at the pictures if you want to. Ready*?"

APPENDIX E

FIDELITY CHECKLIST

Examiner: _________________ Observer: ____________________
Date: _____________________ Child's ID: ___________________

Test of Early Grammatical Impairment

	Yes	No
1. Gives an introduction as instructed	1	0
2. Show the practice page and present practice item	1	0
3. Record responses properly	1	0
4. Use prompts when needed	1	0
5. Gives appropriate guidance when needed	1	0

Total number of points:
Comments:

Test for Reception of Grammar-2

	Yes	No
1. Gives an introduction as instructed	1	0
2. Show the practice page and present practice item	1	0
3. Record responses properly	1	0
4. Use prompts when needed	1	0
5. Gives appropriate guidance when needed	1	0

Total number of points:
Comments:

Story Retelling Task

	Yes	No
1. Gives an introduction as instructed	1	0
2. Record responses properly	1	0
3. Use prompts when needed	1	0
4. Gives appropriate guidance when needed	1	0

Total number of points:
Comments: