# THE EFFECT OF A COMPREHENSIVE ENGLISH LANGUAGE/LITERACY INTERVENTION IN BILINGUAL CLASSROOMS ON THE DEVELOPMENT OF ENGLISH READING FLUENCY FOR ENGLISH-LANGUAGE LEARNERS, GRADES 2-3

A Dissertation

by

ELIZABETH PAULINE TREVINO

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

| | |
|---|---|
| Chair of Committee, | Rafael Lara-Alecio |
| Committee Members, | Beverly J. Irby |
| | Sharolyn Pollard-Durodola |
| | Richard Curry |
| Head of Department, | Victor Willson |

December 2012

Major Subject: Educational Psychology

## ABSTRACT

English-language learners (ELLs) demonstrate lower levels of English reading proficiency than do native English-speaking students. Oral reading fluency (ORF), the number of words read correctly in 1 min, is one indicator of reading proficiency. Within second language (L2) reading research, there have been few studies of L2 ORF development. The purposes of this study were to: (a) model the trajectory (i.e., initial status and growth) of English ORF in Grades 2 and 3 for Spanish-speaking ELLs in bilingual education programs, and (b) determine the effect of a 4-year structured intervention in English language and reading on L2 ORF development.

Data were archived from Project ELLA, a longitudinal, randomized study documenting ELLs' acquisition of English language and reading from kindergarten through third grade. Data included 1,470 observations of English ORF from 283 ELLs at 17 schools. Schools were randomly assigned to the intervention (*n*=8) or control (*n*=9) condition. In intervention schools, a one-way dual language program and a comprehensive ESL intervention were implemented.  The intervention emphasized L2 oral language development in kindergarten and first grades, basic L2 reading skills in second grade, and content-area reading skills in third grade. In the control schools, the district's typical transitional bilingual education program and ESL curricula were implemented. L2 ORF was measured using DIBELS ORF on six occasions. Piecewise multilevel growth models were used for data analysis.

In Grades 2 and 3, ELLs followed a two-stage linear growth trajectory in English ORF, with a large decrease in level between grades. Slope parameters were positive in both grades but decreased slightly in third grade. Participating in Project ELLA added 1.52 wcpm per month to students' ORF scores in Grade 2. Both intervention and control groups improved at the same rate in Grade 3; however, intervention students maintained the higher level of ORF that was attained during second grade. Therefore, the ELLA intervention accelerated L2 ORF growth in second grade, such that intervention students read with greater fluency compared to control students throughout second and third grades.

## DEDICATION

To Saul and my parents, for unfailing love and eagle's wings

*But they that wait upon the* LORD *shall renew their strength;*

*they shall mount up with wings as eagles;*

*they shall run, and not be weary;*

*and they shall walk, and not faint. -- Isaiah 40:31*

## ACKNOWLEDGEMENTS

This dissertation would not have been completed without the assistance of many others. First, I would like to thank Dr. Rafael Lara-Alecio, who served as the chair of my committee but who is much more than chair. Dr. Lara, thank you for encouraging me to embark on a doctoral degree and for mentoring me throughout graduate school. You assured me that this project would come to fruition, despite my own deep-seated doubts and despite the life challenges that I experienced during its writing. Thank you for bringing me to the Bilingual Education Program family, believing in me, and guiding me to this point in my academic career. I have reached this destination only by standing on your shoulders, and I am truly grateful for all the opportunities you have given me.

I would also like to thank the other members of the committee, Drs. Beverly Irby, Sharolyn Pollard-Durodola, and Richard Curry. Dr. Irby, your keen insights and comments to early drafts greatly improved this manuscript. Dr. Durodola and Dr. Curry, you both provided strong support throughout the lengthy dissertation process. I appreciate all of you for serving on the committee and seeing me through completion. Additionally, I would like to extend special thanks to Dr. Fuhui Tong for reviewing the methodology and data analysis chapters.

Data used in this study were obtained from Project ELLA (R305P030032) funded by the Institute of Education Sciences, U.S. Department of Education. I am very grateful to Drs. Lara and Irby, Principal Investigators, for granting access to these data. I also acknowledge the superb work of the Project ELLA research team: researchers, school district administrators, project coordinators and staff, teachers, paraprofessionals, students and their parents. Because of their unshakeable commitment to English-language learners and outstanding effort in implementing the project, the entire ELLA study has made a strong contribution to bilingual/ESL education research.

I would not have had the endurance to complete this dissertation without support from my dissertation/thesis writing group at the Texas A&M University Writing Center. I received invaluable feedback and social support from the group and its facilitators, Candace and Matt. Every graduate student should have the opportunity to participate in such an amazing writing group. I must also acknowledge the daily support that I received from the Graduate Student Writing Club and Coach Debra at AcademicLadder.com. Coach Debra and the members of this group followed my daily and weekly progress and cheered me on every step of the way.

Finally, but certainly not least, I honor my family and all of my friends in the Bilingual Education Program for being rivers of encouragement through the peaks and valleys of dissertation writing. Most especially, I thank my husband Saul and my parents, both my parents by birth (Mom and Dad) and my parents by marriage (Jerry Mac, René, and Marie). Their unwavering belief and steadfast encouragement gave me hope, endurance, and determination. I also thank Janie, who was a most ardent cheerleader and whose kind listening ears were always available when I struggled, and Dr. Miranda Walichowski, who was a ready source of personal and professional advice that propelled me through inertia. For me, these friends and family have been beacons of heavenly love, and for that, I am richly blessed.

## NOMENCLATURE

| | |
|---|---|
| AIC | Akaike Information Criterion |
| AOL | Academic Oral Language |
| AOLS | Academic Oral Language in Science |
| AOWLS | Academic Oral and Written Language in Science |
| BIC | Bayesian Information Criterion |
| BOY | beginning of the year |
| CBM | curriculum-based measurement |
| CRISELLA | Content Reading Integrating Science for English Language and Literacy Acquisition |
| DIBELS | Dynamic Indicators of Basic Early Literacy Skills |
| DL | dual language programs |
| EIR | Early Interventions in Reading |
| ELLs | English-language learners |
| EOY | end of the year |
| ESL | English as a second language |
| FML | full maximum likelihood |
| ICC | intraclass correlation coefficient |
| IRB | Institutional Review Board |
| L1 | first language |
| L2 | second language |

| | |
|---|---|
| MLM | multilevel linear modeling |
| MOY | middle of the year |
| NES | native English speakers, native English-speaking |
| NLP | National Literacy Panel on Language-Minority Children and Youth |
| OLS | ordinary least squares |
| ORF | oral reading fluency |
| Project ELLA | Project English Language/Literacy Acquisition |
| RtI | Response to Intervention |
| SEI | structured-English immersion |
| STELLA | Story Retelling and Higher Order Thinking for English Language and Literacy Acquisition |
| TBE | transitional bilingual education |
| TBOP | Transitional Bilingual Observation Protocol |
| wcpm | words correct per minute |

# TABLE OF CONTENTS

Page

Page

## LIST OF FIGURES

Page

## LIST OF TABLES

Page

# CHAPTER I

# INTRODUCTION

## Background of the Study

Numbering more than 4.5 million, English-language learners (ELLs) constituted 9.1% of the students attending U.S. public schools in 2009 (*Quality Counts*, 2009). In Texas, the setting for the present study, 17% of the public school students were ELLs in 2010 (Texas Education Agency, 2010). The number of ELLs is growing rapidly across all regions of the nation as a result of population shifts, particularly in states that historically have educated few ELLs (Capps et al., 2005; Fry, 2007; *Quality Counts*, 2009). The ELL population has increased faster than the non-ELL population. From 2000 to 2006, ELL enrollment in U.S. public schools increased by 18.1%, whereas non-ELL student enrollment increased by only 7.3% (*Quality Counts*, 2009). Given that the population of ELLs is projected to continue growing (Fry, 2007), educating ELLs in U.S. public schools is an issue of national concern, particularly in the area of reading.

ELLs demonstrated much lower levels of English reading proficiency than non-ELLs on the 2009 National Assessment of Educational Progress (*Nation's Report Card*, 2009). Only 29% of fourth-grade ELLs performed at or above the Basic proficiency level, compared to 70% of non-ELLs. ELLs averaged a scale score of 188, 20 points *below* the Basic proficiency level cut score and 35 points below the average non-ELL score (223). Results for eighth-grade students were similar, though the gap between ELLs and non-ELLs was wider: 26% of ELLs performed at or above the Basic proficiency level, compared to 78% of non-ELLs. Eighth-grade ELLs averaged a scale score of 219, 24 points *below* the Basic proficiency level cut score and 47 points below the average non-ELL score (266). From these results, it is markedly apparent that not only do ELLs lag behind their non-ELL peers, but they experience difficulty developing basic English reading proficiency. This is particularly unsettling given the importance of reading skill to academic achievement and successful completion of high school (Reschly, 2010).

Several significant scholastic problems are related to low reading proficiency. Reading difficulties frequently result in grade retention and/or referral for special education services (Bowman-Perrot, Herrera, & Murry, 2010; Reschly, 2010). For English-language learners, there is particular concern about disproportionate representation in special education as well as the initiation of inappropriate referrals for special education evaluation (i.e., mistaking natural language acquisition processes for learning disabilities; Orosco, Almanza de Schonewise, de Onis, Klingner, & Hoover, 2008). Difficulties with reading may also affect students' engagement and motivation, which has the long-term effect of elevating students' risk for dropping out of school (Reschly, 2010). In 2007, the status dropout rate for foreign-born Hispanics ages 16 through 24, many of whom are ELLs, was 34%, nearly 2.5 times higher than the rate for native-born Hispanics (11%; Planty et al., 2009). Indeed, status dropout rates may reflect ELLs' limited proficiency in reading English. In order to prevent the sobering consequences of reading difficulties for ELLs, it is essential to understand the development of second language (L2) reading proficiency and how to provide effective L2 reading instruction to ELLs.

**L2 Reading**

Researchers have suggested that the cognitive processes used in L2 reading are similar to those used in first language (L1) reading (e.g., August & Shanahan, 2006; Chiappe & Siegel, 2006; Francis, Rivera, Lesaux, Kieffer, & Rivera, 2006). Despite the similarities in process, ELLs experience difficulties acquiring second language reading skills. As Francis, Rivera, Lesaux, Kieffer, and Rivera (2006) described:

> developing literacy in a second language is not a trivial task…While simultaneously developing conversational ability and basic reading skills, these learners must quickly begin to develop oral and written academic language skills for the development of academic knowledge and success in content-area classrooms. (p. 7)

Thus, ELLs face the tremendous challenge of learning to read in English while they are learning the language itself. Without a strong foundation of phonemic awareness, syntactic knowledge, or lexical development in English, decoding and constructing meaning from English text is difficult. Along with learning language and reading skills concurrently, ELLs' reading proficiency must be developed such that they are able to profit from content-area instruction in English. Such well-developed reading proficiency is critical to ELLs' school success (Bowman-Perrott et al., 2010; Francis et al., 2006).

Reading skills can be divided into two classes: word-level (or decoding) skills and text-level skills, which include reading connected text fluently and reading comprehension (August & Shanahan, 2006). According to reading theory and research (Chall, 1996; Eldredge, 2005; Perfetti, 1985), these skills are cumulative: "That is, without mastery of decoding, fluency is compromised; if decoding and fluency are not automatic, the reader's ability to extract and construct meaning from text effectively and efficiently is compromised" (August & Shanahan, 2006, p. 57).

In a seminal synthesis of literature related to second language literacy, the National Literacy Panel on Language-Minority Children and Youth (NLP) reported that ELLs can and do develop word-level reading skills at levels equal to their native-speaking peers (August & Shanahan, 2006). In contrast, the Panel found that language-minority students' reading comprehension skills do not approach the level of their native-speaking peers' comprehension. The general findings indicated that reading comprehension was a particularly difficult skill for language-minority students and that its development depended on effective instruction. The Panel described a "dearth of research on the development of text-level skills…" (p. 120) and called for "more research…to provide a better understanding of the development of text-level skills in language-minority students" (p.100).

**Oral Reading Fluency**

One text-level skill is reading fluency. Reading connected texts quickly, accurately, and with expression is essential for comprehension (Perfetti, 1985; National Institute of Child Health and Human Development [NICHD], 2000). Scholars of reading development concur that the ability to read connected text fluently is an important development in one's overall reading proficiency (Adams, 1990; Hasbrouck & Tindal, 2006; Kuhn & Stahl, 2003). Oral reading fluency (ORF) is "the oral translation of text with speed and accuracy" and is measured as the number of words in connected text read correctly in 1 min (Fuchs, Fuchs, Hosp, & Jenkins, 2001, p. 239).

Fuchs, Fuchs, Hosp, and Jenkins (2001) argued that oral reading fluency is a measure of overall reading competence:

> Our proposition is that oral reading fluency represents a complicated, multifaceted performance that entails, for example, a reader's perceptual skill at automatically translating letters into coherent sound representations, unitizing those sound components into recognizable wholes and automatically accessing lexical representations, processing meaningful connections within and between sentences, relating text meaning to prior information, and making inferences to supply missing information…and because oral reading fluency reflects this complex orchestration, it can be used in an elegant and reliable way to characterize reading expertise. (pp. 239-240)

Indeed, oral reading fluency is positively correlated with scores on reading comprehension measures for both native English speakers (Baker et al., 2008; Chard, Pikulski, & McDonagh, 2006; Fuchs et al., 2001; Klauda & Guthrie, 2008; Marston, 1989) and English-language learners (Baker & Good, 1995; Muyskens, Betts, Lau, & Marston, 2009; Wiley & Deno, 2005). Given the relationship between oral reading fluency and reading comprehension, it follows that improving reading fluency is an important path to improving ELLs' overall reading proficiency.

## Statement of the Problem

An understanding of how L2 reading skills develop and which circumstances encourage reading skill development is indispensable for improving L2 literacy acquisition for ELLs. However, the literature regarding the development of L2 reading skills is very limited (August & Shanahan, 2006; Genesee, Lindholm-Leary, Saunders, & Christian, 2006). The literature is especially sparse with regard to ELLs' reading fluency (Al Otaiba et al., 2009; August & Shanahan, 2006; Crosson & Lesaux, 2010). In its report, the National Literacy Panel called for more research "to examine precursors to reading fluency and instructional practices that can enhance reading fluency in English-language learners across the school years" (August & Shanahan, 2006, p.67). The present study contributes to an understanding of L2 reading proficiency by investigating the developmental trajectory of reading fluency and the effect of instructional setting on ELLs' reading fluency development.

Oral reading fluency scores are used widely in research and in practice. During early reading instruction, ORF scores are used for monitoring progress and determining eligibility for supplemental reading instruction. ORF benchmarks and the typical developmental trajectory for ORF scores for native English speakers (NES) are well documented (Baker et al., 2008; Chard et al., 2008; Crowe, Connor, & Petscher, 2009; Deno, Fuchs, Marston, & Shin, 2001; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993; Hasbrouck & Tindal, 1992; Hasbrouck & Tindal, 2006; Kim, Petscher, Schatschneider, & Foorman, 2010; Speece & Ritchey, 2005; Stage & Jacobsen, 2001; Stage, Sheppard, Davidson, & Browning, 2001). In contrast, few researchers have investigated the trajectory and/or development of ORF with ELL samples (Al Otaiba et al., 2009; Baker, Park, & Baker, 2012; Betts, Bolt, Decker, Muyskens, & Marston, 2009; Dominguez de Ramirez & Shapiro, 2006; Yesil-Dagli, 2011). Only Al Otaiba et al. (2009) have documented fluency development in the same group of students across multiple academic years. Al Otaiba et al. noted that "reasonable growth expectations in oral reading fluency do not yet exist for the rapidly growing number of Latino students who attend high-poverty schools and who receive English-only reading instruction" (p. 324). Thus, without knowledge of typical L2 ORF development patterns, ELLs' L2 reading fluency scores are typically interpreted according to native-speaker norms in fluency growth, without regard for the effect of limited English proficiency on concurrent English reading acquisition (Linan-Thompson, Cirino, & Vaughn, 2007).

Furthermore, the L2 literacy field is in want of effective interventions for developing L2 text-level reading skills, particularly fluency and comprehension skills (Shanahan & Beck, 2006). Many recommendations about instruction of text-level skills for ELLs are derived from the research on monolingual populations rather than from research with second-language learners (Francis et al., 2006). Peer-assisted learning strategies (Calhoon, Al Otaiba, Cihak, King, & Avalos, 2007; McMaster, Kung, Han, & Cao, 2008; Saenz, Fuchs, & Fuchs, 2005) and small-group supplemental instruction (Gunn, Biglan, Smolkowski, & Ary, 2000; Gunn, Smolkowski, Biglan, & Black, 2002; Linan-Thompson, Vaughn, Hickman-Davis, & Kouzekanani, 2003; Vaughn, Cirino, et al., 2006; Vaughn, Mathes, et al., 2006) have shown promise for developing ELLs' fluency. However, these interventions focused on direct instruction of reading skills, and few addressed ELLs' oral language proficiency in tandem with reading skills (e.g., Tong, Irby, Lara-Alecio, & Mathes, 2008; Vaughn, Cirino, et al., 2006; Vaughn, Mathes, et al., 2006), despite strong research findings suggesting the importance of L2 oral language for L2 reading (Acosta, 2010; August & Shanahan, 2006; Droop & Verhoeven, 2003; Geva, 2006; Gottardo & Mueller, 2009; Lindsey, Manis, & Bailey, 2003; Nakamoto, Lindsey, & Manis, 2007, 2008; Proctor, Carlo, August, & Snow, 2005; Roberts & Neal, 2004; Saunders & O'Brien, 2006; Swanson, Rosston, Gerber, & Solari, 2008; Tong, 2006). According to Shanahan and Beck (2006) in the NLP report:

> …providing high-quality instruction in [reading skills] alone would be insufficient to support equal academic success for language-minority students. It may be that what is needed is sound reading instruction combined with simultaneous efforts to increase the scope and sophistication of [ELLs'] oral language proficiency. There is a need for research testing that hypothesis. (p. 448)

In order to improve reading outcomes for ELLs, it is imperative to conduct more research on instructional interventions that incorporate the teaching of reading skills in conjunction with explicit oral language development.

## Purpose of the Study

The purposes of this study were two-fold: (a) to model the trajectory (i.e., initial status and growth) of second language (L2) oral reading fluency for Spanish-speaking English-language learners from second through third grades and (b) to determine the effect of a 4-year structured intervention in English language and reading on the development of L2 oral reading fluency in ELLs. The present study modeled ORF development over second and third grades, two academic years during which reading fluency development is the most critical (Chall, 1996). In addition, the effectiveness of a comprehensive instructional intervention, which addressed both oral language and reading skills in L2, was evaluated.

Because these purposes entail studying skill development longitudinally, multilevel linear modeling (MLM) was used to model students' growth in English reading fluency. MLM, also known as hierarchical linear modeling, provides "an integrated approach for studying the structure and predictors of individual growth" (Raudenbush & Bryk, 2002, p. 161). To fulfill the two purposes of the study, a two-level model was hypothesized and tested. Level 1 consisted of within-individual repeated-measures of oral reading fluency from the beginning of second grade through the end of third grade. Level 2 consisted of the between-student variation in oral reading fluency and included the experimental condition (intervention vs. control) as a predictor of ORF initial status and growth.

## Significance of the Study

This study was conceptualized from the recommendations of the National Literacy Panel, which noted the paucity of studies focused on L2 text-level reading skills and called for more investigations showing the effects of instructional interventions on those skills (August & Shanahan, 2006). The Panel's report specifically mentioned the need for longitudinal and developmental studies of L2 reading and the need for more study of L2 oral reading fluency (August & Shanahan, 2006). The present study advances the field's understanding of second language reading by investigating the development of L2 oral reading fluency longitudinally and the effect of a comprehensive instructional intervention on L2 reading fluency.

This dissertation study contributes uniquely to the literature pertaining to ELLs' oral reading fluency by modeling fluency development in the same cohort of students from second through third grades. To my knowledge, only one published study, Al Otaiba et al. (2009), has depicted L2 fluency development in the same students across multiple academic years. Baker, Park, and Baker (2012) modeled ELLs' reading fluency in first through third grades but modeled the trajectory for each grade with a separate sample of students. Likewise, Dominguez de Ramirez and Shapiro (2006) reported rates of ORF gain for ELLs in first through fifth grades, but the design was cross-sectional rather than longitudinal. Additionally, the present study adds to the literature describing effective reading instruction for ELLs because the effects of an ESL intervention that combined explicit oral language instruction, context-embedded vocabulary development, direct instruction in reading, and content-area reading skill instruction are examined.

## Definition of Terms

### English-language Learner

An English-language learner (ELL) is a student whose first language is not English and who entered school without proficiency in English. ELLs are learning to speak, read and write English as an additional language in school. Synonyms used in the literature include limited-English proficient students, language-minority students, and English learners.

## Oral Reading Fluency

Oral Reading Fluency (ORF) is the number of words in a connected text that a student can read correctly in 1 min.

## Transitional Bilingual Education Program

Transitional bilingual education (TBE) is a program model for developing ELLs' English language and literacy skills. In TBE, students share a home language (in this study, Spanish), and both L1 and L2 are languages of instruction. Initially, L1 is the primary language of instruction for content and literacy, and daily English as a second language (ESL) instruction is incorporated. English content and literacy instruction is added gradually as students acquire L2 proficiency. Early-exit versions of TBE aim to transition students to English-only classrooms within 2 to 3 years. Late-exit versions of TBE aim to transition students to English-only instruction within 4 to 6 years.

## One-way Dual Language Program

One-way dual language (DL) programs are similar to TBE programs in that students share the same native language and are schooled bilingually. However, programmatic goals for one-way DL include developing both L1 and L2 and producing students who are bilingual and biliterate. Instructional characteristics include: separation of L1 and L2 during instruction; cognitively demanding lessons using grade-level core curriculum; and collaborative learning environments throughout the curriculum (Collier & Thomas, 2004).

## Multilevel Models

Multilevel linear modeling (MLM) was used as the method of data analysis in this study. Multilevel linear models, also called hierarchical linear models, mixed-effects models, random-coefficient regression models, and covariance components models, are statistical models used to analyze data with a nested, or clustered, structure (Beretvas, 2004; Raudenbush & Bryk, 2002). For this study, the nested data structure was as follows: repeated-measures nested in individuals, individuals nested in classroom and schools. In MLM, total variance is partitioned according to each level of clustering, which allows the researcher to investigate all levels simultaneously (Beretvas, 2004; Raudenbush & Bryk, 2002).

# Theoretical Framework

One purpose of this study was to model reading fluency development in English-language learners. Several theories of reading converge to support this investigation: LaBerge and Samuel's (1974) model of automatic reading and Chall's (1996) and Ehri's (1995) stage models of reading. In their model, LaBerge and Samuel explained how readers develop the ability to perform component skills of reading automatically and simultaneously. One implication of this model is that automatic word recognition, or fluent reading, frees up attentional resources for comprehension processes. The process of fluency development is also elucidated in stage models of reading. Chall (1996) located fluency within a continuum of developmental stages ranging from pre-literacy skills to mature reading. Lastly, Ehri (1995) delineated stages of sight-word recognition and showed how readers of alphabetic orthographies use the graphophonic system as a mnemonic for learning to identify words on sight. Ehri contended that automatic recognition of sight-words is necessary for reading fluently. Thus, in order to develop the efficient sight-word recognition that is necessary for fluent reading, readers must rely on phonological and lexical oral language skills. In this way, Ehri connected reading fluency to oral language development, which is especially important for young ELLs as they learn how to speak and read in L2 simultaneously.

## Research Questions

The specific research questions answered by this study are:

1. What are the average initial level and rate of growth in English oral reading fluency for Spanish-speaking ELLs in Grades 2 and 3?

2. What is the effect of Project ELLA on Spanish-speaking ELLs' initial level and growth in English oral reading fluency in Grades 2 and 3?

## Limitations

The generalizability of this study's results is limited in the following ways. First, this study occurred in an urban school district with a highly mobile student body primarily from low-income households. The participants were Spanish-speaking ELLs who had: (a) entered kindergarten with limited English proficiency and (b) received instruction in bilingual education programs. Therefore, results should be generalized only to those school districts and/or students who share these characteristics. Generalizability is also limited to Grades 2 and 3.

Furthermore, student mobility resulted in attrition over the span of the present study. The 4-year research project from which the present study's data were archived had an annual attrition rate of 13.1%, which is comparable to other longitudinal studies in urban schools (Tong, Irby, Lara-Alecio, Yoon, & Mathes, 2010). As a consequence, data could not be collected from the most mobile portion of the sample. This limitation is inherent in longitudinal research with panel designs (Gall, Borg, & Gall, 1996).

A third limitation is related to the holistic nature of the instructional intervention under study. The intervention was a comprehensive and multidimensional L2 intervention with multiple components. Because the components were implemented concurrently, each separate component's effect on reading fluency is unknown. Therefore, the findings presented in this dissertation indicate the effect of the whole intervention, with all curricular components implemented, on ELLs' English reading fluency.

## Delimitations

The major delimitation of this study is its singular focus on reading fluency. Although no other reading outcomes were included, I do not imply that reading is a simple process, nor that fluent reading should be the primary goal of reading instruction. Reading is indeed a complex process that requires multiple skills (Chall, 1996; Fuchs et al., 2001; Perfetti, 1985). The focus of this study was reading fluency because L2 reading fluency development for ELLs has not been found sufficiently in the research literature (Al Otaiba et al., 2009; August & Shanahan, 2006). Fluent reading is critical for developing reading comprehension, a skill that is problematic for ELLs (August & Shanahan, 2006).

## Organization of the Study

In Chapter I, I have provided the background information and outlined the major aspects of the study. In Chapter II, I explain the theoretical framework and review the relevant literature. In Chapter III, I describe the method used to conduct the study. I summarize the study’s results in Chapter IV. Finally, in Chapter V, I discuss the findings along with implications and recommendations emanating from the study.

# CHAPTER II

# REVIEW OF THE LITERATURE

In this longitudinal study, I investigated English-language learners' (ELLs) development of oral reading fluency and the effect of a comprehensive English language and reading intervention on ELLs' oral reading fluency. The purpose of Chapter II is to present the theoretical framework and to review the relevant literature. In the first half of the chapter, I describe the theoretical framework of the study. I begin by defining reading fluency. Next, I explain theories of reading fluency development, the contributions of oral language to fluency development, and the implications of reading fluency theories for ELLs reading in L2. In the second half of the chapter, I review the literature related to both of the study's purposes. I discuss growth patterns in reading fluency for both native English speakers and ELLs. Then, I discuss the effects of reading fluency interventions for ELLs.

## Theoretical Framework

### Reading Fluency

**Definition.** Reading has been characterized as the product of two processes, linguistic comprehension and decoding (Gough & Tunmer, 1986; Hoover & Gough, 1990), both being "necessary for reading success, neither being sufficient by itself" (Hoover & Gough, 1990, p. 128). Linguistic comprehension involves interpreting meaning from words, sentences, and discourse in spoken words or written text. Decoding is the ability to translate printed words into their phonological representations.

A reader accomplishes decoding through knowledge of letter-sound correspondences, analogy, and sight word recognition (Ehri, 1995; Hoover & Gough, 1990). Fluent reading, or reading connected text quickly and accurately, is the "bridge" between decoding and comprehension (Pikulski & Chard, 2005, p. 511). Automatic decoding frees a reader's attention to focus on comprehension; skilled comprehension enables a reader to utilize textual cues, which in turn makes decoding more efficient.

In its early conceptions, fluency was defined as reading connected text with speed and accuracy (Chard, Vaughn, & Tyler, 2002; Rasinski, 2006). During the past 2 decades, the concept of fluency has been expanded beyond reading rate and accuracy. Some re-conceptualizations emphasize comprehension processes whereas others emphasize decoding processes.

Among the first to add comprehension to the fluency construct, Schreiber (1991) and Dowhower (1991) argued for including prosody, or reading with expression. Later, the influential National Reading Panel report also included prosody by naming *expression* alongside *speed* and *accuracy* as components of fluency (NICHD, 2000). Prosodic reading mimics the tones and rhythms of natural speech by including the suprasegmental features (e.g., intonation, stress, and duration) that occur in speech (Dowhower, 1991). Reading with prosody indicates comprehension because the reader must organize text into meaningful units in order to apply appropriate inflections and expression.

However, for Pikulski (2006), reading with speed, accuracy, and expression is not a sufficient definition of fluency. Although the incorporation of prosody implicitly included comprehension, Pikulski argued for explicitly including comprehension in fluency definitions and proposed recasting fluency as the “process [of developing] efficient, effective decoding skills that permit a reader to comprehend text…Fluency is manifested in accurate, rapid, expressive oral reading and is applied during, and makes possible, silent-reading comprehension” (p.73). Thus, a fluent reader would be defined as one who not only reads quickly, accurately, and with expression but also comprehends the text. In this view, it is not sufficient to decode quickly and accurately unless the reader understands what is being read; therefore, true fluency must include comprehension.

Others have called for expanding fluency beyond the original speed and accuracy dimensions. Rather than emphasize comprehension, their expansions have emphasized decoding (Kame'enui & Simmons, 2001; Wolf & Katzir-Cohen, 2001). Wolf and Katzir-Cohen’s (2001) definition of fluency in connected text reading included sublexical skills, such as visual perception, letter identification, and phonemic awareness. Likewise, Kame’enui and Simmons (2001) asserted that fluency should be “extended to index the speed and accuracy not just of words but of the constituent phonologic and alphabetic elements that compose words” (p. 206). In other words, students’ speed and accuracy in segmenting phonemes, naming letter-sound correspondences, or other component tasks of decoding should also be considered as fluency alongside connected text fluency. This broader notion of fluency provides a path toward progress monitoring for emergent readers (Kame'enui & Simmons, 2001). Emergent readers are not skilled enough to read connected text fluently, but they can complete component decoding tasks (e.g., identifying letters, naming letter-sounds). Teachers can monitor students' speed and accuracy at these tasks and provide early reading intervention to at-risk students.

**Implications for measurement.** Defining reading fluency, whether the definition emphasizes comprehension or decoding, is important because construct definition has implications for educational measurement. Researchers must operationalize and reliably measure the constructs under study (Kline, 2009). Assessment instruments should measure a construct in a manner that accurately reflects the construct's definition. In the case of reading fluency, if expression or prosody is included in the definition, then expression should be measured in assessments of fluency. If text comprehension is part of fluency, then text comprehension should be included in fluency assessments. Likewise, if component decoding skills are fluency, then those skills should be part of fluency assessments.

However, incorporating comprehension and/or decoding expansions into fluency assessments is challenging. Prosody is difficult to assess efficiently and with strong inter-rater reliability (Allington, 1983; Torgesen & Hudson, 2006). Passage comprehension assessments are generally large-group, standardized tests, which are expensive and time-consuming to administer. Moreover, passage comprehension assessments do not measure reading speed and accuracy. Comprehension, speed, and accuracy should be assessed simultaneously for stronger construct validity (Samuels, 2006). Speed and accuracy in sublexical decoding skills can be assessed alongside connected text reading (e.g., DIBELS), but doing so is not without problems. These assessments are individually administered, and each fluency skill is measured with a separate assessment (e.g., initial sounds fluency, phoneme segmentation fluency). The additional fluency probes lengthen the one-on-one administration time. Furthermore, tests of decoding skill fluency have been criticized as tests of speed rather than fluency (Samuels, 2006). Given these measurement challenges, most fluency researchers have operationalized the fluency construct along the original speed and accuracy dimensions (Torgesen & Hudson, 2006). If comprehension and sublexical decoding dimensions are measured, they are most often measured separately.

**Oral reading fluency.** Despite scholarly disagreement on its theoretical definition, reading fluency's operational definition has been consistent: number of words in connected text read correctly in 1 min (words correct per minute, wcpm), called oral reading fluency (ORF). Students are asked to read a passage of connected text aloud for 1 min while an examiner tabulates the number of words read correctly. The time limit and the counting of only words read correctly make the measure one of reading speed and reading accuracy. Though this does not capture the multidimensionality of broader reading fluency, the rationale for using oral reading fluency as a measure of reading fluency is strong.

Researchers and practitioners use ORF as a measure of reading fluency for several reasons. First, ORF scores have proven robust under empirical scrutiny. They are psychometrically sound with strong findings of reliability, validity, and sensitivity to growth (Baker & Good, 1995; Betts, Muyskens, & Marston, 2006; Deno et al., 2001; Deno, Mirkin, & Chiang, 1982; Fuchs, Fuchs, & Maxwell, 1988; Marston, 1989; Muyskens et al., 2009). In addition, ORF scores are highly related to reading comprehension and overall reading proficiency (Deno et al., 1982; Fuchs et al., 2001; Fuchs et al., 1988; Shinn, Good, Knutson, Tilly, & Collins, 1992). Most recently, research findings have indicated that ORF scores predict outcomes on state reading assessments for both native English speakers (Baker et al., 2008; Chard et al., 2008; Stage & Jacobsen, 2001) and ELLs (Muyskens et al., 2009; Wiley & Deno, 2005). Furthermore, ORF scores are pragmatic assessments for research and practice. ORF assessments are inexpensive and can be frequently and easily administered by classroom teachers or teacher assistants. ORF's efficiency and practicality, along with its relationship to overall reading proficiency, make ORF assessments very useful for screening and progress monitoring.

## Theories of Reading Fluency

The theoretical underpinnings of reading fluency and reading development are pertinent to the present dissertation study. I now discuss the concept of automaticity, stage theories of reading development, and how readers achieve automaticity and fluency through sight-word recognition.

**Automaticity theory.** LaBerge and Samuels (1974) posited a model of fluency based on information processing learning theory. For beginning readers, word recognition is laborious, as the reader must utilize visual, phonological, and semantic processing and memory systems separately. To identify a word, the reader must first perceive and discriminate letters, then recode the letters into the appropriate phonemes, combine the phonemes into the correct word, and finally retrieve the meaning from semantic memory. According to LaBerge and Samuels, automatic processing, the ability to complete a task without explicitly attending to it, is “critical for the successful operation of multicomponent, complex skills such as reading” (p. 295). Fluent readers perform the component skills of reading (e.g., visual perception, phonological recoding, semantic retrieval, etc.) simultaneously; thus, they are able to read accurately and automatically (LaBerge & Samuels, 1974). Automatic processing enables readers to attend to the text as a whole:

> So long as word meanings are automatically processed, the focus of attention remains at the semantic level and does not need to be switched to the visual system for decoding, nor to the phonological level for retrieving the semantic meanings. (LaBerge & Samuels, 1974, p. 320)

When letter recognition, phonological recoding, and meaning retrieval are performed automatically, then the reader’s attention is free to focus on global comprehension of the text. With automaticity theory, LaBerge and Samuels accounted for how reading fluency relates to reading comprehension.

**Stage theories of reading development.** Automatic processing of text is the mechanism for reading fluently, but how do beginning readers achieve automaticity? The developmental path to automaticity is illuminated by stage theories of reading development. Chall (1996) and Ehri (1995, 1998, 1999) proposed stages of reading development. Both stage theories are useful for understanding reading fluency development. Chall provided a coherent framework for understanding broad reading development by describing how reading subskills, including reading fluency, coalesce into mature reading over time. In contrast, Ehri narrowly focused on word recognition by explaining how readers develop efficient, automatic word recognition skills via sight word development.

***Chall's stages of reading development.*** Chall (1996) outlined six stages of reading development from the pre-school, emergent literacy stage (Stage 0) to the collegiate-level, mature literacy stage (Stage 5). Of the six stages, Chall's second and third stages (Stages 1 and 2, respectively) are most relevant to fluency development (Kuhn & Stahl, 2003). Stage 1 includes the beginning of formal literacy instruction when readers are introduced to sound-symbol correspondences. For reading in alphabetic languages, this is known as the alphabetic principle, the concept that letters represent sounds of spoken words. Readers in Stage 1 use knowledge of the alphabetic principle to develop decoding skills. As readers decode the letter-sounds with greater accuracy, they enter Stage 2, the confirmation, fluency, and "ungluing from print" stage (Chall, 1996, p. 18). In this stage, beginning readers "use their decoding knowledge, the redundancies of the language, and the redundancies of the stories read…[to] gain fluency and speed" (Chall, 1996, pp. 18-19). Readers in Stage 2 confirm what they already know about print and letter-sound correspondences, develop quick and accurate word recognition abilities, and extend beyond print to attend to meaning and comprehension of text. They develop automaticity not only in decoding and word recognition but also in constructing meaning from text.

***Ehri's stages of sight-word recognition.*** In her stage model, Ehri (1995, 1999) dissected readers' progress in Stages 1 and 2 with more detail. In describing how readers achieve fluent, automatic decoding during early reading development, Ehri (1998) contended that fluent readers are able to identify words instantly by sight rather than by decoding each letter (i.e., identifying each letter's sound and then blending the sounds together into the spoken word). The more words a reader can identify on sight, the faster and more fluently he can read. Thus, Ehri (1995) argued that automaticity is achieved through sight-word learning and that readers progress through four predictable stages of sight-word learning: pre-alphabetic, partial alphabetic, full alphabetic, and consolidated alphabetic phases. In each phase, readers learn sight-words using different types of connections, from visual to alphabetic to consolidated multi-letter connections (Ehri, 1998).

In the pre-alphabetic phase, emergent readers recognize words primarily through visual cues rather than phonetic cues. In other words, they do not use letters or decipher letter-sounds to *read* print. Instead, they identify words in environmental print by remembering specific visual features without attending to phoneme-grapheme relationships (i.e., letter-sound correspondences in the alphabetic system; also called the graphophonic system). For example, a pre-alphabetic reader might identify a *Target* department store by recognizing the circles in its logo rather than recognizing the word *Target*. Similarly, he or she might *read* a stop sign by recognizing its octagonal shape and red color rather than the letters *STOP* written on it.

During the partial alphabetic phase, emergent readers begin to use phonetic cues to read words. As they begin to decode letters into sounds, they connect with only the most salient sounds in the word, such as a word's initial or final sound. In the partial alphabetic stage, a reader might identify the word *soup* as *stop* because he has connected the first and last letters to the initial /s/ and final /p/ sounds in the word *stop*.

As readers become more proficient with the graphophonic system, they move into the full alphabetic phase. Whereas partial-alphabetic readers utilize only the most salient letter-sound connections in a word, fully alphabetic readers have complete knowledge of the graphophonic system and utilize all letters when decoding the pronunciation of a word. According to Ehri (1998), readers

> [form] connections between graphemes in the spellings and phonemes underlying the pronunciations of individual words. The connections are formed out of readers' general knowledge of the grapheme-phoneme correspondences…they apply their graphophonic knowledge to analyze how letters symbolize individual phonemes detectable in the word's pronunciation. This secures the sight word in memory. (p.13)

Knowing the graphophonic system then becomes a "powerful mnemonic…that bonds the written forms…to their pronunciations in memory" (Ehri, 1998, p. 15). Words that are encountered repeatedly become sight words, recognizable in an instant.

As sight vocabulary expands, "letter patterns that recur across different words become consolidated" in memory (Ehri, 1995, p. 121). In this consolidated alphabetic phase, a reader synthesizes frequently encountered word chunks (e.g., -AN, -END, -ING, -ER) and stores them as units in memory. For example, rather than learning the sight word *send* by connecting and memorizing four separate grapheme-phoneme combinations (S-/s/, E-/ε/, N-/n/, and D-/d/), a consolidated alphabetic reader with –END unitized in memory would only need to make two connections—S-/s/ and END-/ɛnd/. Unitizing letter patterns in this manner facilitates both word recognition and sight word learning and enables readers to read more quickly, accurately, and fluently (Ehri, 1995).

**Oral Language Contributions to Fluency Development**

Ehri's theory of sight word development has been called "elegant" (Chard et al., 2006, p. 44). In it, she not only accounted for readers' transitions from alphabetic knowledge to automatic word recognition, but she also accounted for the role of oral language proficiency in word recognition processes. Pikulski (2006) asserted:

> Ehri show[ed] that progress in reading beyond the beginning stages is dependent on oral-language development, pointing out that reading words, particularly reading them fluently, is dependent on familiarity with them in their oral form. If the syntactic and meaning aspects of the word are to be activated, they must be part of what the reader knows through oral-language development. For the word-recognition process as proposed in Ehri's theory to be complete, it must connect with meaning that has been developed as another aspect of language development. (p. 82)

For readers to develop automatic sight word recognition efficiently, they must have oral language skills (Ehri, 1998). They must be familiar with the phonemic system, or the sounds of the spoken language, and be able to segment phonemes in order to decode letters into sounds. They must also have syntactic knowledge—knowledge of the language-specific rules for how to connect words into phrases, clauses, and sentences—in order to construct meaning from text. Additionally, they must have lexical (or vocabulary) knowledge, a mental dictionary of familiar words and their meanings, which, like syntactic knowledge, is necessary for constructing meaning from text.

Without phonemic, syntactic, and lexical knowledge of spoken language, beginning readers are unable to form strong connections among the written word, its pronunciation, and its meaning. They do not benefit fully from the graphophonic system's function as a mnemonic for learning sight words and have difficulties accessing word meanings due to underdeveloped lexical memory. Thus, for beginning readers with oral language lacunae, word identification and meaning construction processes will function more slowly and inaccurately, impeding the development of automatic, fluent reading ability.

## Theoretical Implications for L2 Reading Fluency

Although there is no theory of reading fluency that specifically applies to L2 reading, the theoretical background discussed here has two important implications for ELLs reading in English. The first is how underdeveloped L2 oral language proficiency affects L2 reading fluency. The second involves the association between reading fluency and reading comprehension for L2 readers.

**L2 oral language proficiency and L2 reading fluency.** Oral language proficiency "includes both receptive and expressive skills, and can also encompass knowledge or use of specific aspects of oral language, including phonology, vocabulary, morphology, grammar, discourse features, and pragmatic skills" (August & Shanahan, 2006, p. 55). By definition, ELLs have limited English oral language proficiency. ELLs learn to read in English despite limited capacity to communicate orally in their second language (L2). They likely have deficits in the oral language skills (e.g., phonemic awareness, syntactic knowledge, and lexical knowledge) that, according to Ehri's theory, may contribute to fluent reading.

The relationship between L2 oral language proficiency and L2 reading fluency has received little attention in research (August & Shanahan, 2006). Jackson and Lu (1992) found that gifted bilingual readers, though they had less proficiency in oral English, read as fluently as a comparison group of monolingual readers of English. In this case, English oral language proficiency did not appear to predict reading fluency skills because students' reading speed and accuracy were at a higher level than their oral proficiency would suggest. However, Geva and Yaghoub Zadeh (2006) found that oral language proficiency was a significant predictor of text reading efficiency for English-as-a-second-language second graders but not for a NES comparison group. Moreover, specific aspects of L2 oral language proficiency predict L2 reading fluency. Yesil-Dagli (2011) reported that receptive vocabulary was a strong predictor of ORF initial level and growth for first-grade ELLs, and Yaghoub Zadeh, Farnia, and Geva (2012) reported that first-grade listening comprehension predicted third-grade reading fluency for ELLs. From these few studies, it appears that L2 oral language proficiency may be important for the development of L2 reading fluency.

**L2 reading fluency and L2 reading comprehension.** According to automaticity theory, when decoding and meaning retrieval are done automatically, the reader's attention is free to focus on global comprehension of the text. Thus, fluent readers have more attentional resources available for comprehending text resulting in better comprehension of text. As would be predicted with automaticity theory, researchers have documented a strong, positive relationship between oral reading fluency and reading comprehension in NES students (Baker et al., 2008; Chard et al., 2006; Deno et al., 1982; Fuchs et al., 2001; Fuchs et al., 1988; Klauda & Guthrie, 2008; Marston, 1989; Shinn et al., 1992). This positive relationship also appears to extend to ELLs who are reading in L2 (Baker et al., 2012; Baker & Good, 1995; Muyskens et al., 2009; Wiley & Deno, 2005). As L2 reading fluency increases, L2 reading comprehension increases.

However, researchers recently have suggested that the relationship between L2 reading fluency and L2 reading comprehension is less straightforward for ELLs reading in L2. For L2 reading, the association between reading fluency and reading comprehension depends upon other precursory reading skills, particularly L2 oral proficiency. Crosson and Lesaux (2010) found that L2 oral proficiency moderated the relationship between L2 reading fluency and L2 reading comprehension. In their study, only ELLs who read fluently *and* had strong oral language proficiency exhibited strong reading comprehension. Similarly, Yaghoub Zadeh et al. (2012) reported that "once the prerequisite reading skills (phonological awareness, naming speed, word reading, and language proficiency) that underlie [reading fluency and reading comprehension] are modeled, the association between them becomes nonsignificant" (p. 182). This finding led Yaghoub Zadeh and colleagues to conclude that, for ELLs, the relationship between L2 reading fluency and L2 reading comprehension is best understood through their shared underlying factors: word-level reading skills and language proficiency.

## Literature Review

In the second half of this chapter, I review the extant literature related to the two purposes of the present dissertation. First, I describe what is known about the development of reading fluency for both native English speakers and ELLs. Then, I conclude the chapter with a discussion of reading fluency interventions for ELLs.

### ORF Development

The first purpose in the present study was to model the trajectory (i.e., initial status and growth) of English oral reading fluency for Spanish-speaking ELLs from second through third grades. Consistent with this purpose, I now review the research regarding ELLs' English oral reading fluency development. However, before discussing ELLs' reading fluency, I review the literature describing native English-speaking students' oral reading fluency development. Doing so provides context to ELLs' reading fluency development. First, much of the available reading fluency research has been conducted with native English-speaking students; researchers have only recently begun to investigate ELLs' reading fluency. Second, although ELLs do not command English language and literacy skills at native-speaker levels, ELLs' performance in English reading is typically interpreted using NES benchmarks or norms (Linan-Thompson et al., 2007). Using NES reading fluency benchmarks may or may not be appropriate for monitoring ELLs' progress (Al Otaiba et al., 2009; Dominguez de Ramirez & Shapiro, 2006; Gersten et al., 2007; Linan-Thompson et al., 2007; Yesil-Dagli, 2011). However, in the absence of benchmarks for L2 reading fluency, NES benchmarks and reading fluency development patterns provide a useful comparison.

**Reading fluency development in NES students.** Studies of NES students have characterized oral reading fluency growth in three ways: (a) grade-level performance benchmarks by a normative sample (Hasbrouck & Tindal, 1992; Hasbrouck & Tindal, 2006), (b) point-estimates of improvement rates (Fuchs et al., 1993; Deno et al.,2001) and (c) the trajectory of growth over time (Fuchs et al., 1993; Speece & Ritchey, 2005; Chard et al, 2008; Kim et al., 2010).

Hasbrouck and Tindal (1992, 2006) published two sets of ORF performance benchmarks for use in progress monitoring. The normative samples in these two studies were predominantly English-proficient students. ELLs were included in the Hasbrouck and Tindal (2006) sample, but the exact numbers or proportions were unknown. Results indicated the performance levels at three times per year (fall, winter, spring) for Grades 2-6 (Hasbrouck & Tindal, 1992) and Grades 1-8 (Hasbrouck & Tindal, 2006). Expected rates of gain can be calculated from the performance levels, assuming 30 weeks of instruction between fall and spring assessments. According to the 2006 norms, the average student in Grade 1 read 23 wcpm in winter and 53 wcpm in spring. (Note: Grade 1 norms begin in winter rather than fall.) Thus, the average student gained 30 words over approximately 15 weeks of instruction, an average gain of 2 wcpm per week. From fall to spring, the average student in Grade 2 and Grade 3 gained 38 wcpm and 36 wcpm, respectively. Assuming 30 weeks of instruction between fall and spring, students in Grades 2 and 3 gained approximately 1.25 wcpm per week. For students in Grades 4 and 5, the gain from fall to spring was 29 wcpm, an average weekly gain of just under 1 wcpm. In sum, to reach the spring performance levels of Hasbrouck and Tindal (2006), students would have to make large weekly gains in first grade (≥ 2 wcpm per week). Rates of gain could slow in second and third grades (> 1 wcpm per week) and even more in fourth and fifth grades (approximately 1 wcpm per week).

Fuchs et al. (1993) and Deno et al. (2001) reported average reading fluency growth rates for general education students at specific grade levels. Table 1 shows how the average growth rates across studies, including Hasbrouck and Tindal (2006), are similar. Typical ORF growth rates for native speakers in general education decrease over time. During early reading instruction in Grade 1, the reading fluency growth rate is approximately 2 wcpm per week. In Grades 2 and 3, respectively, reading fluency growth decreases to between 1.25 and 1.5 wcpm per week and then to between 1 and 1.25 wcpm per week. In the later grades, students continue increasing in fluency speed but increase at a slower pace. Growth decreases to less than 1 wcpm per week in Grades 4 and 5.

Findings regarding reading fluency growth trajectory confirm these patterns. Within a single year, growth follows a linear trajectory (Fuchs et al., 1993; Stage et al., 2001; Stage & Jacobsen, 2001). However, across multiple academic years, growth follows a quadratic trajectory (Fuchs et al., 1993; Speece & Ritchey, 2005; Baker et al., 2008; Kim et al., 2010). The typical NES reading fluency growth pattern is acceleration through first and second grades followed by deceleration at the end of second grade and through third grade (Baker et al., 2008; Chard et al., 2008; Crowe et al., 2009; Kim et al., 2010). Students continue weekly improvements in fluency, but their rates of improvement slow near the beginning of third grade.

Table 1
*Expected Weekly Oral Reading Fluency (ORF) Gains* [a] *for Native English-Speaking Students*

| Grade | Hasbrouck & Tindal (1992) [b] | Fuchs et al. (1993) | Deno et al. (2001) | Hasbrouck & Tindal (2006) [b] |
|---|---|---|---|---|
| 1 | | 2.0 | 1.8 | 2.0 |
| 2 | 1.37 | 1.5 | 1.66 | 1.27 |
| 3 | 1.17 | 1.0 | 1.18 | 1.2 |
| 4 | 0.63 | 0.85 | 1.01 | 0.97 |
| 5 | 0.76 | 0.5 | 0.58 | 0.97 |

[a] ORF growth rate is expressed in wcpm per week.
[b] Hasbrouck & Tindal (1992, 2006) reported ORF benchmarks for a normative sample at Fall, Winter, and Spring observations within each grade (First grade: Winter and Spring only). The growth rates shown here are the expected rates of gain for a student at the 50th percentile, assuming 30 weeks between Fall and Spring observations.

**Reading fluency development in ELLs.** Longitudinal ORF growth patterns for English-language learners have received less attention in research, but research on this topic is burgeoning. ELLs' reading fluency development has been the subject of seven studies (Al Otaiba et al., 2009; Baker et al., 2012; Baker & Good, 1995; Betts et al., 2009; Dominguez de Ramirez & Shapiro, 2006; Graves, Plasencia-Peinado, Deno, & Johnson, 2005; Yesil-Dagli, 2011). ORF growth rates observed in the seven studies are synthesized in Table 2.

Table 2
*Expected Weekly Oral Reading Fluency (ORF) Gains* [a] *for English-language Learners*

| Grade | Baker & Good (1995) | Graves et al. (2005)[b] | Dominguez de Ramirez & Shapiro (2006) | Al Otaiba et al. (2009) | Betts et al. (2009) | Yesil-Dagli (2011)[c] | Baker, Park, & Baker (2012) [c] |
|---|---|---|---|---|---|---|---|
| 1 | | L: 2.8<br>A: 3.6<br>H: 1.8<br>Total: 2.75 | 0.57 | | | 1.26 | |
| 2 | 1.3 | | 0.75 | 1.23 | | | 1.52 |
| 3 | | | 0.48 | 1.31 | L1 Spanish: 1.27<br>L1 Somali: 1.17 | | 1.54 |
| 4 | | | 0.44 | | | | |
| 5 | | | 0.71 | | | | |

[a] ORF growth rate is expressed in wcpm per week.
[b] Sample was divided into low-achieving (L), average (A), and high-achieving (H) readers.
[c] Weekly rate calculated from reported annual gains reported by the researchers, assuming 30 weeks of instruction between first and last measurement.

Dominguez de Ramirez and Shapiro (2006) compared ORF growth for ELLs and NES students. They studied a cross-sectional sample of ELLs (*N*=68) in a transitional bilingual education program in Grades 1-5. English ORF was measured three times in 1 year, and average slopes were reported for each grade (see Table 2). At all grades, ELL students grew at a slower rate than the NES comparison group in the study and well below the typical NES growth rates reported in Table 1. Although acceleration could not be statistically tested in the study, ELLs appeared to accelerate growth in second grade, decelerate in third and fourth, and accelerate again in fifth, which would depart from the typical NES pattern. However, the study's small samples in each grade and cross-sectional design limit inferences about typical longitudinal growth patterns for ELLs.

Graves, Plasencia-Peinado, Deno and Johnson (2005) investigated the use of ORF to formatively assess ELLs' reading skills. Graves and colleagues assessed first-grade ELLs for 6 weeks at the end of Grade 1. The sample was separated into low-, average-, and high-achieving readers. Low readers (*n*=27) gained 2.8 wcpm per week, average readers (*n*=23) gained 3.6 wcpm per week, and high readers (*n*=27) gained 2.75 wcpm per week. These gains of >2 wcpm surpass the first-grade rates of gain for native English speakers in Table 1.

Baker and Good (1995) studied the validity and reliability of ORF scores for second-grade English learners (*N*=50) and found evidence that, like NES students, ELLs slow their L2 reading fluency growth during second grade. During the 10-week observation period, students gained fluency at a rate of 1.3 wcpm per week. This rate is consistent with the expected rate of gain for second-grade NES students.

ELLs’ ORF growth trajectory was modeled statistically in four studies. Yesil-Dagli (2011) modeled L2 ORF growth in first grade, Al Otaiba et al. (2009) and Baker et al. (2012) modeled L2 ORF growth in second and third grades, and Betts et al. (2009) modeled L2 ORF growth in third grade.

Yesil-Dagli (2011) studied 2,481 first-grade ELLs and identified a positive curvilinear (quadratic) trajectory in ELLs’ fluency development, indicating that students accelerated their fluency growth during first grade. Linear growth, or rate of improvement, was 7.92 wcpm every 2 months. Quadratic growth, or rate of acceleration, was 0.5 wcpm every 2 months. ELLs’ average yearly gain was 38 wcpm, similar to expected first-grade gains for native speakers (Hasbrouck & Tindal, 2006).

Al Otaiba et al. (2009) modeled Latino students’ fluency growth over second and third grades for 1,767 Latino ESL students who were instructed in general education classrooms. During second grade, ESL students improved their fluency in a strong, linear pattern but showed slight deceleration at the end of the school year. In third grade, students’ initial growth followed a steep, positive linear trend (+2 wcpm per week in the first 2 months of the school year). But growth decelerated in a strong, negative quadratic trend as third grade progressed. In spite of growth over both years, these ESL students in general education continued to be classified as struggling readers according to ORF benchmarks.

Like Al Otaiba et al. (2009), Baker et al. (2012) modeled L2 fluency development in second and third grades. However, unlike Al Otaiba et al., they did not follow the same students across 2 years but collected data from separate samples of second- and third-grade ELLs who received bilingual reading instruction. English ORF was assessed three times during the year, and initial status and growth were estimated in a model for each grade. Second-grade ELLs began the year reading 30 wcpm and gained 46 wcpm during the year. Third-grade ELLs began the year reading 54 wcpm and also gained 46 wcpm during the year. Assuming 30 weeks of instruction between the first and last ORF assessment, second- and third-grade ELLs gained approximately 1.5 wcpm per week, slightly outpacing expected gains for NES students.

Betts et al. (2009) studied ELLs in third grade and found that a linear model of fluency growth fit the data well. The sample included 300 beginning-level ELLs, 207 Spanish-speakers and 93 Somali-speakers. Students' English ORF was assessed three times during the year, and a multiple-group structural equation model was applied to the data. Both groups of students made comparable average weekly gains of just over 1 wcpm. Spanish-speaking students made weekly gains of 1.27 wcpm. Somali-speaking students made weekly gains of 1.17 wcpm. On average, students gained 37 wpm over the course of the third grade year. These weekly and yearly gains are comparable to the Hasbrouck and Tindal (2006) norms for third-grade native English speakers.

These findings provide tentative growth expectations for ELLs in first through third grades. Although only one of the four studies of trajectory characterized growth in the same ELLs through multiple grades, it appears that the general pattern of L2 reading fluency development is similar to that of monolingual readers: acceleration during first and second grades followed by deceleration beginning in third grade. Nevertheless, the number of studies is limited. More studies of reading fluency development are needed to confirm the similarities of growth rates and developmental patterns for ELLs and NES students.

**ORF Interventions for ELLs**

The second purpose of this study was to determine the effect of a comprehensive, 4-year language and reading intervention for ELLs on L2 reading fluency. In the final section of this chapter, I review studies of reading interventions for ELLs that measured ORF as an outcome.

Several interventions have been studied for their effects on ELLs' English oral reading fluency: peer-assisted learning strategies (Calhoon et al., 2007; McMaster et al., 2008; Saenz et al., 2005), direct instruction of reading skills delivered in small groups (Gunn et al., 2000; Gunn et al., 2002; Linan-Thompson et al., 2003; Vaughn, Cirino, et al., 2006; Vaughn, Mathes, et al., 2006), and individual tutoring (Li & Nes, 2001; Santoro, Jitendra, Starosta, & Sacks, 2006). These interventions varied in mode of delivery from whole-class to small-group to individual, but all were delivered as supplements to students' core reading curriculum.

**Peer-assisted learning strategies.** Findings from three studies support peer-assisted learning strategies (PALS), a version of classwide peer-tutoring, as a method for developing ELLs' English reading skills. The three studies have been conducted with English learners across grades and program models: kindergarteners receiving ESL instruction in a pull-out model (Calhoon et al., 2007), first graders in two-way bilingual immersion programs (McMaster et al., 2008), and ELLs in Grades 2 through 6 in transitional bilingual education programs (Saenz et al., 2005). In all three studies, PALS was implemented in a whole-class, general education setting and included explicit, structured instruction in developmentally appropriate reading skills. Kindergarten PALS (K-PALS) focused on phonemic awareness, letter-sound correspondence, and decoding (Calhoon et al., 2007). The Grade 1 and Grades 2-6 versions of PALS focused on fluency, comprehension, and strategic reading (McMaster et al., 2008; Saenz et al., 2005). No statistically significant differences between the contrast groups and treatment groups were found in any of the three studies. However, positive effect sizes (ES) favoring the PALS treatment groups were reported. McMaster et al. (2008) reported the smallest effects of 0.18 and 0.10 on two oral reading passages. Calhoon et al. (2007) reported an effect size of 0.38 for the intervention group's pre-to-posttest growth in ORF. Saenz et al. (2005) reported the largest effect of 0.60.

**Individual tutoring.** In two studies, individual tutoring was investigated. Li and Nes (2001) investigated the effects of paired reading with 4 English learners of Chinese origin in Grades 1, 2 and 3. The study design was a modified AB (baseline-intervention) single-subject design. The intervention consisted of 55 40-min sessions of paired reading with a skilled reader. By the end of the intervention, students had improved in both reading fluency and reading accuracy and continued that improvement during a modified maintenance phase when the intervention was partially withdrawn. However, this study had several methodological weaknesses that undermined its conclusions regarding the effectiveness of paired reading. First, the study's relatively weak AB design did not allow strong causal inferences of intervention efficacy. In addition, information about the equivalency of the probes used to measure oral reading fluency, which is a central requirement for single-subject research, was not provided. Finally, no statistical analyses were performed on the data. Because the study's effects were not quantified with effect sizes, it is difficult to compare the outcomes with other studies.

In a second study of individual tutoring with ELLs, Santoro, Jitendra, Starosta, and Sacks (2006) also utilized a single-subject design, though a much stronger multiple-probe, multiple-baseline design, to investigate the impact of *Read Well* on the English reading skills of 4 second-grade English learners with low reading performance. *Read Well* (Sprick, Howard, & Fidanque, 1998-2000) is a commercially available reading curriculum that incorporates systematic, explicit instruction of phonological awareness, phonics, fluency, vocabulary, and comprehension. At the end of the intervention, which was given in daily, 30-min sessions for between 7 and 14 weeks, all students had improved their mean level of oral reading fluency. The mean effect size across the 4 students was 0.71; however, this mean was unduly influenced by one student's effect size of 1.67. The effect sizes for the other 3 students were small to moderate, ranging from 0.24 to 0.63.

Although individual tutoring interventions appear to have positive effects on ELLs' oral reading fluency, one cannot draw firm conclusions about the effectiveness of such interventions from only two studies. Additionally, as acknowledged by Santoro et al. (2006), individual tutoring is difficult to implement in schools. One-to-one tutoring often requires time and extra resources (e.g., for hiring and training additional personnel) that many schools do not have.

**Small-group direct instruction in reading.** Supplemental direct instruction in small groups is the third type of reading intervention with evidence for improving ELLs' ORF (Gunn et al., 2000, 2002; Linan-Thompson et al., 2003; Vaughn, Cirino, et al., 2006; Vaughn, Mathes, et al., 2006). These interventions featured direct and explicit instruction of reading skills in groups of 2 to 5 students. Each intervention included a combination of the following skills: phonemic awareness, letter knowledge and letter-sound correspondence, word recognition and decoding, connected text fluency, comprehension, and vocabulary. Furthermore, the interventions included an English language support component to make language more comprehensible to second language learners. The level of English language support ranged from instructor-provided vocabulary assistance and background knowledge (Gunn et al., 2000) to a 10-min oral language development component (Pollard-Durodola, Mathes, Vaughn, Cardenas-Hagan, & Linan-Thompson, 2006; Vaughn, Cirino, et al., 2006; Vaughn, Mathes, et al., 2006).

Gunn, Biglan, Smolkowski, and Ary (2000) concluded that students of very limited English proficiency were able to improve their reading fluency through 2 years of participation in a supplemental reading intervention in English. The intervention included 30 min of daily "explicit instruction to develop word recognition skills, accompanied by clear feedback, active engagement, and cumulative review" (Gunn, Smolkowski, Biglan, Black, & Blair, 2005, p. 68). Hispanic students constituted 62% of the sample (*n*=195), but only 19 students were identified as having very limited English proficiency. These students were placed with bilingual instructional assistants, who spent extra time developing vocabulary and background knowledge during supplemental intervention sessions. Information about the English language proficiency of the remaining Hispanic students was not reported. A sub-analysis of the 19 students who spoke little or no English at study outset indicated a statistically significant difference in ORF for non-English speakers who received the intervention compared to non-English speakers in the control group. The effect size for this difference was not reported.

In a 1-year follow-up study, Gunn, Smolkowski, Biglan, and Black (2002) investigated whether the benefits of supplemental instruction were maintained. One year post-intervention, Hispanic students in the treatment condition maintained higher performance on oral reading fluency (ES=0.46) when compared with Hispanic students in the control group. No interaction effects between ethnicity and treatment condition were observed, indicating that Hispanic students benefited from the supplemental instruction as much as non-Hispanic students did. In an additional analysis, the authors found no interaction between treatment condition and initial English language proficiency (English-speaking vs. non-English speaking), though they cautioned that a Type II error is likely given that so few (only 19) students were non-English speaking. Gunn, Smolkowski, Biglan, Black, and Blair (2005) followed up on student progress 2 years post-intervention. Oral reading fluency scores of intervention students continued to improve relative to control group students. Analyses of post-intervention oral reading fluency data for students who began with limited- or non-English proficiency were not reported.

In the Gunn et al. (2000, 2002, 2005) studies, ELLs were only a small portion of the sample. Although almost two-thirds of the sample were Hispanic students, initial assessments of language proficiency indicated that most of the Hispanic students spoke both Spanish and English. Only 19 students were identified as non-English speaking at the outset of the study. Thus, the analyses conducted to ascertain intervention effects for this small portion of the sample should be interpreted with caution. In addition, the procedures to determine language proficiency of students were imprecise. Project assessors only spoke with students to determine whether they could converse in English and Spanish. Students who conversed in Spanish only were considered non-English proficient. Students who could converse in English were considered English proficient. Thus, the assessments of language proficiency were not based on rigorous, objective measures, and analyses of subgroups based on the subjective assessment of language proficiency are questionable.

Linan-Thompson, Vaughn, Hickman-Davis, and Kouzekanani (2003) found that second-grade ELLs who were at risk for reading problems made statistically significant gains in ORF from pretest to posttest (standardized mean difference (SMD)=1.61). Students received daily instruction for 30 to 35 min over a 13-week period in groups of 2 or 3. Students were assessed at the end of the intervention and on two follow-up occasions, at 4 weeks and again at 4 months after intervention. On average, students increased their reading fluency by approximately 2 words per week while receiving the intervention. At the 4-week follow-up, students continued to increase their fluency, but the increase was not statistically significant (SMD=0.43). However, at the 4-month follow-up, students' gain in reading fluency was statistically significant (SMD=0.62). Despite this post-intervention gain in mean level, students did not maintain the rate of gain in fluency (2 words per week) that they had achieved during the intervention, which lends support to the intervention as causal mechanism. Nevertheless, without a control group, such an inference is tenuous at best.

Vaughn, Cirino, et al. (2006) and Vaughn, Mathes, et al. (2006) investigated supplemental English reading instruction with first-grade ELLs who were at risk for reading problems. Unlike the Linan-Thompson et al. (2003) study, both of these studies included a contrast group that received the schools' regular intervention for struggling readers. In both studies, the intervention was 7 months in duration and consisted of 50 min of daily instruction in small groups of 3 to 5 students. Reading instruction consisted of systematic, explicit instruction in the five strands of reading: phonemic awareness, letter knowledge, word recognition, connected text fluency, and comprehension. The reading intervention, which was designed for monolingual English speakers with reading difficulties, was supplemented with a structured 10-min English oracy intervention to develop English oral language (Pollard-Durodola et al., 2006). In addition, the reading curriculum was modified to include language support activities. The results of both studies were similar. Comparing the ORF of intervention group to the contrast group, the researchers did not find statistically significant differences between groups in either study; however, they reported effect sizes favoring the intervention groups. For two ORF assessments, Vaughn, Cirino, et al. reported 0.32 and 0.27, and Vaughn, Mathes, et al. reported 0.16 and 0.18.

**Comprehensive second language/literacy intervention.** Tong, Irby, Lara-Alecio, and Mathes (2008) studied the effects of a multiyear, comprehensive L2 intervention on 589 second-grade ELLs in bilingual education programs. The intervention included explicit, structured L2 oral language development in kindergarten and first grade, direct instruction of L2 reading skills in the second semester of first grade and throughout second grade, and context-embedded vocabulary instruction in all three grades. In each grade, the intervention was 7 months in duration and consisted of 75 min of daily instruction in kindergarten and 90 min daily in first and second. At the end of second grade, the researchers found statistically significant differences between the intervention and control groups on English ORF. The reported effect size was 0.16, favoring the intervention group.

The intervention studied by Tong, Irby, and colleagues (2008) incorporated practices found in other interventions reviewed here: direct instruction of L2 reading skills, oral language development, whole-class instruction, and supplemental instruction for struggling students. However, the Tong, Irby, et al. intervention is distinguished by its nature as a comprehensive second language and reading intervention with a developmental focus over multiple years. In addition, the Tong, Irby, et al. study as a whole is distinguished from other studies in this review by its longitudinal design, a large sample, and the presence of control group.

**Summary of ORF interventions for ELLs.** These studies comprise the burgeoning body of literature on the effects of instructional interventions on English ORF with samples of English-language learners. The studies are diverse in design and vary in methodological quality. With the exception of Tong, Irby, et al. (2008), the studies had smaller samples of ELLs. Saenz et al. (2005) used a larger sample (*N*=119), but the analyses were conducted at the teacher level (*N*=12). Thus, with such a small teacher-level sample, the researchers failed to find statistically significant differences in spite of the intervention's moderately large effect size (0.60). Generally, researchers reported positive effect sizes for ORF. Linan-Thompson et al. (2003) found statistically significant pre-post gains in ORF for ELLs who received small-group direct instruction in reading, but there was no contrast group in the study. In studies where the intervention group was compared to a contrast group on ORF assessments, only Tong, Irby, et al. (2008) reported statistically significant differences. Though there is evidence that peer-assisted learning strategies, individual tutoring, direct instruction in reading, and the comprehensive L2 intervention studied by Tong, Irby, et al. each have practical effects on ELLs' ORF, additional research findings could confirm these effects and/or identify other effective interventions for ELLs' reading fluency.

## Conclusion

In this chapter, I have outlined the theoretical background for this dissertation. I defined reading fluency as reading with speed, accuracy, and expression, although some expand the definition to include comprehension and component decoding skills. I described oral reading fluency, the number of words read correctly in 1 min of connected text reading. Using ORF as an operational definition of reading fluency follows the convention of the field and has strong support in the literature. After defining the construct under study in this dissertation, I discussed the relevant theories of reading fluency, showing the importance of automatic processing for fluency (LaBerge & Samuels, 1974) and the role of fluency in broad literacy development (Chall, 1996). Next, I explained the stage model of sight-word development, in which Ehri (1995) described how the letters of a word are imprinted into a reader's memory and connected to lexical memory through phonological representations. I concluded the theoretical background by discussing the contributions of oral language development and the implications of reading fluency theories for L2 reading fluency.

In addition to providing theoretical background, I also reviewed the relevant literature with the following findings. Expected performance levels and growth of ORF for native English speakers are well established. These are the standards by which ELLs are typically measured because expectations for ELLs' fluency development are not well defined. Seven recent studies have explored L2 fluency growth in first, second, and third grades. Though findings from these studies suggest that ELLs' fluency growth is similar to NES students' growth, more research is warranted to provide stronger evidence for similarities or to investigate possible differences within the same general patterns. For example, ELLs' may accelerate fluency over a longer time period than NES students accelerate (Al Otaiba et al., 2009), or ELLs may experience two periods of acceleration (Dominguez de Ramirez & Shapiro, 2006). Finally, it is important to understand how ELLs respond to reading instruction by analyzing ORF outcomes. L2 reading fluency instruction has been moderately addressed by researchers, and positive effects have been associated with PALS and supplemental direct instruction of reading skills. But, this body of intervention studies would be more robust with additional studies, like Tong, Irby, et al. (2008), that: (a) have larger samples (>100 participants), (b) include control or contrast groups, and (c) include comprehensive second language and literacy development within the intervention. In Chapter III, I describe the methodology used in this dissertation study, which contributes to the extant knowledge related to: (a) ELLs' L2 reading fluency development and (b) ELLs' response to language and reading instruction.

# CHAPTER III

# METHODOLOGY

The purposes of this study were to: (a) model the trajectory (i.e., initial status and growth) of second language (L2) oral reading fluency for Spanish-speaking English-language learners from second through third grades and (b) determine the effect of a 4-year structured intervention in English language and reading on the development of L2 oral reading fluency in ELLs. In this chapter, I describe the methodology for the investigation. I outline the study design and describe the context, participants, and instrumentation. Then, I explain the data collection and intervention procedures. Finally, I describe the data analysis methods.

## Design of the Study

This study is part of a larger, federally funded research project, Project ELLA (English Language/Literacy Acquisition; R305P030032).[1] The primary goal of Project ELLA was to compare typical and enhanced models of structured-English immersion (SEI) and transitional bilingual education (TBE). The enhanced models' treatment, which will be described in more detail later in this chapter, was a multi-tiered intervention focused on developing English oral language, vocabulary, and reading. A sample of Spanish-speaking English-language learners from Kindergarten through Grade 3 was followed, and students' language and reading skills in English and Spanish were documented before, during, and after the project.

[1] Data for this dissertation were obtained from the data archives of Project ELLA, a completed project funded by grant from the U.S. Department of Education, Institute of Education Sciences (R305P030032).

Participating schools, where TBE and/or SEI programs were already operating, were randomly assigned to either the treatment (enhanced practice) or the control (typical practice) condition. Randomizing at the school level accomplished two goals. First, assigning an entire school to a condition minimized contamination of the intervention between treatment and control classrooms on the same campus. Second, the participating schools were able to observe Texas state law requiring a Language Proficiency Assessment Committee to determine limited-English proficient students' instructional placement in bilingual education or ESL programs. Because randomization was carried out at the school level, Project ELLA was experimental at the school level and quasi-experimental at the student level.

In the present study, archived data from the database of language and literacy outcomes measured during Project ELLA were analyzed. Specifically, the current study was focused on English oral reading fluency for students in typical and enhanced bilingual education conditions (e.g., No SEI students were included.). Figure 1, a graphic representation of the study's design, includes the sequence of treatment and measurement occasions for the two groups in the present study, enhanced-practice TBE (intervention condition) and typical-practice TBE (control condition). A multilevel growth model was used to analyze students' growth in English oral reading fluency during second and third grades.

| | Kindergarten | | | Grade 1 | | | Grade 2 | | | Grade 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | BOY | MOY | EOY | BOY | MOY | EOY | BOY | MOY | EOY | BOY | MOY | EOY |
| Intervention | X---------------------- | | | X---------------------------- | | | X------------------------------ ORF1 | ORF2 | ORF3 | X---------------------------- ORF4 | ORF5 | ORF6 |
| Control | | | | | | | ORF1 | ORF2 | ORF3 | ORF4 | ORF5 | ORF6 |

*Figure 1*. Research design. BOY=Beginning of the (school) year, MOY=Middle of the (school) year, EOY=End of the (school) year, X=3-tiered Project ELLA intervention, ORF=DIBELS oral reading fluency assessment.

## Context

Project ELLA was conducted in a large urban school district in Southeast Texas. The district was chosen because of its positive reputation and experience educating ELLs. It is one of the top districts for educating Latino students in Texas (Melton, 2009) and has been the winner of the Broad Prize for Urban Education. During the project, the school district served a diverse student body: 64% Hispanic, 31% African-American, 4% White, and 2% Other ethnicities.[2] In addition, 80.1% was classified as economically disadvantaged (qualifying for free or reduced-price lunch), and 31% had limited English proficiency. The district mobility rate was 24.7% (Texas Education Agency, 2008).

## Participants

### Schools

The present study included data from 17 schools, 8 intervention schools and 9 control schools. Table 3 depicts a comparison of intervention and control schools on key demographic characteristics. Intervention and control schools were similar on most characteristics listed, with the exception of percentage of ELLs. Intervention schools had enrolled a larger percentage of ELLs, resulting in a larger percentage of students served in bilingual/ESL programs and necessitating more teachers certified in bilingual education and/or ESL.

[2] Sum is greater than 100 due to rounding.

Table 3
*Descriptive Statistics for School Characteristics by Group*

| | Intervention Schools (*n*=8) | | Control Schools (*n*=9) | |
|---|---|---|---|---|
| Demographics | Mean | SD | Mean | SD |
| Student enrollment | 776 | 111 | 773 | 90 |
| Hispanic student enrollment (%) | 76.7 | 9.9 | 52.1 | 20.9 |
| Economically disadvantaged (%) | 86.2 | 2.9 | 81.5 | 6.3 |
| ELLs (%) | 60.4 | 9.3 | 37.0 | 16.4 |
| Mobility (%) | 23.0 | 6.4 | 27.7 | 8.1 |
| Student/teacher ratio | 14.8 | 0.9 | 15.1 | 0.9 |
| Minority staff (%) | 63.8 | 5.9 | 61.5 | 7.2 |
| Average teaching experience (years) | 9.9 | 1.8 | 9.6 | 2.2 |
| Students participating in ELL programs (%) | 58.1 | 10.1 | 34.9 | 16.3 |
| Teachers certified in bilingual education or ESL (%) | 45.4 | 8.3 | 32.9 | 11.8 |

**Teachers and Paraprofessionals**

In each grade, teachers at the participating campuses were randomly selected to participate in the project. Upon teachers' selection, participation in the project was voluntary. However, all teachers who were selected chose to participate. All teachers were highly qualified teachers with state bilingual and/or ESL certification. During the project, intervention teachers received professional development as outlined in the section describing the intervention. In the present study, data were collected from the students of 29 second-grade teachers and 22 third-grade teachers.

Paraprofessionals were hired to assist with classroom instruction during the intervention. All were highly-qualified with a minimum of 45 college credit hours, and all were Spanish-English bilinguals. They were trained biweekly to deliver intervention components. In addition, paraprofessionals assisted with student assessment. They were trained to give standardized language and literacy assessments and had to demonstrate their proficiency in administering specific assessments before being permitted to assess students.

**Students**

Project ELLA began with 822 kindergarten ELLs who had received parental approval to participate in the study in either bilingual education or SEI conditions. All students had Home Language Surveys indicating Spanish as the home language, and all were classified as limited-English proficient according to criteria specified in Chapter 89, Subchapter BB of the Texas Administrative Code (1996). The bilingual education intervention and control conditions contained a total of 472 students. Attrition at the end of kindergarten left 374 students (20.8% attrition) in the bilingual education sample. The complete project sample (students in both bilingual education and SEI conditions) was augmented with 120 students at the beginning of Grade 1. Fifty-one of those students were added to the bilingual education control condition. No students were added to the bilingual education intervention condition. At the end of Grade 3, 390 students remained in Project ELLA. Of those 390, 224 students had been instructed either the enhanced or typical-practice bilingual education conditions. There was 52.5% attrition in both the overall project and the bilingual education condition. Over 4 years, the average annual attrition rate was 13.1%, which is similar to other longitudinal research projects conducted in urban areas (Tong et al., 2010).

**Subsample for Present Study**

This study used a subsample from the typical-practice and enhanced bilingual education conditions. Students were selected if they participated in either the intervention or control bilingual education conditions and had at least one English oral reading fluency measurement during second and third grades. Based on these criteria, 283 students qualified for inclusion, 151 from the intervention condition and 132 from the control condition. The two groups' average age and gender composition are compared in Table 4.

**Sample Size at Each Level**

This study was conceptualized as a multilevel investigation: repeated oral reading fluency measures were nested in students who were nested in classrooms and schools. In Table 5 the sample of repeated measures at Level 1 is summarized. There were 1,470 oral reading fluency measurements taken in second and third grades. Those measurements were nested in 283 students, 76% of whom had measurements on all six measurement occasions. In Table 6, the sample size is further analyzed by group (intervention or control) and by level (students, classrooms, and schools). There were 17 schools, 8 intervention schools and 9 control schools. In Grade 2, there were 29 classrooms, 11 intervention classrooms and 18 control classrooms. In Grade 3, there were 22 classrooms, 10 intervention classrooms and 12 control classrooms.

Table 4
*Mean Age (Beginning of Grade 2) and Gender Composition by Group*

| | | Age, Grade 2 | | Gender | |
|---|---|---|---|---|---|
| | *n* | Mean | SD | Female (%) | Male (%) |
| Intervention (ELLA) | 151 | 7.59 | .37 | 71 (47.0) | 80 (53.0) |
| Control | 132 | 7.59 | .35 | 68 (51.5) | 64 (48.5) |
| Total | 283 | 7.59 | .37 | 139 (49.1) | 144 (50.9) |

Table 5
*Summary of Sample Size at Level 1*

| Number of ORF observations | Number of participants | % of Sample | Cumulative frequency of participants |
|---|---|---|---|
| 1 | 19 | 6.7 | 19 |
| 2 | 6 | 2.1 | 25 |
| 3 | 31 | 10.9 | 56 |
| 4 | 4 | 1.4 | 60 |
| 5 | 8 | 2.8 | 68 |
| 6 | 215 | 75.9 | 283 |
| Total ORF observations | 1470 | | |

Table 6
*Full Count of Students Nested within Classrooms within Schools*

| | Intervention | Control | Total |
|---|---|---|---|
| Schools | 8 | 9 | 17 |
| Students, beginning of Grade 2 (Time=0) | 151 | 132 | 283 |
| Average number of students per school | 18.9 | 14.7 | 16.7 |
| Students, beginning of Grade 3 (Time=3) | 117 | 110 | 227 |
| Average number of students per school | 14.6 | 12.2 | 13.4 |
| Classrooms (Teachers), Grade 2 | 11 | 18 | 29 |
| Average number of teachers per school | 1.4 | 2.0 | 1.7 |
| Average number of students per teacher | 13.7 | 7.3 | 8.7 |
| Classrooms (Teachers), Grade 3 | 10 | 12 | 22 |
| Average number of teachers per school | 1.3 | 1.3 | 1.3 |
| Average number of students per teacher | 11.7 | 9.2 | 10.3 |

## Instrumentation

The Dynamic Indicators of Basic Early Literacy Skills DIBELS (Good & Kaminski, 2002a) assess reading and reading-related skills. The DIBELS measures include subtests of phonemic awareness (Initial Sound Fluency, Phoneme Segmentation Fluency), decoding (Nonsense Word Fluency), accuracy and fluency (Oral Reading Fluency), and comprehension (Retell Fluency). Each subtest is individually administered following standardized procedures for curriculum-based measurement (Good & Kaminski, 2002b). In this study, only the Oral Reading Fluency (ORF) subtest was used.

DIBELS ORF measures students' ability to read grade-level connected English texts accurately and fluently. Students read a passage aloud for 1 min while an examiner records words read correctly and incorrectly. Omissions, substitutions, and hesitations of more than 3 seconds are counted as errors. Self-corrections (within 3 seconds) are counted as correct. The ORF score is the number of words read correctly in 1 min. Passage difficulty is equivalent to grade-level texts. Passages within grades are considered alternate forms; passages become slightly more difficult with each successive grade. In the present study, ORF benchmark passages were administered on six occasions: the beginning, middle, and end of both second and third grades.

ORF scores elicited via CBM procedures have strong technical adequacy (Deno et al., 2001; Marston, 1989). Test-retest reliability ranges from .92 to .97 (Good, Simmons, & Kame'enui, 2001; Marston, 1989). Interscorer reliability was reported to be .99 (Tindal, Marston, & Deno, 1983). Alternate-form reliability ranges from .89 to .94 (Tindal et al., 1983). Baker and Good (1995) investigated the reliability of English ORF scores for English-language learners. They reported an alternate-form reliability coefficient of .92 for ORF scores elicited from bilingual students in their sample. In addition, the reliability for estimates of ORF level was .99, whereas estimates of ORF slope were less reliable at .49. Baker and Good concluded that English ORF scores are sensitive to growth and adequately capture ELLs’ progress in reading. Alternate-form reliability coefficients were calculated using ORF scores from the present sample. The average alternate-form reliability coefficient was .84 for both the second-grade and third-grade ORF scores.

Along with strong evidence for ORF score reliability, there is a preponderance of validity evidence for the interpretation of ORF scores as indicators of reading proficiency. Results reported by Shinn, Good, Knutson, Tilly, and Collins (1992) provided evidence for the construct validity of ORF scores for readers in third grade. Two measurements of ORF were the observed variables most highly related ($r$=.88 and .90) to the latent variable representing reading competence in a confirmatory factor analysis. Several studies investigated the concurrent validity of ORF scores with common standardized assessments of reading achievement (e.g., Stanford Achievement Test-Reading Comprehension Subtest, Woodcock Reading Mastery Passage Comprehension) and reported correlation coefficients ranging from .73 to .91 (Deno et al., 1982; Fuchs & Deno, 1992; Fuchs et al., 1988; Marston, 1989). More recently, validity research has focused on the predictive validity of ORF scores for outcomes on high-stakes state reading tests. Seven studies reported correlations ranging from .43 to .80 ($M$=.66, $N$=18 coefficients) for ORF scores and seven different states' exams (Baker et al., 2008; Barger, 2003; Buck & Torgesen, 2003; Shaw & Shaw, 2002; Stage & Jacobsen, 2001; Vander Meer, Lentz, & Stollar, 2005; Wilson, 2005).

Furthermore, there is evidence for the validity of English ORF scores as indicators of L2 reading proficiency for English-language learners. Baker and Good (1995) examined the convergent construct validity by estimating the correlation between second-grade ELLs' English ORF scores and their scores on other reading measures, including the Stanford Diagnostic Reading Test, the Stanford Reading Comprehensive Subtest pre- and posttests, and teacher rating of reading skill. Correlation coefficients were .53, .73, .76, and .80, respectively. Likewise, Betts et al. (2006) investigated the correlations between scores on an ORF assessment and a standardized, norm-referenced reading test given around the same time to second-grade ELLs. The correlation between the spring ORF assessment and the Northwest Achievement Levels Test was .69, providing preliminary evidence of concurrent validity of English ORF scores for ELLs. Finally, several recent studies have yielded evidence for the predictive validity of ORF scores for ELLs' performance on end-of-year standardized reading exams or high-stakes state assessments (Betts et al., 2006; Muyskens et al., 2009; Wiley & Deno, 2005; Wilson, 2005). Coefficients ranged from .60 to .78 for linguistically diverse ELLs in second, third, and fifth grades.

## Data Collection

At the outset of Project ELLA, IRB approval was given for the entire research project, and parental consent was obtained for all student participants. Consent forms were provided in Spanish and English so that parents could give consent in the language in which they were most proficient. Teachers and paraprofessionals also consented to participation in the study. The present study using data archived from the completed project was also approved by the IRB.

Although students were assessed using a variety of language and literacy measures for the duration of Project ELLA, oral reading fluency assessments during second and third grades are of primary interest here. DIBELS ORF was given at the beginning (September), middle (January), and end (May) of Grades 2 and 3. At each measurement occasion, the assessment was administered by trained paraprofessionals or testers. Students were assessed individually in a quiet location in the school.

## Intervention

The ELLA intervention was a comprehensive intervention that was implemented over 4 years, from kindergarten through third grade. The intervention foci shifted over time in a developmentally appropriate progression of L2 emphases: oral language, basic reading skills, and then content-area reading skills. The intervention was implemented at two levels: Level I, the teacher level, and Level II, the student level.

Level I consisted of professional development for teachers and paraprofessionals. Teachers attended biweekly workshops for a total of 6 hr per month. Paraprofessionals attended 4 hr of professional development per month. Topics of professional development included: reviewing and practicing scripts for intervention lessons, reflecting on student learning, self-assessing their pedagogical growth resulting from participation as teachers in the intervention, and training in specific ESL instructional strategies. These strategies were incorporated into the scripted lessons and are described in Appendix A. In addition, teachers developed professional portfolios and reflected weekly on their practice (Brown & Irby, 2001).

The Level II student intervention was three-tiered. Tier 1 was the regular language arts, math, science, and social studies instruction in kindergarten through third grades. Tier 2 was a structured daily ESL intervention, which was delivered in a separate ESL block for 75 min in kindergarten and 90 min in first through third grades. Tier 3 was small-group supplemental instruction implemented for the lowest performing students. A family involvement component was implemented for enhanced practice students in kindergarten and first grade. The Tier 2 ESL intervention and Tier 3 small-group supplemental instruction are outlined in Figures 2 and 3, respectively, and detailed descriptions are given below.

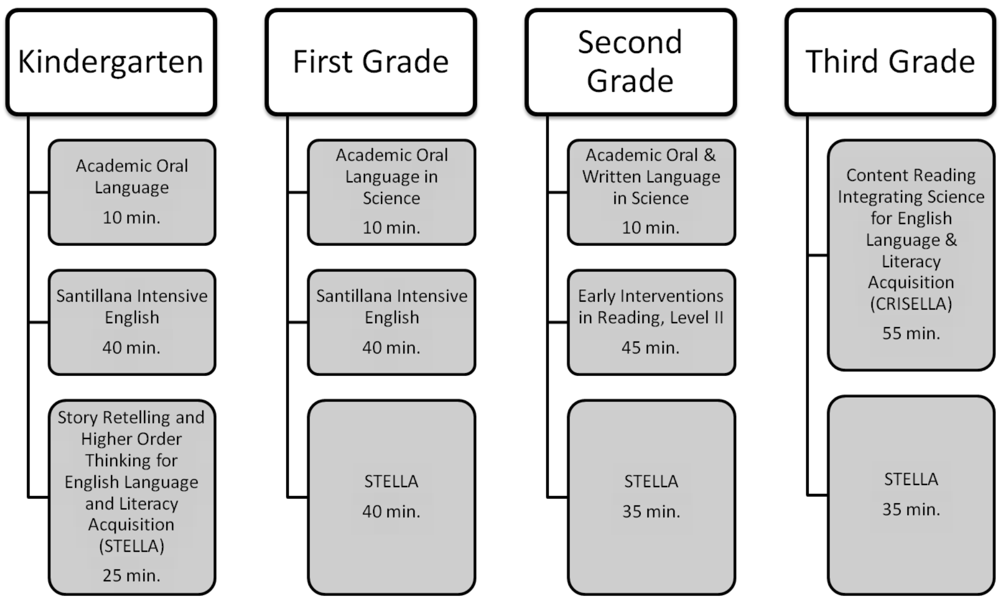


*Figure 2.* Overview of Tier 2 intervention curricula by grade.

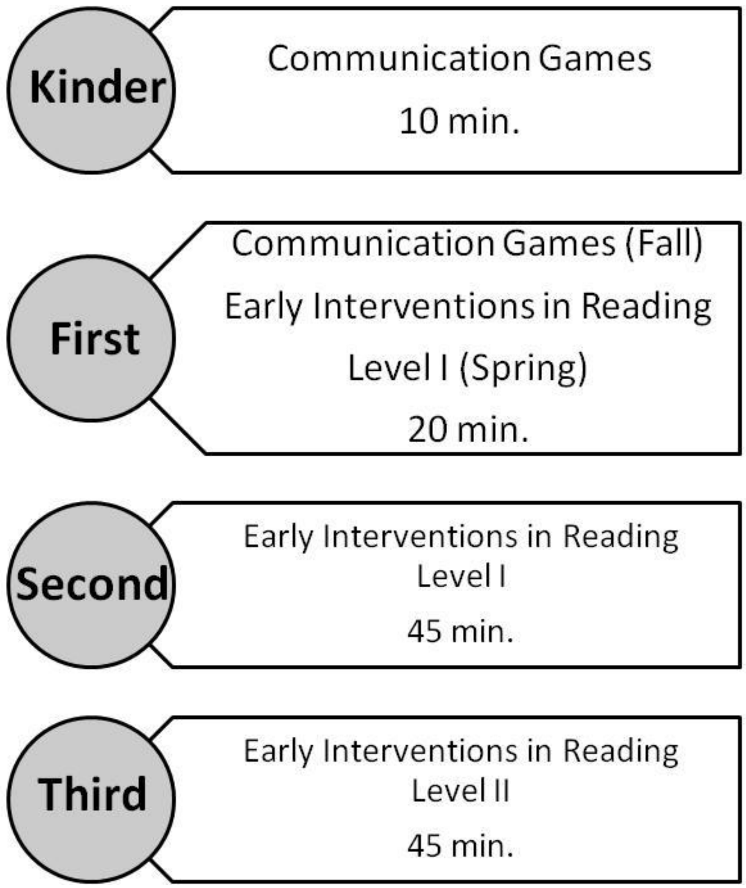


*Figure 3.* Overview of Tier 3 intervention curricula by grade.

**Tier 1**

Tier 1 of the intervention was the content area instruction provided to students in a developmental, one-way dual language education program. Both Spanish and English were used for instruction, and the program followed a native-language maintenance philosophy. The Spanish/English language distribution was 70/30 in kindergarten, 60/40 in first grade, 50/50 in second grade, and 40/60 in third grade. In kindergarten and first grade, all content areas—Spanish language arts, math, science, and social studies—were taught in Spanish. The block of ESL instruction provided in the intervention was the primary English instruction during kindergarten and first grades. Beginning in the second semester of second grade, math was taught in English. In third grade, English science instruction was added through the CRISELLA intervention component described below. All content instruction was aligned to the Texas state standards for each curriculum area.

**Tier 2**

The block of ESL instruction involved multiple components, which shifted in emphasis from year to year according to second language literacy development patterns. Oral language was emphasized in kindergarten and first grades, reading skills were emphasized in second grade, and content-area reading was emphasized in third grade. Figure 2 shows the curriculum components for each grade. In kindergarten, the 75-min ESL block consisted of three components: *Santillana Intensive English* (Ventriglia & González, 2000) for 40 min, Story Retelling and Higher Order Thinking for English Language and Literacy Acquisition (STELLA; Irby, Quiros, Lara-Alecio, Rodriguez, & Mathes, 2008) for 25 min, and an Academic Oral Language (AOL) activity for 10 min. In first grade, the ESL block time was increased to 90 min. The same three instructional components were used with minor adjustments. The AOL activity was modified to integrate science content (Academic Oral Language in Science, AOLS), and an additional 15 min was allotted to STELLA instruction (40 min). In second grade, the ESL block remained 90 min long. STELLA continued to be an integral instructional component (35 min), but Santillana was replaced by *SRA Early Interventions in*

*Reading, Level II* (Mathes & Torgesen, 2005b) for 45 min. The 10-min oracy component (AOLS) was adapted to Academic Oral and Written Language in Science (AOWLS), which elicited students' writing in addition to oral language. In third grade, the final year of intervention, the curriculum consisted of two components: 55 min of Content Reading Integrating Science for English Language and Literacy Acquisition (CRISELLA; Irby, Lara-Alecio, Mathes, Rodriguez, & Guerrero, 2008), an adaptation of Scott Foresman's (Cooney, 2006) third-grade science text, and 35 min of STELLA instruction. Each curriculum component is described below.

**Santillana Intensive English.** *Santillana Intensive English* is a research-based, systematic curriculum for English language instruction. Lessons use math, science, and social studies topics to provide instruction in phonemic awareness, phonics, vocabulary development, reading fluency, and reading comprehension. Topics were presented using lesson (picture) cards from the curriculum. Students listened to a topic-related story, answered leveled comprehension questions, and practiced new vocabulary words with the teacher. Other activities included role-playing conversations in pairs and small groups and working in the Santillana workbooks individual and with partners. Teachers used a 4-day sequence for each topic. Day 5 of each week was a make-up day, when teachers reviewed concepts that students had not mastered during the week, or an extension day, when the teacher led extension activities from the lesson cards. Santillana was used for 40 min in kindergarten and first grade.

**STELLA.** STELLA, developed by Irby, Quiros, Lara-Alecio, Rodriguez, and Mathes (2008), was a shared reading intervention implemented during all 4 years of the project. STELLA lessons used children's literature (one book per week) to expand L2 vocabulary knowledge, develop L2 listening and reading comprehension, and provide opportunities for critical thinking. The scripted lessons featured explicit vocabulary instruction (3 words/week in kindergarten, 6-8 words in first grade, 10-12 words in second grade, 10-14 words in third grade) and the use of preselected ESL strategies to teach language (1-2 strategies in kindergarten, 4-6 strategies in first grade, 4-7 strategies in second grade, and 6-10 strategies in third grade). Science concepts were integrated in STELLA for first, second, and third grades. Each lesson also included leveled questions, each question identified by its level according to Bloom's taxonomy, and teachers used these questions to engage students orally. Additionally, in kindergarten and first grade, interactive read-aloud strategies were used. In second and third grade, students participated in choral reading for fluency practice. Kindergarten STELLA lessons were 25 min in length, first grade STELLA lessons were 40 min, and second and third grade lessons were 35 min. For additional description of STELLA lessons, the interested reader is referred to Irby, Quiros, et al. (2008) and Tong, Lara-Alecio, Irby, Mathes, and Kwok (2008).

**Academic Oral Language (AOL, AOLS, AOWLS).** The Academic Oral Language (Irby, Lara-Alecio, Tong, Rodriguez, & Guerrero, 2009) component was used in kindergarten, first, and second grades (10 min), although its format was adjusted each year to incorporate science content and, later, writing. In kindergarten, the activity was structured as a daily oral language question, modeled after the *Question of the Day* materials from Lakeshore Learning (1997). Question topics were aligned with the

Santillana Intensive English lesson themes. Students were asked a question and given 3-4 possible answer choices. Teachers used a pocket chart to display cards with the questions and possible choices. Each student took a turn answering the question of the day, selecting one of the answer choices and using it in a complete sentence. Student name cards were placed on the pocket chart underneath their chosen answer choice, which created a graph that the teacher then used to elicit comparisons or generalizations about the data from the students. In first grade, Academic Oral Language in Science was created based on the "Question of the Day" format. At the request of the school district, science concepts aligned to district and state curriculum standards were incorporated into the oral language activity. In second grade, the oral language component was modified to Academic Oral and Written Language in Science by incorporating writing activities (Trevino et al., 2007). The research team created mini-lessons that used science-related visuals to elicit students' oral language and writing related to science concepts.

**Early Interventions in Reading, Level II.** In second grade, the Project ELLA intervention shifted in focus from oral language development to reading skill development. *Santillana Intensive English* was replaced by *SRA Early Interventions in Reading, Level II* (*EIR Level II*; Mathes & Torgesen, 2005b) for 45 min daily. *EIR Level II* was designed according to principles of direct instruction (Carnine, Silbert, & Kame'enui, 2004), and its 120 lessons integrated five strands of reading (NICHD, 2000): phonemic awareness, letter-sound correspondence, word recognition and spelling, fluency, and comprehension. The phonemic awareness strand included activities in phoneme discrimination, segmentation, and blending. In the letter-sound correspondence strand, a new letter-sound was introduced every 2-3 days. The word recognition strand taught word recognition strategies using word lists and included practice with phonetically regular and irregular high-frequency words. For the fluency strand, students practiced word recognition strategies by reading decodable text. Finally, the comprehension strand engaged students in both pre-reading activities (e.g., previewing the story, predicting, setting a purpose for reading, and activating prior knowledge) and post-reading activities (e.g., story grammar, summarizing, graphic organizers, main idea and story details, and making inferences) related to the text. *EIR Level II* was designed for small-group instruction; however, Project ELLA teachers taught the lessons in whole-class instruction. One of *EIR*'s developers, who was also a Project ELLA researcher, provided guidance in adapting the program for whole-class instruction.

**CRISELLA.** In third grade, the intervention emphasized reading and academic language in the content areas, specifically in science. *EIR Level II* was replaced by Content Reading Integrating Science for English Language and Literacy Acquisition (CRISELLA; Irby, Lara-Alecio, et al., 2008) for 55 min. CRISELLA was a scripted enhancement of Scott Foresman's (Cooney, 2006) third-grade science textbook . Using the textbook as a basis, researchers developed lessons that integrated reading skills and expository text strategies for English-language learners. For each chapter, teachers received scripted lesson plans that included pre-reading strategies, vocabulary development activities, partner reading, graphic organizers, hands-on inquiry activities, cooperative grouping, scaffolded questions, vocabulary extensions, fluency practice, and direct teaching of reading skills. Among the reading skills taught were: sequencing, comparing and contrasting, drawing conclusions, identifying main idea and details, identifying cause and effect, and making inferences.

**Parental involvement.** In kindergarten and first grade, a parental education component was implemented for parents of students in the intervention classrooms. In addition, the project provided bilingual, take-home activity books and/or learning activities that were aligned to the intervention curriculum.

### Tier 3

Tier 3 intervention was provided each year for the lowest performing students. Students were pulled from ESL block for daily small-group instruction (groups of 3-5 students) delivered by highly trained paraprofessionals. Students were selected for Tier 3 instruction based on low performance on DIBELS (Good & Kaminski, 2002a) assessments given at the beginning of each year. In kindergarten, students participated in communication games for 10 min. The communication games were researcher-developed and focused on vocabulary development, phonemic awareness (e.g., segmenting and blending words, identifying initial and word-final sounds), and English listening and speaking. In first grade, students received 20 min of small-group instruction. In the first semester, they continued participating in communication games. In the second semester, communication games were replaced by *SRA Early Interventions in Reading, Level I* (*EIR Level I*; Mathes & Torgesen, 2005a), a more aggressive reading intervention that developed phonemic awareness, reading fluency, and comprehension. *EIR Level I* was continued throughout second grade for 45 min daily. In third grade, *SRA*

*Early Interventions in Reading, Level II* was the curriculum for the small-group instruction (45 min).

**Intervention Fidelity**

To ensure that teachers appropriately and accurately delivered the ELLA intervention, Project ELLA coordinators conducted classroom observations three times per year (beginning, middle, and end). The coordinators were trained by the project investigators to use a 4-point, Likert-type rating scale to assess teachers' fidelity to intervention lessons. Observers obtained strong interrater reliability of .98 when applying the rating scale to videotaped lessons of intervention teachers implementing the curriculum. The 4-point Likert rating scale assessed the following five areas: (a) knowledge of the lesson content and script, (b) management of instructional materials, (c) student involvement, (d) leveled questioning, and (e) classroom management. Out of 96 possible points, the overall mean score was 83.85 (SD=12.13). The average observation time for each teacher through the 4-year project was 73 min annually.

In addition, observations of intervention and control classrooms were conducted using the Transitional Bilingual Observation Protocol (TBOP; Lara-Alecio & Parker, 1994). The TBOP was used to document pedagogical language use (e.g., language of instruction, language of students, language modalities, light and dense cognitive content, and activity structures). These observations were conducted four times yearly (September, November, February, April), and interrater reliability of protocol scores was .97.

## Control Condition

Students in classrooms assigned to the control condition participated in a late-exit, transitional bilingual education program as typically provided in the district. In kindergarten, 80% of instruction was delivered in Spanish and 20% in English (80/20 model). Through each successive academic year, instruction gradually transitioned until a 50/50 language distribution was reached in third grade. ESL instruction was provided in a 45-min block with no support from the research team. All content areas were taught in Spanish in kindergarten and first grade. Formal English reading instruction began in the second semester of first grade for students who had passing grades in Spanish language arts. In second grade, English language arts instruction was added for all students, and English math instruction began in the second semester. Instruction was aligned to Texas curricular standards. The Houghton-Mifflin Spanish language arts series and the Spanish editions of Harcourt science and social studies were utilized in the control classrooms. Teachers held state certification for bilingual education.

In addition to content instruction, daily ESL instruction was provided in a 45-min block with no support from the research team. In order to document typical instructional practices, the research team conducted observations, which indicated that instruction was aligned to district benchmarks and state ESL standards and that ESL instruction varied widely among classrooms. In particular, variation was noted in (a) curricula used (scripted direct instruction in some schools, district-adopted ESL or English language arts curricula in others), (b) methods of vocabulary instruction (included practices such as incidental teaching of words, flash cards, activity pages), and (c) methods of reading skill development (included oral reading of sentences, decoding practice, counting and reading syllables). Often, the prescribed 45-min ESL block was not fully implemented because other tasks (e.g., restroom breaks, lessons in other curriculum areas) intruded into ESL time. A typical lesson included a focus, instruction, guided practice, and independent practice. In addition, teachers were observed to use code-switching to make clarifications during ESL time. As in the enhanced practice group, struggling students in the control condition received supplemental tutorials; however, the content of supplemental instruction was left to each individual teacher's discretion.

## Data Analysis

Data were analyzed using multilevel models. Educational data frequently contain dependence associated with the levels of clustering (repeated measures within students, students within classrooms or schools). Ignoring this dependence can inflate the Type I error rate. Also called hierarchical linear models, multilevel models account for dependence by partitioning the variance in the outcome according to the levels of clustering (Hox, 2002). Predictor variables can be entered into the model at any level of the model. Thus, the researcher can examine effects at different levels—between schools, between students within schools, or repeated measures within students—simultaneously (Beretvas, 2004; Hox, 2002). In this study, the clustering structure is as follows: ORF measurements clustered within students who were clustered within schools.

In addition, multilevel models flexibly handle unbalanced designs. This is particularly advantageous for longitudinal studies where the number of measurement occasions often varies across participants due to attrition. In the present longitudinal sample, only 76% of students were assessed in ORF at all six measurement occasions. One method of handling missing data is listwise deletion, deleting the cases without complete measurements. Deleting cases causes loss of information that potentially biases results toward students who have complete data. Multilevel modeling can easily incorporate all available measurements without deletion of cases, which results in more accurate model parameters.

**Piecewise Modeling**

In this study, multilevel linear growth models were specified to estimate students' growth in ORF over second and third grades. Specifically, a piecewise linear growth model was fit to the data. In linear growth models, change or development over time is estimated in a linear regression equation. Piecewise modeling separates the growth rate into multiple phases with each phase having a different rate, which allows for the testing of multiple phases of growth within one model. In the present study, two phases of growth were modeled.

Modeling two distinct phases of growth is appropriate for three reasons. First, previous research findings with native English speakers have indicated that reading fluency growth rates are positive and linear within a single year but decelerate each year after first grade (Baker et al., 2008; Chard et al., 2008; Crowe et al., 2009; Fuchs et al., 1993; Kim et al., 2010; Speece & Ritchey, 2005; Stage & Jacobsen, 2001; Stage et al., 2001). There is some evidence that this occurs with ELLs as well (Al Otaiba et al., 2009; Dominguez de Ramirez & Shapiro, 2006). Thus, I expected ORF growth to be slower in third grade compared to second grade. Second, the ESL intervention emphasized different aspects of L2 literacy in each grade level. The second-grade curriculum focused on basic reading skills, whereas the third-grade curriculum emphasized content-area reading skills. Given the change in focus, it is reasonable to expect changes in ORF growth rate. Finally, the two-phase growth model accounts for discontinuities in growth rate associated with measurement. The DIBELS ORF passages used in this study are equivalent probes within grades; however, the passage difficulty increases across grades (i.e., Third-grade passages are slightly more difficult than second-grade passages). Thus, students' growth is discontinuous from second to third grade. Modeling separate rates in second and third grades accounts for this discontinuity.

In addition to differences in growth rate, I expected a change in level, or shift in intercept, between second and third grades. This study included summer vacation between second and third grades. Typically, students regress in skill during summer vacation. Figure 4 contains empirical growth plots for 12 students selected at random. These plots clearly illustrate the discontinuity between the third and fourth measurement before and after summer vacation. I added an intercept shift for beginning of third grade in the piecewise growth model to account for a decrement in reading fluency skill during the summer months.

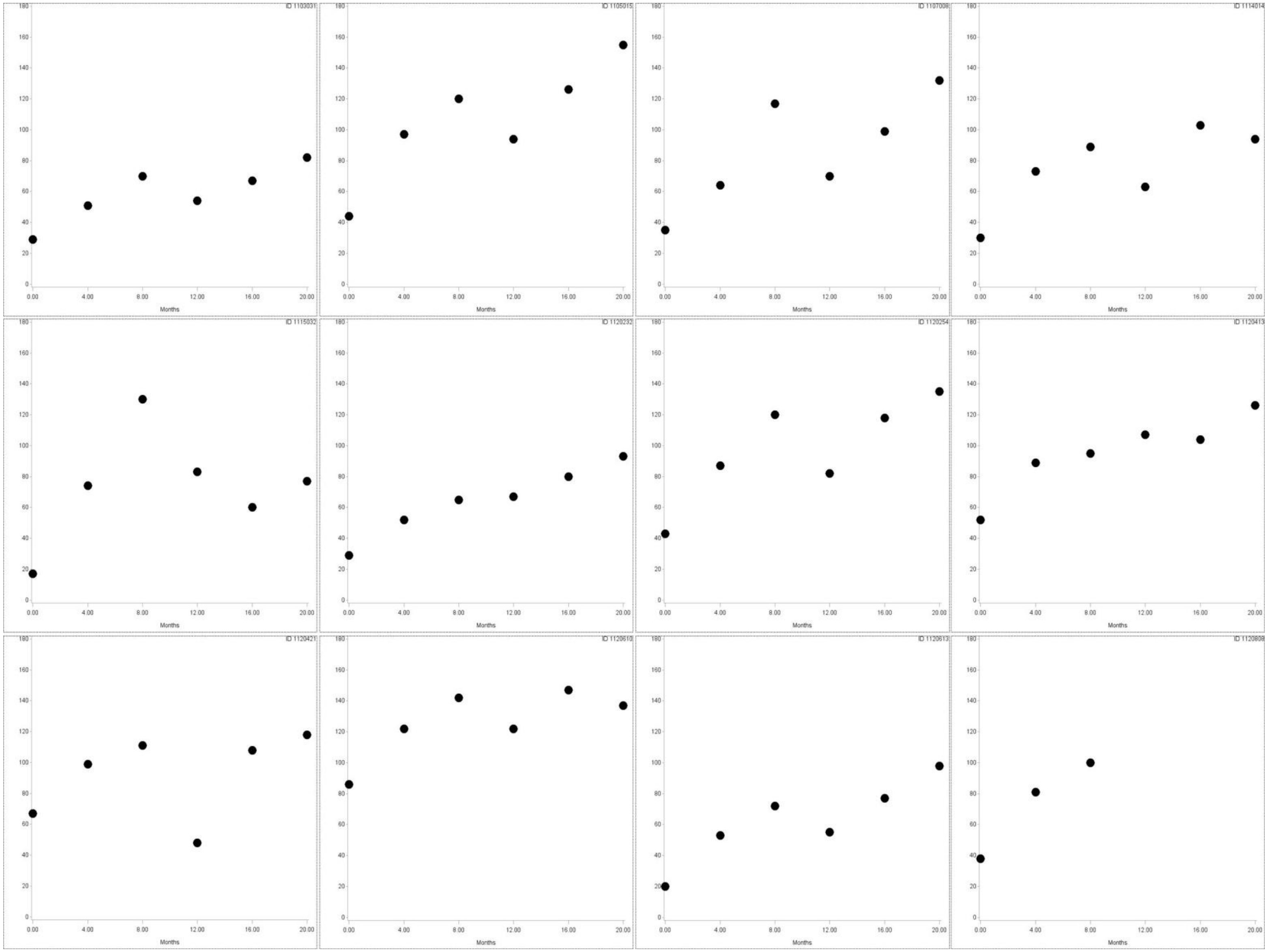


*Figure 4.* Empirical growth plots for 12 students selected at random.

## Model Specification

For these data, repeated measures were nested in students, who were nested in second- and third-grade classrooms, which were then nested in schools. In order to assess the amount of variance at each clustering level, two preliminary unconditional three-level models were hypothesized and tested. The first three-level model contained repeated measures at Level 1, students at Level 2, and school at Level 3. The second three-level model was a cross-classified model, with students cross-classified in second- and third-grade classrooms at Level 3. Model parameters and fit statistics for these preliminary models are listed in Appendix B. There was no statistically significant variance at Level 3 for either of the two models. Thus, the third level was dropped from subsequent analyses.

Data analysis proceeded with two-level models, repeated measures at Level 1 nested in students at Level 2. Following Singer & Willett's (2003) recommendations for analyzing longitudinal data, three 2-level models were fit to the data.

**Model A.** First, a two-level unconditional means model was fit. An unconditional means model partitions the total variance in the outcome into variance components for each level of the model. The model equations were:

Level 1: $ORF_{ij} = \pi_{0i} + e_{ij}$

Level 2: $\pi_{0i} = \beta_{00} + r_{0i}$

Combined: $ORF_{ij} = \beta_{00} + r_{0i} + e_{ij}$,

where $ORF_{ij}$ is the oral reading fluency score for the *i*th student on occasion *j*, $\beta_{00}$ is the grand mean ORF score across students and measurement occasions, $\pi_{0i}$ is the person-specific mean ORF for student *i*, $r_{0i}$ represents the between-student deviation between $\pi_{0i}$ (person-specific mean) and $\beta_{00}$ (the grand mean) for student *i*, and $e_{ij}$ represents the within-student deviation between $\pi_{0i}$ and $ORF_{ij}$. Results for this preliminary model indicated that there was statistically significant variance within (Level 1) and between (Level 2) students that could be explained in subsequent multilevel models.

**Model B.** Next, an unconditional piecewise growth model was fit to the data. Specifying the piecewise linear growth model required recoding data into new variables. Table 7 shows the data and recoded time variables for 1 student in the sample. Time was recoded in months (the Month variable) to facilitate the interpretation of slope coefficients. Data were centered at the beginning of second grade (Month = 0). The Month variable was recoded to create the pieces of the growth curve and the shift in intercept (discontinuity) at the beginning of third grade. A variable (Gr2Slope) was created to represent the second-grade slope and coded as 0, 4, 8, 8, 8, 8. A second variable (Summer) represented the summer decrement, or intercept shift at beginning of third grade, and was coded as 0, 0, 0, 1, 1, 1. A third variable (Gr3Slope) represented the third-grade slope and was coded as 0, 0, 0, 4, 8, 12.

Table 7
*ORF Data and Recoded Time Variables for 1 Student*

| Student | ORF | Time | Month | Gr2Slope | Summer | Gr3Slope |
|---|---|---|---|---|---|---|
| 1067 | 30 | 1 | 0 | 0 | 0 | 0 |
| 1067 | 68 | 2 | 4 | 4 | 0 | 0 |
| 1067 | 83 | 3 | 8 | 8 | 0 | 0 |
| 1067 | 66 | 4 | 12 | 8 | 1 | 4 |
| 1067 | 88 | 5 | 16 | 8 | 1 | 8 |
| 1067 | 108 | 6 | 20 | 8 | 1 | 12 |

The unconditional piecewise growth model equations were as follows:

Level 1: $ORF_{ij} = \pi_{0i} + \pi_{1i}Gr2Slope_{ij} + \pi_{2i}Summer_{ij} + \pi_{3i}Gr3Slope_{ij} + e_{ij}$

Level 2: $\pi_{0i} = \beta_{00} + r_{0i}$

$\pi_{1i} = \beta_{10} + r_{1i}$

$\pi_{2i} = \beta_{20} + r_{2i}$

$\pi_{3i} = \beta_{30} + r_{3i}$

Combined: $ORF_{ij} = \beta_{00} + \beta_{10}Gr2Slope_{ij} + \beta_{20}Summer_{ij} + \beta_{30}Gr3Slope_{ij} + r_{0i} +$

$r_{1i}Gr2Slope_{ij} + r_{2i}Summer_{ij} + r_{3i}Gr3Slope_{ij} + e_{ij}$,

where $ORF_{ij}$ is the oral reading fluency score for the *i*th student on occasion *j*; $\pi_{0i}$ is the intercept, or expected ORF at beginning of second grade for student *i*; $\pi_{1i}$ is the expected monthly growth rate for second grade; $\pi_{2i}$ is the expected change in ORF level when students returned from summer vacation (i.e., at the beginning of third grade); $\pi_{3i}$ is the expected monthly growth rate for third grade, and $e_{ij}$ is the within-student, random variation at occasion *j*. In the Level 2 and combined models, $\beta_{00}$ is the estimated intercept, the mean ORF score at the beginning of second grade; $\beta_{10}$ is the estimated mean monthly growth rate across students in second grade; $\beta_{20}$ is the estimated mean change in ORF level after summer vacation; and $\beta_{30}$ is the estimated mean monthly growth rate in third grade. The remaining terms represent between-student random variation: $r_{0i}$ is the random effect for intercept, $r_{1i}$ is the random effect for second-grade growth rate, $r_{2i}$ is the random effect for shift in ORF level after summer vacation, and $r_{3i}$ is the random effect for third-grade growth rate.

**Model C.** After the piecewise growth model was fit, the effect of intervention was added as a Level-2 predictor. The Intervention variable was coded as 1 for students who received instruction in the Project ELLA intervention and 0 for students who were members of the control group. Intervention was added as a predictor for three of the four growth parameters. Based on results from the unconditional growth model, Intervention was not entered as a predictor of $\pi_{3i}$, the third-grade slope parameter, nor was $\pi_{3i}$ allowed to vary randomly (i.e., $\pi_{3i}$ was a fixed effect with no predictors). Model equations were as follows:

Level 1: $\mathrm{ORF}_{ij} = \pi_{0i} + \pi_{1i}\mathrm{Gr2Slope}_{ij} + \pi_{2i}\mathrm{Summer}_{ij} + \pi_{3i}\mathrm{Gr3Slope}_{ij} + e_{ij}$

Level 2: $\pi_{0i} = \beta_{00} + \beta_{01}\mathrm{Intervention} + r_{0i}$

$\pi_{1i} = \beta_{10} + \beta_{11}\mathrm{Intervention} + r_{1i}$

$\pi_{2i} = \beta_{20} + \beta_{21}\mathrm{Intervention} + r_{2i}$

$\pi_{3i} = \beta_{30}$

Combined: $ORF_{ij} = \beta_{00} + \beta_{01}\text{Intervention} + \beta_{10}\text{Gr2Slope}_{ij} + \beta_{11}\text{InterventionGr2Slope}_{ij} + \beta_{20}\text{Summer}_{ij} + \beta_{21}\text{InterventionSummer}_{ij} + \beta_{30}\text{Gr3Slope}_{ij} + r_{0i} + r_{1i}\text{Gr2Slope}_{ij} + r_{2i}\text{Summer}_{ij} + e_{ij}$,

where $\beta_{01}$ is the difference in ORF between intervention and control groups at the beginning of second grade, $\beta_{11}$ is the difference in ORF growth between intervention and control groups during second grade, and $\beta_{21}$ is the difference in summer decrement between intervention and control groups.

**Error covariance structure.** The Level-1 error covariance structure was examined. The standard multilevel model for change assumes that the within-subject (Level-1) residuals are "independently and identically distributed with mean zero and homogeneous variance $\sigma^2$ for all participants" (Kwok, West, & Green, 2007, p. 558). However, longitudinal data are expected to be serially dependent (Singer & Willett, 2003). Within an individual, later observations are likely to be correlated with earlier observations. Therefore, the within-subject covariance structure of the standard multilevel model—the identity structure of $\sigma^2\mathbf{I}$, which assumes serial independence and homogeneous residual variance at each measurement occasion—may not be appropriate for longitudinal data (Kwok et al., 2007). Correctly specifying the covariance structure yields more precise estimates of fixed effects, which affect the results of hypothesis testing and the width of confidence intervals (Singer & Willett, 2003). Singer and Willett (2003) recommended testing alternative covariance structures to determine which covariance structure best fits the longitudinal data.

Model C ("final" model) was tested using four error covariance structures that are common in longitudinal data: first-order autoregressive, AR(1); second-banded Toeplitz, TOEP(2); first-order autoregressive moving average, ARMA(1,1); and banded main diagonal (heterogeneous residual variances), UN(1) (Kwok et al., 2007). Using the goodness-of-fit indices described in the next section, UN(1) was selected as the covariance structure for these data.

## Evaluation of Model Integrity

**Fit indices.** Model fit was assessed using goodness-of-fit indices: the deviance statistic, Akaike Information Critierion (AIC; Akaike, 1973), and Bayesian Information Criterion (BIC; Schwartz, 1978). The deviance statistic compares the log-likelihood statistic for the current model against the log-likelihood statistic for the saturated (perfectly fit) model and "quantifies *how much worse* the current model is in comparison to the best possible model" (Singer & Willett, 2003, p. 117); therefore, smaller values are preferred for deviance statistics. For two nested models fit to identical data sets, model fit was compared using a likelihood ratio test, which tests the difference between the two models' deviance statistics against a $\chi^2$ distribution. Models that were not nested were compared using AIC and BIC. Like the deviance statistic, both AIC and BIC are based on the log-likelihood statistic and smaller values are preferred, but each exacts penalties based on the model structure. AIC penalizes based on the number of parameters in the model, whereas BIC penalizes based on the number of parameters and sample size.

**Assumptions about errors.** Model residuals were evaluated to confirm that they were normally distributed and homoscedastic. To evaluate the tenability of the normality assumption, normal probability plots were examined for level-1 and level-2 raw residuals. Plots of standardized residuals were also examined. The assumption of homoscedasticity for level-1 within-student residual variances was relaxed via the UN(1) error covariance structure, which assumed heterogeneous residual variances at each time point. Homoscedasticity for level-2 residuals was evaluated by plotting level-2 raw residuals against predictors and visually inspecting the residual variability at each value of the predictor.

**Sensitivity of parameter estimates to outliers.** To identify outliers, plots of standardized residuals by ID number were examined. Students who were ±3 standard deviations from the mean residual (0) were identified as outliers. From the level-2 standardized residual plots, 5 students (1 from control group, 4 from experimental group) were identified as outliers at Level 2. From the level-1 standardized residual plots, 13 individual scores (8 from students in control group, 6 from students in experimental group) were identified as outliers at Level 1. Outliers were deleted from the data set, and the models were fit again to verify that parameter estimates were not unduly influenced by outlier students or scores.

## Estimation and Inference

**Model estimation.** All models were estimated in SAS 9.3 (SAS Institute, 2011) using the Newton-Raphson algorithm for full maximum likelihood (FML) estimation. Under FML estimation, the likelihood of the sample data is maximized, and thus the deviance statistic quantifies the fit of the entire model, both the structural portion (fixed effects) and the stochastic portion (random effects; Singer & Willet, 2003). In contrast, restricted maximum likelihood (RML) estimation maximizes the likelihood of the sample residuals, and the deviance statistic characterizes the fit of the stochastic part of the model only (Singer & Willett, 2003). In this study, the structural portion (fixed effects) was the primary interest, and therefore FML was used.

**Inferences about variance components and fixed effects.** Hypothesis tests for variance components used simple *z*-tests, which were conducted by dividing an estimate by its standard error. For fixed effects, hypothesis tests were conducted using *t*-tests with degrees of freedom calculated using the between-within method. Confidence intervals (95%) were constructed around fixed effect estimates by adding and subtracting 1.96 times the standard error.

## Summary

In this chapter, I outlined the methodology used in this dissertation. First, I described the study design, context, and participants. Next, I detailed the instrumentation, data collection, and intervention curricula. Finally, I explained the data analysis methods for multilevel modeling. In Chapter IV, I present the results of the study.

# CHAPTER IV

# PRESENTATION AND ANALYSIS OF DATA

The purposes of this study were to model the developmental trajectory of L2 oral reading fluency for Spanish-speaking English-language learners from Grade 2 through Grade 3 and to determine the effect of a comprehensive instructional intervention, which emphasized oral language and reading skills in L2, on ELLs' oral reading fluency. These aims were achieved by building a series of multilevel growth models for L2 reading fluency during second and third grades. In Chapter IV, I report the results of these analyses. First, I present the results of descriptive and preliminary analyses. Preliminary results are followed by the primary results addressing the two research questions, which were:

1. What are the average initial level and rate of growth in English oral reading fluency for Spanish-speaking ELLs in Grades 2 and 3?

2. What is the effect of Project ELLA on Spanish-speaking ELLs' initial level and growth in English oral reading fluency in Grades 2 and 3?

## Preliminary Results

### Descriptive Statistics

Univariate descriptive statistics are presented to show the central tendency, dispersion, and shape of the oral reading fluency data analyzed in this study. Table 8 contains the mean, standard deviation, skewness, and kurtosis statistics for ORF at each measurement occasion in second and third grades and for all 1,470 ORF observations.

Table 9 provides means and standard deviations by intervention and control groups. The correlations among ORF at each measurement occasion are presented in Table 10.

Table 8
*Descriptive Statistics for Oral Reading Fluency*

| | *n* | Portion of Sample Assessed | *M* | *SD* | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| ORF, Grade 2 | | | | | | |
| Fall | 279 | 98.6% | 44.33 | 24.74 | 0.69 | 0.06 |
| Winter | 263 | 92.9% | 72.44 | 29.80 | 0.22 | 0.07 |
| Spring | 257 | 90.8% | 89.07 | 31.95 | 0.02 | -0.25 |
| ORF, Grade 3 | | | | | | |
| Fall | 227 | 80.2% | 75.06 | 27.01 | 0.39 | 0.03 |
| Winter | 223 | 78.8% | 92.96 | 30.61 | 0.24 | 0.10 |
| Spring | 222 | 78.4% | 104.15 | 29.55 | 0.02 | 0.30 |
| ORF, all observations [a] | 1470 | | 78.30 | 34.90 | 0.18 | -0.35 |

[a] Descriptive statistics for all ORF observations in the person-period data set.

Table 9
*Descriptive Statistics by Experimental Condition*

| | Intervention (ELLA) | | | Control | | |
|---|---|---|---|---|---|---|
| | *n* | *M* | *SD* | *n* | *M* | *SD* |
| ORF, Grade 2 | | | | | | |
| Fall | 147 | 42.52 | 22.14 | 132 | 46.34 | 27.29 |
| Winter | 139 | 75.33 | 27.08 | 124 | 69.20 | 32.39 |
| Spring | 138 | 92.75 | 28.97 | 119 | 84.79 | 34.73 |
| ORF, Grade 3 | | | | | | |
| Fall | 117 | 76.10 | 24.03 | 110 | 73.95 | 29.94 |
| Winter | 116 | 96.03 | 26.65 | 107 | 89.63 | 34.21 |
| Spring | 116 | 109.08 | 25.91 | 106 | 98.76 | 32.35 |

Table 10
*Correlations among ORF Measurements at Each Occasion*

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| ORF, Grade 2 | | | | | | |
| 1. Fall | -- | | | | | |
| 2. Winter | .81*** | -- | | | | |
| 3. Spring | .76*** | .85*** | -- | | | |
| ORF, Grade 3 | | | | | | |
| 4. Fall | .80*** | .80*** | .82*** | -- | | |
| 5. Winter | .79*** | .80*** | .83*** | .89*** | -- | |
| 6. Spring | .71*** | .75*** | .80*** | .82*** | .84*** | -- |

*** $p < .001$.

Because this study focused on *initial status* and *changes* in ORF during second and third grades, the intercept and slope parameters for ORF are the technical outcome variables for the study (Ferron et al., 2008). Therefore, trends for intercept and slope should also be described. Ordinary least squares (OLS) regression trajectories of ORF for each student provide descriptive information about the trends for intercept and slope among the sample. Figure 5 contains the OLS regressions of ORF on time for 28 students (10%) selected randomly from the sample. OLS intercepts clustered between 35 and 70 wcpm at the beginning of Grade 2. The mean OLS slope was 10.34 wcpm (*SD*=6.41) per measurement occasion. Table 11 provides the $r^2$ values for the 28 OLS regressions. The mean $r^2$ value was .63 (*SD*=.25), indicating that individual OLS regression models explained approximately 63% of the within-student temporal variance, on average.

**Model A, Unconditional Means Model**

Singer and Willet (2003) recommended estimating the unconditional means model as a preliminary analysis. The unconditional means model estimates the mean level in the outcome across all time points and partitions the outcome variability for each level of the model. At Level 1, the within-student variance ($e_{ij}$) is the temporal variation within an individual student, or dispersion among an individual student's scores over time around his/her individual mean. At Level 2, the between-student variance ($r_{0j}$) is the dispersion of individual student means around the grand mean ($\beta_{00}$) for the sample.

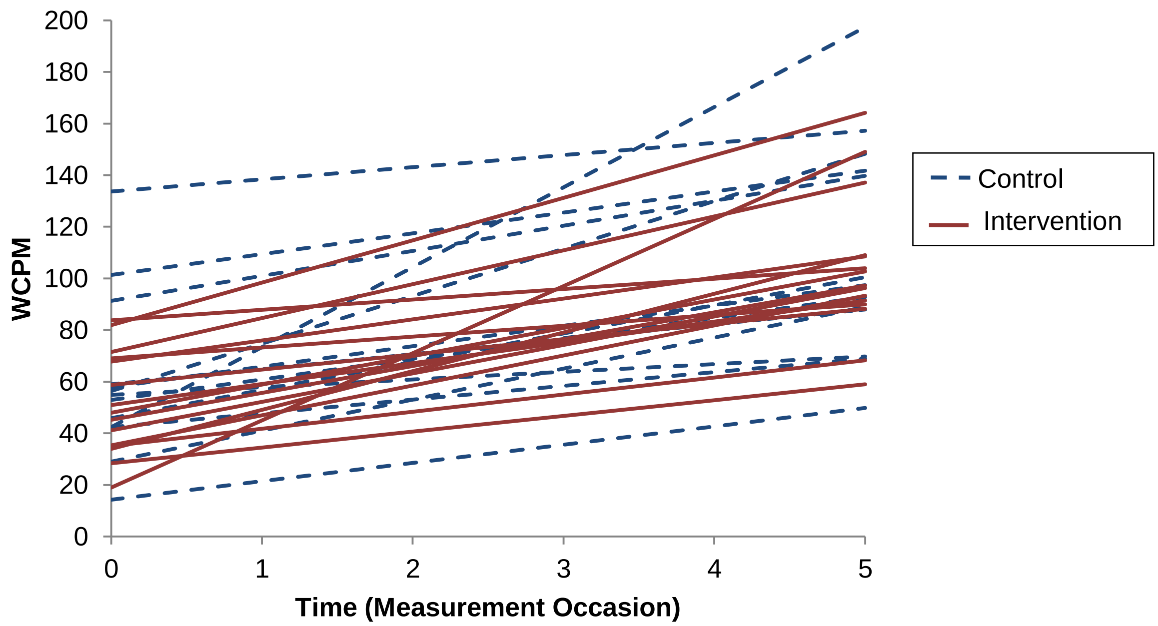


*Figure 5*. Fitted OLS regressions of oral reading fluency on time for a random subsample of 28 students (10% of sample). WCPM=words read correctly in 1 min. Time: 0-2=Fall,Winter, and Spring of Grade 2, respectively; 3-5=Fall, Winter, and Spring of Grade 3, respectively.

Table 11

*Stem-and-leaf Plot of $r^2$ Values for Fitted OLS Regressions of Oral Reading Fluency on Time for a Randomly Selected Subsample (n=28)*

| Stem | Leaves |
|---|---|
| .9 | 98651 |
| .8 | 75 |
| .7 | 87651 |
| .6 | 9320 |
| .5 | 84 |
| .4 | 966542 |
| .3 | 6 |
| .2 | 9 |
| .1 | 41 |

Model A (see Table 12) is the unconditional means model for these data. The mean ORF score ($\beta_{00}$) across all time points through the 2-year study was 75.92 wcpm. There was statistically significant variance ($p < .001$) in Levels 1 and 2 of the model, which indicated that further analyses exploring this variation (Models B and C) were warranted. The within-student variance accounted for 47% of the total variance, and the between-student variance accounted for 53%.

**Intraclass Correlation Coefficient**

The intraclass correlation coefficient (ICC) is the proportion of total variance in the outcome that is between-subject variation (Singer & Willett, 2003). The ICC also describes the amount of residual autocorrelation for the unconditional means model and, therefore, estimates the average correlation between composite residuals (errors) for any two observations within each person (Singer & Willett, 2003). The ICC was computed via the following equation:

$$\rho = \sigma^2_0 / (\sigma^2_0 + \sigma^2_\varepsilon),$$

where $\sigma^2_0$ is the between-student variance and $\sigma^2_\varepsilon$ is the within-student variance. The ICC for the unconditional means model was .53, indicating that 53% of the total variance in ORF was attributable to variation between students and that the estimated residual autocorrelation was large. The magnitude of the ICC provided support for analyzing these data with multilevel modeling rather than OLS regression, which would assume zero residual autocorrelation.

Table 12

*Parameter Estimates for Two-Level Growth Curve Model of L2 Oral Reading Fluency Development*

| | Model A | | Model B | | Model C | |
|---|---|---|---|---|---|---|
| | Estimate | 95% CI | Estimate | 95% CI | Estimate | 95% CI |
| **Fixed Effects** | | | | | | |
| Initial level, Grade 2, $\pi_{0i}$ | | | | | | |
| Intercept, $\beta_{00}$ | $75.92^{***}$ (1.67) | [72.65, 79.19] | $46.20^{***}$ (1.48) | [43.30, 49.10] | $46.80^{***}$ (2.14) | [42.61, 50.99] |
| Intervention, $\beta_{01}$ | | | | | -3.46 (2.93) | [-9.20, 2.28] |
| Rate of change, Grade 2, $\pi_{1i}$ | | | | | | |
| Intercept, $\beta_{10}$ | | | $5.49^{***}$ (0.16) | [5.18, 5.80] | $4.73^{***}$ (0.23) | [4.28, 5.18] |
| Intervention, $\beta_{11}$ | | | | | $1.52^{***}$ (0.31) | [0.91, 2.13] |
| Summer shift (change in level), $\pi_{2i}$ | | | | | | |
| Intercept, $\beta_{20}$ | | | $-30.47^{***}$ (1.51) | [-33.43, -27.51] | $-28.88^{***}$ (1.77) | [-32.35, -25.41] |
| Intervention, $\beta_{21}$ | | | | | -2.49 (2.05) | [-6.51, 1.53] |
| Rate of change, Grade 3, $\pi_{3i}$ | | | | | | |
| Intercept, $\beta_{30}$ | | | $3.65^{***}$ (0.14) | [3.38, 3.92] | $3.76^{***}$ (0.14) | [3.49, 4.03] |
| **Variance Components** | | | | | | |
| Level 1 (within-student) | | | | | | |
| Temporal variation, $e_{ij}$ | $580.26^{***}$ (23.86) | [533.49, 627.03] | $135.87^{***}$ (8.74) | [118.74, 153.00] | T0 $52.42^{*}$ (27.69) | [-1.85, 106.69] |
| | | | | | T1 $190.15^{***}$ (20.34) | [150.28, 230.02] |
| | | | | | T2 $75.46^{*}$ (33.50) | [9.80, 141.12] |
| | | | | | T3 $89.04^{***}$ (12.94) | [63.68, 114.40] |
| | | | | | T4 $125.41^{***}$ (15.89) | [94.27, 156.55] |
| | | | | | T5 $193.92^{***}$ (22.23) | [150.35, 237.49] |
| Level 2 (between-student) | | | | | | |
| Grade 2 initial level, $r_{0i}$ | $659.71^{***}$ (67.27) | [527.86, 791.56] | $505.40^{***}$ (52.71) | [402.09, 608.71] | $552.98^{***}$ (55.56) | [444.08, 661.88] |

Table 12 Continued

| | Model A | | Model B | | Model C | |
|---|---|---|---|---|---|---|
| | Estimate | 95% CI | Estimate | 95% CI | Estimate | 95% CI |
| Grade 2 rate of change, $r_{1i}$ | | | $2.52^{***}$ (0.66) | [1.23, 3.81] | $4.20^{***}$ (0.92) | [2.40, 6.00] |
| Summer shift, $r_{2i}$ | | | $105.25^{*}$ (56.87) | [-6.22, 216.72] | $141.49^{***}$ (36.38) | [70.19, 212.79] |
| Grade 3 rate of change, $r_{3i}$ | | | 0.41 (0.52) | [-0.61, 1.43] | | |
| **Fit Indices** | | | | | | |
| Deviance | 14057.80 | | 12494.90 | | 12448.20 | |
| AIC | 14063.80 | | 12524.90 | | 12486.20 | |
| BIC | 14074.80 | | 12579.50 | | 12555.40 | |

*Note.* Full ML, SAS Proc Mixed. All relevant covariance components, though not displayed above to conserve space, were estimated in each model. For Model C, the error covariance structure was UN(1), heterogeneous residual variances. T0…T5=Time 0 through Time 5, respectively.
* $p < .05$, ** $p < .01$, *** $p < .001$.

## Primary Results

### Research Question #1

The first research question was: *What are the average initial level and rate of growth in English oral reading fluency for Spanish-speaking ELLs in Grades 2 and 3?* This question was answered using Model B, an unconditional piecewise growth model. In Model B, parameters for ORF initial level and growth in Grades 2 and 3 were estimated for all students in the sample. Table 12 provides parameter estimates, standard errors, confidence intervals, and fit indices for Model B. Adding growth parameters in Model B accounted for 77% of the within-student variance and 7% of the between-student variance that was present in Model A. Large decreases in deviance ($\chi^2$ (12) = 32.91, $p$ < .001), AIC, and BIC indicated that the fit of Model B improved over Model A.

**Grade 2 initial level and growth.** On average, ELLs read 46.20 wcpm at the beginning of second grade ($\beta_{00}$). The second-grade growth parameter estimate ($\beta_{10}$) was positive and statistically significant, indicating improvement in oral reading fluency during second grade. On average, ELLs gained 5.49 wcpm per month (weekly gain of 1.37 wcpm; $t$ (1184) = 33.87, $p$ < .001) during second grade and read 90.12 words per minute on average [46.20 + 5.49 * 8] at the end of the school year.

**Grade 3 initial level and growth.** The $\beta_{20}$ parameter indicated the magnitude and direction of the shift in ORF level at the beginning of third grade after summer vacation. Oral reading fluency levels decreased by an average of 30.47 words per minute between second and third grades ($\beta_{20}$; $t$ (1184) = -20.20, $p$ < .001). Thus, the average student began third grade reading 59.65 wcpm [(46.20 + 5.49 * 8) – 30.47].

The third-grade growth parameter ($\beta_{30}$) was positive and statistically significant, indicating that students increased oral reading fluency skill throughout third grade. On average, ELLs increased ORF by 3.65 wcpm per month (0.91 wcpm per week; $t$ (1184) = 25.29, $p$ < .001)) during third grade. At the end of the school year, the average student read 88.85 wcpm [59.65 + 3.65 * 8].

**Research Question #2**

The second research question was: *What is the effect of Project ELLA on Spanish-speaking ELLs' initial level and growth in English oral reading fluency in Grades 2 and 3?* This question was answered using Model C, a conditional piecewise growth model. As in Model B, parameters for ORF initial level and growth in Grades 2 and 3 were estimated in Model C. In addition, Model C included parameters for the effect of intervention on Grade 2 initial level and growth and Grade 3 initial level (or change in level after summer vacation). A parameter for the effect of intervention on Grade 3 growth was not included in Model C because the results of Model B indicated that the variance in Grade 3 growth rate was minute and not statistically significant ($r_{3i}$ = 0.41, $p$ = .21). Thus, students in the sample increased ORF at similar rates during Grade 3, and including predictors of Grade 3 growth was inappropriate. Table 12 displays parameter estimates, standard errors, and confidence intervals for Model C. AIC and BIC for Models B and C were compared. Large decreases in both fit indices indicated that model fit improved for Model C.

**Group differences in Grade 2 initial level and growth.** The intercept parameter ($\beta_{00}$) indicated the average initial level for students in the control group. The average control student read 46.80 wcpm at the beginning of Grade 2 ($\beta_{00,}$ $t$ (281) = 21.91, $p$ < .001). The intervention parameter for intercept ($\beta_{01}$) was negative, indicating that intervention students began second grade reading slightly less fluently than control students; however, that difference was not statistically significant ($\beta_{01}$ = -3.46, $t$ (281) = -1.18, $p$ = .24). Therefore, both groups began second grade reading at similar levels of fluency.

Like the intercept parameter in this model, the second-grade growth parameter estimate ($\beta_{10}$) indicated the average ORF growth for control group students only. $\beta_{10}$ was positive and statistically significant, indicating improvement in oral reading fluency during second grade. On average, control students gained 4.73 wcpm per month (1.18 wcpm per week; $t$ (1182) = 20.92, $p$ < .001) during Grade 2 and were reading a fitted mean of 84.64 wcpm at the end of the school year [46.80 + 4.73 * 8]. The interaction effect between intervention and Grade 2 growth ($\beta_{11}$) was positive and statistically significant. Therefore, the $\beta_{11}$ parameter estimate quantified the added value of the ELLA intervention on ORF growth during second grade. On average, intervention students gained 1.52 wcpm ($t$ (1182) = 4.89, $p$ < .001) more than control students gained, for a total gain of 6.25 wcpm per month (1.56 wcpm per week). Intervention students read a fitted mean of 96.8 wcpm at the end of second grade [46.8 + 6.25 * 8].

**Group differences in Grade 3 initial level and growth.** The $\beta_{20}$ parameter indicated the magnitude and direction of the shift in ORF level for control group students at the beginning of third grade. On average, control students decreased in ORF level by 28.88 wcpm after summer vacation ($\beta_{20}$; $t$ (1182) = -16.35, $p < .001$). Intervention students decreased fluency by 2.49 words more than control students ($\beta_{21}$), but the difference was not statistically significant ($t$ (1182) = -1.21, $p = .23$). Thus, both groups decreased in fluency level by similar amounts after summer vacation. On average, control students read 55.76 wcpm [84.64 – 28.88] and intervention students read 67.92 wcpm [96.8 – 28.88] at the beginning of Grade 3.

The effect of intervention on Grade 3 ORF growth was not tested because preliminary results showed no statistically significant variance in Grade 3 growth rates. Students generally increased ORF at the same pace regardless of group membership. On average, all students gained 3.76 wcpm per month in Grade 3 ($\beta_{30}$; $t$ (1182) = 27.36, $p < .001$). At the end of the year, control students read a fitted mean of 85.84 wcpm [55.76 + 3.76 * 8] and intervention students read a fitted mean of 98 wcpm [67.92 + 3.76 * 8].

**Effect of intervention.** Figure 6 is a graphic representation of each group's ORF initial status and growth trajectory for second and third grades as estimated in the model. The intervention and control groups were reading at equivalent levels of fluency in the fall of second grade. However, during second grade, the intervention group's fluency growth outpaced the control group's growth such that the average intervention student read 12 wcpm more than the average control student at the end of second grade. In third grade, both groups experienced equivalent declines in level of fluency at the fall assessment and equivalent growth rates during the school year. Nevertheless, the intervention group read at higher levels of fluency throughout third grade, maintaining the higher level of fluency achieved in second grade.

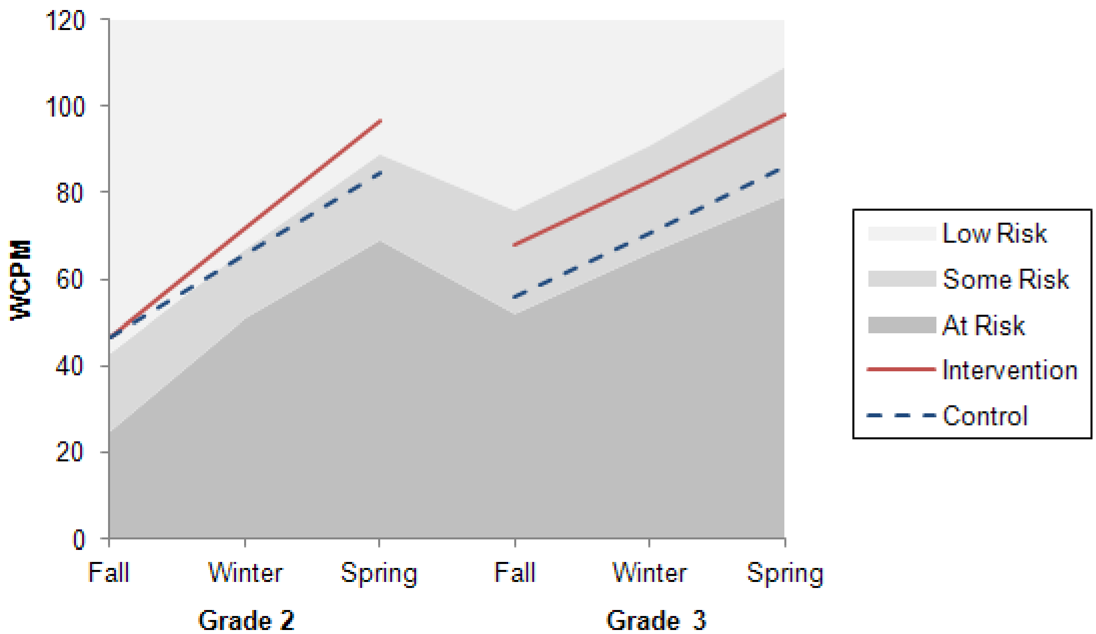


*Figure 6.* Fitted trajectories for second- and third-grade ELLs' English oral reading fluency growth by group. The shaded background indicates risk status according to DIBELS ORF benchmark goals for three assessment periods per year ("DIBELS Benchmark Goals," n.d.). wcpm=words read correctly in 1 min.

## Summary

In this chapter, I presented the results of the study. Preliminary results provided descriptive information about the data and indicated that multilevel growth models were appropriate for analyzing the data. Primary results from multilevel models were used to address the two research questions. Results from the unconditional piecewise linear growth model (Model B) indicated that ELLs' trajectory of ORF growth in Grades 2 and 3 could be described as a two-stage pattern of linear growth with a decrease in ORF level between grades. Slope parameters were positive in both grades but decreased slightly in third grade. Results from the conditional piecewise growth model (Model C) indicated that adding intervention as a predictor improved model fit. The intervention parameters quantified the effect of Project ELLA on students' mean ORF trajectory. During second grade, participating in Project ELLA added 1.52 wcpm per month to intervention students' ORF scores. Although students in both intervention and control groups improved ORF at the same rate in third grade, intervention students maintained the higher level of ORF performance that was reached during second grade. I present a more detailed discussion of the findings and their implications in Chapter V.

# CHAPTER V

# SUMMARY, DISCUSSION, AND CONCLUSIONS

In Chapter V, I provide a summary of the study, along with discussion of the results, implications for practice, and recommendations for future research. I begin by summarizing the study, the methodology employed to address the research questions, and the findings. Next, I discuss the results organized by research question. Finally, I extend the discussion by suggesting implications for educational practice with ELLs and recommendations for future research related to English language and reading acquisition.

## Summary of the Study

### Problem and Purpose

ELLs, who make up approximately 10% of the K-12 student population in U.S. public schools, demonstrate lower levels of English reading proficiency than do native English-speaking students (August & Shanahan, 2006; *Nation's Report Card*, 2009). English reading proficiency is important for academic achievement and completion of high school (Reschly, 2010), and low reading proficiency often results in grade retention and/or referral for special education placement (Bowman-Perrott et al., 2010). Developing grade-level proficiency in English reading is critical for ELLs, who acquire language, literacy, and content knowledge simultaneously.

Oral reading fluency is one indicator of reading proficiency (Fuchs et al., 2001; NICHD, 2000). Although ORF development has been widely researched with native

English-speaking samples (Baker et al., 2008; Chard et al., 2008; Crowe et al., 2009; Deno et al., 2001; Fuchs et al., 1993; Hasbrouck & Tindal, 1992; Hasbrouck & Tindal, 2006; Kim et al., 2010; Speece & Ritchey, 2005; Stage & Jacobsen, 2001; Stage et al., 2001), there have been few studies of ORF development with ELL samples (Al Otaiba et al., 2009; Baker et al., 2012; Betts et al., 2009; Dominguez de Ramirez & Shapiro, 2006; Yesil-Dagli, 2011). Moreover, little attention has been given to L2 reading fluency in general (Al Otaiba et al., 2009; August & Shanahan, 2006). Research on reading fluency instruction for ELLs has emerged in the last decade (Calhoon et al., 2007; Linan-Thompson et al., 2003; McMaster et al., 2008; Saenz et al., 2005; Tong, Irby, et al., 2008; Vaughn, Cirino, et al., 2006; Vaughn, Mathes, et al., 2006). However, with the exception of Tong, Irby, et al. (2008), the interventions implemented in these projects emphasized direct instruction of reading skills with minimal-to-moderate support for comprehensive development of ELLs' English oral language and content knowledge.

The two purposes of the present study emanated from these observations about the literature related to L2 reading fluency. The first purpose was to model English ORF development from the beginning of Grade 2 to the end of Grade 3 (i.e., initial status and growth within each grade) for ELLs receiving instruction in bilingual education programs. The second purpose was to determine the effect of the Project ELLA intervention on ELLs' English reading fluency development.

**Methodology**

Data for this study were archived from Project ELLA, a large-scale, longitudinal, randomized study documenting ELLs' acquisition of English language and reading skills. These data included 1,470 observations of English oral reading fluency from 283 ELLs participating in bilingual education programs at 17 schools. Schools were randomly assigned to the intervention (*n*=8) or control (*n*=9) condition.

In the intervention schools, the Project ELLA intervention was implemented in kindergarten through third grades. The intervention was a comprehensive, multi-tiered intervention that included professional development for teachers and paraprofessionals, content-area curricula as typically provided in the school district, enhancement of the transitional bilingual education program to one-way dual language program, structured daily ESL intervention, and supplemental instruction for the lowest performing students. The structured daily ESL intervention curricula were delivered in 75-min (kindergarten) and 90-min blocks (Grades 1-3) and emphasized L2 oral language development in kindergarten and first grades, basic L2 reading skills in second grade, and content-area reading skills in third grade.

In the control schools, the school district's typical transitional bilingual education program, content-area instruction, and ESL curricula were implemented. ESL instruction was provided in a 45-min block with no support from the research team. ESL instruction varied widely across schools and classrooms, but lessons typically included a focus, instruction, guided practice, and independent practice.

Students' oral reading fluency was measured using DIBELS oral reading fluency assessments on six occasions during second and third grades. ORF was assessed in the fall, winter, and spring of both second and third grades, and 76% of the sample had complete data for all six measurement occasions. There were 1,470 ORF measurements nested within 283 students nested in 29 second-grade classrooms and 22 third-grade classrooms within 17 schools. Given the nested structure of the data, multilevel models were used to analyze the data.

Two 3-level preliminary unconditional models were hypothesized, each with repeated measures nested within students at Levels 1 and 2. In one preliminary model, students were nested within schools at Level 3, whereas in the other preliminary model students were cross-classified in second- and third-grade classrooms at Level 3. In each preliminary model, Level 3 had no statistically significant variance; therefore, the third level was dropped from the model. Subsequent analyses utilized two-level models with repeated measures nested in students.

To accomplish the first purpose of the study, an unconditional piecewise growth model was fit. This model depicted ORF development for the whole sample without predictor variables. To accomplish the second purpose of the study, a conditional piecewise growth model, including the effect of intervention as a Level-2 predictor, was fit. This model estimated differences in ORF development for control and intervention students. Models used FML estimation, and model fit was evaluated using deviance-based likelihood ratio tests, AIC, and BIC. Fixed effects were of primary interest and were tested using *t*-tests. Degrees of freedom were calculated using the between-within method, and 95% confidence intervals were constructed around the fixed effect estimates.

## Discussion of the Findings

### Research Question #1

**Summary of the findings.** In Grades 2 and 3, ELLs in this study followed a two-stage linear trajectory of growth in English oral reading fluency. The trajectory consisted of statistically significant positive linear growth in each grade with a decrease in ORF level between second and third grades. Slope parameters decreased in third grade indicating that students slowed their acquisition of L2 reading fluency skill over time. On average, ELLs in this sample read ~46 wcpm at the beginning of second grade and grew ~5.5 wcpm per month (1.37 wcpm per week) in ORF during second grade. Between second and third grades, ELLs in this sample decreased ORF by ~30 wcpm and began third grade reading at ~60 wcpm on average. During third grade, ELLs in this sample improved ORF at a rate of ~3.7 wcpm per month (<1 wcpm per week).

**L2 reading fluency development in ELLs.** These findings are generally consistent with those reported by Al Otaiba et al. (2009). Compared to ELLs in the Al Otaiba study, ELLs in the present study read at a slightly higher initial level in second grade (46 wcpm vs. 38 wcpm) and at a comparable initial level in third grade (60 wcpm vs. 61 wcpm). For growth trajectory, the findings presented here are similar to those reported by Al Otaiba et al. in that non-linear models of ORF growth fit the data well. For second-grade ELLs, Al Otaiba and colleagues reported a positive, strongly linear trend with slight deceleration (negative quadratic trend). For third-grade ELLs, they reported a strong linear trend in the first 2 months, followed by strong deceleration for the remainder of the year. In the present analysis, the quadratic trend was not tested statistically; instead, non-linearity in ORF trend was accounted for using the two-rate piecewise model with intercept shift. In this two-rate model, the third-grade growth parameter was smaller than the second-grade growth parameter; therefore, similar to the ELLs sampled by Al Otaiba et al., the ELLs sampled here showed some deceleration of ORF growth in third grade. This pattern is not surprising given that the typical NES reading fluency growth pattern is acceleration through first and second grades followed by deceleration at the end of second grade and through third grade (Baker et al., 2008; Chard et al., 2008; Crowe et al., 2009; Kim et al., 2010).

Practitioners may be particularly interested in ELLs' average weekly growth rates for English ORF because weekly growth rates are useful for progress monitoring. The average weekly growth rates for the present sample were within the range of weekly growth rates reported in other studies of ELL samples. In Grade 2, the weekly rate of growth for the ELLs in this study, 1.37 wcpm, was comparable to the 1.3 wcpm per week reported by Baker and Good (1995), slightly higher than the 1.23 wcpm per week reported by Al Otaiba et al. (2009), and almost twice the 0.75 wcpm per week reported by Dominguez de Ramirez and Shapiro (2006). In Grade 3, the weekly rate of growth for this study (0.91 wcpm) was somewhat lower than the rates reported by Al Otaiba et al. (1.31 wcpm) and Betts et al. (1.27 and 1.17 wcpm; 2009) but nearly twice that reported by Dominguez de Ramirez and Shapiro (0.48 wcpm). In addition, ELLs' weekly gains in ORF were similar to gains reported for NES students in Fuchs et al. (1993) and Deno et al. (2001) in Grade 2 (1.25-1.5 wcpm per week). In Grade 3, ELLs fell slightly short of typical NES weekly gains (1-1.25 wcpm per week). It appears that the average ELL might be expected to gain between 1.2 and 1.4 wcpm per week during Grade 2. During Grade 3, ELLs can be expected to decelerate their ORF growth slightly but continue gaining ORF at rates between 0.9 and 1.3 wcpm per week.

**Comparison to ORF benchmarks and norms.** In comparison to DIBELS ORF benchmark goals and the norms published by Hasbrouck and Tindal (2006), ELLs in this study read at expected levels for the beginning and end of Grade 2. According to DIBELS benchmark risk status, students were generally at *low risk* for reading difficulties, and their average fluency level was approximately the 50$^{th}$ percentile compared to the normative sample in Hasbrouck and Tindal (2006). In Grade 3, the students' risk for reading difficulties elevated slightly. Students' ORF levels at the beginning of the year and end of the year placed them at *some risk* on DIBELS and between the 25$^{th}$ and 50$^{th}$ percentiles compared to the 2006 norms. As a whole, this sample of ELLs performed at levels of ORF comparable to NES students in Grade 2 but slightly below the NES average in Grade 3.

**Research Question #2**

**Summary of the findings.** At the beginning of Grade 2, regardless of membership in the control group or intervention group, ELLs read a fitted mean of ~47 wcpm. Receiving ESL instruction in Project ELLA positively influenced students' ORF growth rate. Intervention students improved ORF more rapidly than control students by ~1.5 wcpm per month. Whereas control students improved by 4.73 wcpm per month (1.18 wcpm per week), intervention students improved by 6.25 wcpm per month (1.56 wcpm per week). At the end of second grade, intervention students read 12 wcpm more than control students read (~97 wcpm vs. ~85 wcpm). Between second and third grades, all students decreased ORF by ~29 wcpm regardless of group membership. Thus, at the beginning of Grade 3, control students read a fitted mean of 55.76 wcpm, and intervention students read a fitted mean of 67.92. During third grade, all students, regardless of group membership, improved ORF at a rate of 3.76 wcpm per month (<1 wcpm per week). Although both groups improved at the same rate in Grade 3, intervention students maintained the higher level of ORF achieved as a result of greater growth in second grade. At the end of Grade 3, the average intervention student read 12 wcpm more than the average control student read. Thus, the ELLA intervention produced positive effects in ELLs' English ORF level and rate of growth.

**Impact of the intervention.** The ELLA intervention was successful at accelerating ORF growth in second grade, such that intervention students read with greater fluency compared to control students throughout second and third grades. These findings are consistent with those from other studies of the ELLA intervention's effect on ELLs' L2 language and reading skills. Tong, Lara-Alecio, et al. (2008) and Tong et al. (2010, 2011) reported that ELLA intervention students accelerated English oral language acquisition, phonological awareness, decoding, and reading comprehension skills from kindergarten through second grade compared to the control group. Tong, Irby, et al. (2008) concluded that bilingual education intervention students outperformed control group students on second-grade posttests of English language and reading skills, specifically phonological processing, oral language comprehension, picture vocabulary, and reading fluency. The present study extends previous findings supporting the efficacy of ELLA as a comprehensive L2 intervention by documenting positive effects on L2 oral reading fluency development. To date, the ELLA intervention has demonstrated a positive impact on the full range of L2 reading skills, from precursors to literacy (e.g., phonological awareness and oral language proficiency) to word-level skills (e.g., decoding) to text-level skills (e.g., reading fluency and comprehension).

The ELLA intervention was a holistic intervention with multiple curricular components. The effects reported here reflect the impact of the entire intervention with all curricular components implemented concurrently. Because oral reading fluency is a "complicated, multifaceted performance" of reading skill (Fuchs et al., 2001, p. 239), it is fitting that a multidimensional intervention improved L2 reading fluency for ELLs. The effects of the individual components on reading fluency are unknown; however, each component included instructional procedures that have been shown to improve reading fluency.

Researchers have identified procedures for improving reading fluency: repeated reading of text, guided repeated oral reading of text, modeled reading of text by teachers, and various forms of assisted reading (e.g., echo reading, choral reading, partner reading; Chard et al., 2002; Palumbo & Willcutt, 2006; NICHD, 2000). For ELLs, peer-assisted reading activities and direct instruction of English reading have produced positive effects on L2 reading fluency (Calhoon et al., 2007; Linan-Thompson et al., 2003; McMaster et al., 2008; Saenz et al., 2005; Vaughn, Cirino, et al., 2006; Vaughn, Mathes, et al., 2006). Francis et al. (2006) suggested that fluency instruction should also emphasize vocabulary instruction, increase students' exposure to print, and include oral discussion of texts. Given these findings and recommendations, several aspects of the ELLA intervention possibly contributed to the positive effects for oral reading fluency. Table 13 provides an analysis of the fluency-building instructional activities featured in the ELLA intervention.

Table 13
*Analysis of Fluency-building Instructional Procedures in Three Components of the ELLA Intervention*

| Instructional Procedure | STELLA | EIR-Level II | CRISELLA |
|---|---|---|---|
| Repeated reading of text | + | + | + |
| Modeled reading of text by teachers | + | + | |
| Assisted reading (e.g., echo reading, choral reading) | + | + | + |
| Peer-assisted reading activities | + | + | + |
| Direct instruction of reading skills | | + | + |
| Explicit vocabulary instruction | + | | + |
| Oral discussion | + | | + |

The STELLA story reading component, employed in all 4 years of the intervention, integrated several features of effective fluency instruction using authentic children's literature as texts. First, daily, explicit, and context-embedded vocabulary instruction using picture cards, word walls, sentence stems, and semantic maps expanded ELLs' English vocabulary knowledge. Along with its emphasis on vocabulary, STELLA instruction also emphasized repeated readings of each week's featured book. By the end of each week, students had read the featured book many times using a variety of repeated reading techniques: teacher read-aloud and modeling, choral reading, and partner reading. In addition, comprehension and higher-order thinking questions related to the text were presented in order to elicit oral discussion from students. In these ways, STELLA met the specific literacy needs of ELLs by integrating effective fluency practice with explicit linguistic comprehension instruction.

*EIR (Level II)*, which was implemented throughout second grade, also possibly influenced reading fluency growth for intervention students. Formerly known as *Proactive Reading*, *EIR* resulted in positive effects for ELLs' reading skills when used as a supplemental intervention in small groups (Vaughn, Cirino, et al., 2006; Vaughn, Mathes, et al., 2006). In second-grade ELLA classrooms, *EIR Level II* was adapted for whole-class instruction. Thus, it appears that, even with the adaptations for large-group instruction, *EIR* contributed to improved reading fluency for ELLs. Using *EIR*, intervention teachers provided direct instruction of the five strands of reading: phonemic awareness, letter-sound correspondence, word recognition and spelling, fluency, and comprehension. Fluency instruction was focused on the repeated reading of decodable text. Together with explicit instruction in word recognition strategies, the repeated readings possibly helped intervention students increase their reading accuracy and speed.

With the addition of the CRISELLA component in third grade, the ELLA invention included a stronger focus on content-area reading skills, specifically the reading of expository text in science. Although there were no differences in reading fluency growth in third grade, intervention students maintained the higher level of fluency that they achieved in second grade. It is possible that CRISELLA factored in students' maintenance of higher fluency achievement. CRISELLA lessons integrated literacy instruction into science instruction using the third-grade science textbook as a foundation. A variety of reading skills and activities were included in lesson plans. Among them were partner reading, vocabulary development and extension activities, and repeated readings of textbook passages, all of which are known to support reading fluency development.

## Implications for Practice

### Assessment of ELLs

Early intervention is important in the prevention of future reading difficulties for native English speakers and ELLs (Linan-Thompson, Vaughn, Prater, & Cirino, 2006; Mathes, n.d.; Simmons et al., 2008). Many schools utilize a response-to-intervention (RtI) approach for identifying students who need additional intervention and for monitoring student progress during supplemental interventions. ORF is one assessment that is recommended for use in RtI with ELLs (Gersten et al., 2007). However, the available norms for interpreting ORF progress were derived largely from NES samples ("DIBELS Benchmark Goals," n.d.; Hasbrouck & Tindal, 2006), and these norms have been applied to ELLs' progress with scant evidence for the appropriateness of doing so

(Al Otaiba et al., 2009; Linan-Thompson, 2010). The average ELLs in this study demonstrated ORF levels and weekly growth comparable to average NES students in second grade and somewhat comparable to, though slightly below, average NES students in third grade. The present findings, alongside those of Al Otaiba et al. (2009), Betts et al. (2009), and Baker et al. (2012), shed light on expected levels and rates of growth for L2 ORF in second and third grades and provide tentative evidence that ELLs can meet, or at least approximate, average NES levels of ORF. Thus, existing benchmarks may be appropriate for interpreting ELLs' ORF in Grades 2 and 3.

This recommendation is advanced with a caveat: ELLs' ORF scores should not be considered in isolation but should be one indicator among multiple measures of reading and language skill. In addition to ORF, teachers should collect comprehensive data describing student background and educational experiences, L2 oral proficiency and reading skills, and, when possible, native language and literacy skills. Decisions regarding eligibility for supplemental reading intervention, progress during intervention, and referral for special education services should be made using all available evidence and not ORF alone.

**Summer Regression in ORF**

For all ELLs in this study, summer vacation from school resulted in large decreases in ORF. All students regressed by a similar amount, ~30 wcpm, such that they returned to middle-of-second-grade levels. Regression during summer is typical, and ORF benchmarks account for this ("DIBELS Benchmark Goals," n.d.; Hasbrouck & Tindal, 2006). However, ELLs must accelerate acquisition of L2 language and literacy skills to compensate for limited proficiency at school entry. Therefore, a decrease of the magnitude observed here is disconcerting because the necessary acceleration is hindered when students spend valuable instructional time regaining skills during the initial months of the school year. Increased exposure to L2 reading and language during summer would likely mitigate the regression in skill. Summer English literacy programs, either as formal summer school or as informal literacy programs sponsored by schools, libraries, or communities, would likely benefit ELLs' L2 reading fluency development as well as overall second language and literacy development.

**ESL Instruction in Bilingual Education Programs**

The results of this study give insight into effective approaches for ESL instruction within bilingual program models. Both control and intervention students were instructed within bilingual education programs. The control classrooms followed a TBE model, whereas the intervention classrooms followed a developmental, one-way dual language model. The primary differences between the two models were: (a) an additive maintenance philosophy in the one-way dual language model (compared to the subtractive, transitional philosophy of TBE); (b) ongoing professional development in second language acquisition and ESL instructional strategies for intervention teachers; (c) increased instructional time allotted to the ESL block in the intervention (75/90 min vs. 45 min); and (d) direct, explicit, and tightly structured ESL intervention curricula integrating oral language, context-embedded vocabulary instruction, reading, and science content. These enhancements to the district's typical TBE model produced positive effects in students' acquisition of L2 reading skill. The benefits of additive language programs and ongoing professional development have long been described in the literature (Cummins, 2000; Torres-Guzmán, 2002). The findings presented here and in other Project ELLA studies provided evidence that explicit, direct, well planned, and structured ESL instruction improved students' English language and reading acquisition. Furthermore, despite the increase in English instructional time in the intervention condition compared to control condition, intervention students have shown strong acquisition of Spanish oral language and reading (Tong, Irby, et al., 2008; Tong et al., 2011). Thus, the enhancements to typical TBE resulted in a bilingual education model that promoted strong English acquisition while maintaining and strengthening native language skills.

To ensure strong English acquisition within bilingual education programs, ESL instruction should be explicit, direct, and structured and should emphasize vocabulary acquisition and academic language development. Moreover, ESL instruction should proceed in a developmentally appropriate progression, beginning with a focus on oral language in kindergarten and first grades, adding direct instruction in English reading in first and second grades, and then progressing to content-area reading in third grade. Within this progression, the early focus on oral language and vocabulary lays a foundation for strong text-level English reading skills, such as fluency and comprehension. Direct instruction in reading provides students with knowledge of the English graphophonic system and increases decoding skills. Content-area reading instruction integrates reading instruction within content-area learning, providing valuable instruction and practice for reading expository text in English.

## Recommendations for Further Research

Several recommendations for future research derive from the present study. Researchers of L2 reading fluency should focus on three topics: (a) precursory skills that affect L2 reading fluency initial status, (b) L2 reading fluency growth patterns and how they compare to patterns observed in monolingual English speakers, and (c) the relationship between L1 and L2 reading fluency for students developing biliteracy.

First, the precursory skills to L2 reading fluency are not well understood (August & Shanahan, 2006; Crosson & Lesaux, 2010; Yesil-Dagli, 2011). Although the multilevel models fit well for the data in the present study, there was substantial variability in the second-grade intercept that was not explained even after intervention was included as a predictor in the model. The addition of prereading skills (e.g., letter identification, phonological awareness, oral language proficiency) and early reading skills (e.g., decoding) would likely explain a portion of this variance (Crosson & Lesaux, 2010; Yaghoub Zadeh et al., 2012) and would give researchers insight into which precursory skills support L2 reading fluency in ELLs. This knowledge would inform early L2 literacy instructional practices.

Second, researchers should continue investigating ELLs' L2 reading fluency growth patterns. Additional longitudinal studies, particularly those spanning more than two school years, would provide a more complete picture of expected L2 reading fluency growth patterns for ELLs and would enable researchers to ascertain with greater certainty whether NES expectations are appropriate for ELLs. Direct comparisons of growth between ELL and NES students would be useful for similar purposes. There have been few studies making direct comparisons (Baker & Good, 1995; Dominguez de Ramirez & Shapiro, 2006).

Finally, although this study included only L2 reading fluency, future studies might include L1 reading fluency as well. Many ELLs are educated in bilingual education programs with biliteracy as an explicit programmatic goal. For students with strong literacy skills in both languages, crosslinguistic transfer between L1 and L2 occurs (Riches & Genesee, 2006). Few investigators have studied the influence of L1 reading fluency on L2 reading fluency (Baker et al., 2012; Dominguez de Ramirez & Shapiro, 2007). Knowing if and how L1 and L2 reading fluency interact in reading fluency development would assist bilingual education practitioners as they design and implement high-quality biliteracy instruction and assessments for ELLs in bilingual education programs.

## Conclusion

Young ELLs face daunting educational challenges. If they are to meet these challenges, strong English reading skills are imperative. For educators tasked with ensuring that ELLs acquire English language and literacy, an understanding of how L2 reading skills develop and which instructional practices enhance reading skill development is indispensable for improving ELLs' L2 reading achievement. However, relatively few research studies have documented L2 reading development (August & Shanahan, 2006; Genesee et al., 2006). This is especially true for L2 reading fluency (Al Otaiba et al., 2009; August & Shanahan, 2006). In the absence of research evidence specific to ELLs and L2 reading, many recommendations about instruction of reading fluency and comprehension for ELLs are derived from evidence with monolingual readers rather than from research with second-language learners (Francis et al., 2006). I have attempted to address this lacuna with the present study, in which I investigated the developmental trajectory of L2 reading fluency and the effect of the Project ELLA intervention on reading fluency development.

Findings indicated that ELLs in Grades 2 and 3 exhibited a two-stage linear trajectory in English oral reading fluency growth. Students demonstrated statistically significant positive linear growth within each grade but slowed fluency growth in Grade 3. Between grades, they experienced a large average decrease in fluency level. In Grade 2, ELLs' levels and average weekly growth were comparable to expectations for native English-speaking students. In Grade 3, ELLs performed slightly below native-speaker expectations. Findings for the intervention indicated that ELLs in the intervention condition accelerated fluency development in Grade 2 and were performing at a higher level of reading fluency by the end of the school year compared to ELLs in the control group. In Grade 3, both intervention and control students grew at the same rate, but intervention students maintained the higher level of reading fluency that they had established in second grade.

This study has contributed to the knowledge base in several ways. First, the same cohort of students was followed for 2 years, generating a longitudinal view of L2 reading fluency that is largely absent from the literature. Given the longitudinal nature of language and literacy acquisition, this is an important methodological contribution. Second, these findings have expanded current understanding of L2 reading fluency by providing information about reading fluency level and growth expectations for ELLs. This information will help educators appropriately utilize and interpret ORF assessments with ELLs. Finally, these results confirmed the positive effects of the ELLA intervention for ELLs and extended the effects to L2 reading fluency. As a comprehensive intervention, the ELLA model enabled teachers to provide powerful instruction in oral language, vocabulary knowledge, reading skills, and content-area reading, which resulted in a wide range of positive effects for English language and reading acquisition. These are small steps toward the larger goal of understanding L2 reading and instruction. As research continues and knowledge of L2 reading development broadens, educators will be able to assess ELLs more equitably, interpret assessment results more validly, and instruct ELLs more effectively.

# REFERENCES

Acosta, S. T. (2010). *High-stakes reading assessment and English oral language development: A study of third grade English language learners in a Texas school district* (Doctoral dissertation, Texas A&M University). Retrieved from http://hdl.handle.net/1969.1/ETD-TAMU-2010-05-8043

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Akaike, H. (1974). A new look at the statistical model of identification. *IEEE Transaction on Automatic Control, 19*, 716-723.

Al Otaiba, S., Petscher, Y., Pappamihiel, N. E., Williams, R. S., Dyrlund, A. K., & Connor, C. (2009). Modeling oral reading fluency development in Latino students: A longitudinal study across second and third grade. *Journal of Educational Psychology, 101*, 315-329. doi: 10.1037/a0014698

Allington, R. L. (1983). Fluency: The neglected goal. *The Reading Teacher, 36*, 556-561.

August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, NJ: Lawrence Erlbaum.

Baker, D. L., Park, Y., & Baker, S. K. (2012). The reading performance of English learners in grades 1–3: The role of initial status and growth on reading fluency in

Spanish and English. *Reading and Writing, 25*, 251-281. doi: 10.1007/s11145-010-9261-z

Baker, S. K., & Good, R. (1995). Curriculum-based measurement of English reading with bilingual Hispanic students: A validation study with second-grade students. *School Psychology Review, 24*, 561-578.

Baker, S. K., Smolkowski, K., Katz, R., Fien, H., Seeley, J. R., Kame'enui, E. J., & Beck, C. T. (2008). Reading fluency as a predictor of reading proficiency in low-performing, high-poverty schools. *School Psychology Review, 37*, 18-37.

Barger, J. (2003). *Comparing the DIBELS Oral Reading Fluency indicator and the North Carolina end of grade reading assessment* (Technical Report). Asheville, NC: North Carolina Teacher Academy. Retrieved from https://dibels.uoregon.edu/docs/techreports/NC_Tech_Report.pdf

Beretvas, S. N. (2004). Using hierarchical linear modeling for literacy research under No Child Left Behind. *Reading Research Quarterly, 39*, 95-99.

Betts, J., Bolt, S., Decker, D., Muyskens, P., & Marston, D. (2009). Examining the role of time and language type in reading development for English language learners. *Journal of School Psychology, 47*, 143-166. doi:10.1016/j.jsp.2008.12.002

Betts, J., Muyskens, P., & Marston, D. (2006). Tracking the progress of students whose first language is not English towards English proficiency: Using CBM with English language learners. *MinneTESOL/WITESOL Journal, 23*, 15-37.

Bowman-Perrott, L. J., Herrera, S., & Murry, K. (2010). Reading difficulties and grade retention: What's the connection for English language learners? *Reading and Writing Quarterly, 26,* 91-107. doi: 10.1080/10573560903397064

Brown, G., & Irby, B. J. (2001). *The pincipal portfolio*. Thousand Oaks, CA: Corwin Press.

Buck, J., & Torgesen, J. (2003). *The relationship between performance on a measure of oral reading fluency and performance on the Florida Comprehensive Assessment Test* (FCRR Technical Report #1). Tallahassee, FL: Florida Center for Reading Research. Retrieved from https://dibels.uoregon.edu/docs/techreports/fcrr_fcat_report.pdf

Calhoon, M. B., Al Otaiba, S., Cihak, D., King, A., & Avalos, A. (2007). Effects of a peer-mediated program on reading skill acquisition for two-way bilingual first-grade classrooms. *Learning Disability Quarterly, 30*, 169-184.

Capps, R., Fix, M., Murray, J., Ost, J., Passel, J. S., & Herwantoro, S. (2005). *The new demography of America's schools: Immigration and the No Child Left Behind Act* (Research Report). Washington, DC: The Urban Institute. Retrieved from http://www.urban.org/Publications/311230.html

Carnine, D. W., Silbert, J., & Kame'enui, E. J. (2004). *Direct instruction reading* (4th ed.). Upper Saddle River, NJ: Merrill-Prentice Hall.

Chall, J. S. (1996). *Stages of reading development* (2nd ed.). Fort Worth, TX: Harcourt Brace College Publishers.

Chard, D. J., Pikulski, J. J., & McDonagh, S. H. (2006). Fluency: The link between decoding and comprehension for struggling readers. In T. Rasinski, C. Blachowicz & K. Lems (Eds.), *Fluency instruction: Research-based best practices* (pp. 39-61). New York, NY: Guilford Press.

Chard, D. J., Stoolmiller, M., Harn, B. A., Wanzek, J., Vaughn, S., Linan-Thompson, S., & Kame'enui, E. J. (2008). Predicting reading success in a multilevel schoolwide reading model: A retrospective analysis. *Journal of Learning Disabilities, 41*, 174-188. doi: 10.1177/0022219407313588

Chard, D. J., Vaughn, S., & Tyler, B. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of Learning Disabilities, 35*, 386-406.

Chiappe, P., & Siegel, L. S. (2006). A longitudinal study of reading development of Canadian children from diverse linguistic backgrounds. *The Elementary School Journal*, *107*, 135-152.

Collier, V. P., & Thomas, W. P. (2004). The astounding effectiveness of dual language education for all. *NABE Journal of Research and Practice, 2*(1), 1-20. Retrieved from http://njrp.tamu.edu/2004/PDFs/Collier.pdf

Cooney, T. (2006). *Scott Foresman science, grade 3*. Glenview, IL: Pearson/Scott Foresman.

Crosson, A. C., & Lesaux, N. K. (2010). Revisiting assumptions about the relationship of fluent reading to comprehension: Spanish-speakers' text-reading fluency in English. *Reading and Writing, 23*, 475-494. doi: 10.1007/s11145-009-9168-8

Crowe, E. C., Connor, C. M., & Petscher, Y. (2009). Examining the core: Relations among reading curricula, poverty, and first through third grade reading achievement. *Journal of School Psychology, 47*, 187-214. doi: 10.1016/j.jsp.2009.02.002

Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. Clevedon, England: Multilingual Matters.

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30*, 507-524.

Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children, 49*, 36-45.

DIBELS Benchmark Goals. (n.d.). Eugene, OR: Institute for the Development of Educational Achievement. Retrieved from https://dibels.uoregon.edu/benchmark.php

Dominguez de Ramirez, R., & Shapiro, E. S. (2006). Curriculum-based measurement and the evaluation of reading skills of Spanish-speaking English language learners in bilingual education classrooms. *School Psychology Review, 35*, 356-369.

Dominguez de Ramirez, R., & Shapiro, E. S. (2007). Cross-language relationship between Spanish and English oral reading fluency among Spanish-speaking English language learners in bilingual education classrooms. *Psychology in the Schools, 44*, 795-806. doi: 10.1002/pits.20266

Dowhower, S. L. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory into Practice, 30*, 165-175.

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly, 38*, 78-103.

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading, 18*, 116-125.

Ehri, L. C. (1998). Grapheme-phoneme knowledge is essential to learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 3-40). Mahwah, NJ: Lawrence Erlbaum.

Ehri, L. C. (1999). Phases of development in learning to read words. In J. Oakhill & R. Beard (Eds.), *Reading development and the teaching of reading: A psychological perspective* (pp. 79-108). Oxford, UK: Blackwell Science.

Eldredge, J. L. (2005). Foundations of fluency: An exploration. *Reading Psychology, 26*, 161-181. doi: 10.1080/02702710590930519

Ferron, J. M., Hogarty, K. Y., Dedrick, R. F., Hess, M. R., Niles, J. D., & Kromrey, J. D. (2008). Reporting results from multilevel analyses. In A. A. O'Connell & D. B. McCoach (Eds.), *Multilevel modeling of educational sata* (pp. 391-426). Charlotte, N.C.: Information Age.

Francis, D., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). *Practical guidelines for the education of English language learners: Research-based recommendations for instruction and academic interventions.* Portsmouth, NH: RMC Research Corporation, Center on Instruction. Retrieved from http://www.centeroninstruction.org/practical-guidelines-for-the-education-of-english-language-learners-research-based-recommendations-for-instruction-and-academic-interventions

Fry, R. (2007). How far behind in math and reading are English language learners? (Research Report). Washington, DC: Pew Hispanic Center. Retrieved from http://www.pewhispanic.org/2007/06/06/how-far-behind-in-math-and-reading-are-english-language-learners/

Fuchs, L. S., & Deno, S. L. (1992). Effects of curriculum within curriculum-based measurement. *Exceptional Children, 58*, 232-243.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review*, *22*, 27-48.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*, 239-256.

Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education*, *9*(2), 20-28.

Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction* (6th ed.). White Plains, NY: Longman.

Genesee, F., Lindholm-Leary, K., Saunders, W. M., & Christian, D. (Eds.). (2006). *Educating English language learners: A synthesis of research evidence*. New York, NY: Cambridge University Press.

Gersten, R., Baker, S. K., Shanahan, T., Linan-Thompson, S., Collins, P., & Scarcella, R. (2007). *Effective literacy and English language instruction for English learners in the elementary grades: A practice guide* (NCEE 2007-4011). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/practice_guides/20074011.pdf

Geva, E. (2006). Second-language oral proficiency and second-language literacy. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (pp. 123-139). Mahwah, NJ: Lawrence Erlbaum.

Geva, E., & Yaghoub Zadeh, Z. (2006). Reading efficiency in native English-speaking and English-as-a-second-language children: The role of oral proficiency and underlying cognitive-linguistic processes. *Scientific Studies of Reading, 10*, 31-57.

Good, R. H., & Kaminski, R. A. (Eds.). (2002a). *Dynamic Indicators of Basic Early Literacy Skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Good, R. H., & Kaminski, R. A. (Eds.). (2002b). *Dynamic Indicators of Basic Early Literacy Skills: Administration and scoring guide* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5*, 257-288.

Gottardo, A., & Mueller, J. (2009). Are first- and second-language factors related in predicting second-language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology, 101*, 330-344. doi: 10.1037/a0014320

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*(1), 6-10.

Graves, A. W., Plasencia-Peinado, J., Deno, S. L., & Johnson, J. R. (2005). Formatively evaluating the reading progress of first-grade English learners in multiple-language classrooms. *Remedial and Special Education, 26*, 215-225.

Gunn, B., Biglan, A., Smolkowski, K., & Ary, D. (2000). The efficacy of supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school. *The Journal of Special Education, 34*, 90-103.

Gunn, B., Smolkowski, K., Biglan, A., & Black, C. (2002). Supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school: A follow-up. *The Journal of Special Education, 36*, 69-79.

Gunn, B., Smolkowski, K., Biglan, A., Black, C., & Blair, J. (2005). Fostering the development of reading skill through supplemental instruction: Results for Hispanic and non-Hispanic students. *The Journal of Special Education, 39*, 66-85.

Hasbrouck, J. E., & Tindal, G. (1992). Curriculum-based oral reading fluency norms for students in grades 2 through 5. *Teaching Exceptional Children, 24*(3), 41-44.

Hasbrouck, J. E., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher, 59*, 636-644. doi: 10.1598/RT.59.7.3

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2*, 127-160.

Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah: NJ: Lawrence Erlbaum.

Irby, B. J., Lara-Alecio, R., Mathes, P. G., Rodriguez, L., & Guerrero, C. (2008, February). *Promoting bilingualism and biliteracy: Programmatic difference between one-way dual language and transitional bilingual programs*. Paper presented at the meeting of the National Association for Bilingual Education, Tampa, FL.

Irby, B. J., Lara-Alecio, R., Tong, F., Rodriguez, L., & Guerrero, C. (2009, February). *One-way developmental dual language model: practical application of scientifically-based research components*. Paper presented at the meeting of the National Association for Bilingual Education, Austin, TX.

Irby, B. J., Quiros, A. M., Lara-Alecio, R., Rodriguez, L., & Mathes, P. (2008, August 12). *What administrators should know about a research-based oral language development intervention for English language learners: A description of Story retelling and higher order Thinking for English Language and Literacy*

*Acquisition--STELLA*. Retrieved from the Connexions website: http://cnx.org/content/m17321/1.1/

Jackson, N. E., & Lu, W.-H. (1992). Bilingual precocious readers of English. *Roeper Review, 14*(3), 115-119.

Kame'enui, E. J., & Simmons, D. C. (2001). Introduction to this special issue: The DNA of reading fluency. *Scientific Studies of Reading, 5*, 203-210.

Kim, Y., Petscher, Y., Schatschneider, C., & Foorman, B. (2010). Does growth rate in oral reading fluency matter in predicting reading comprehension achievement? *Journal of Educational Psychology, 102*, 652-667. doi: 10.1037/a0019643

Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology, 100*, 310-321. doi: 10.1037/0022-0663.100.2.310

Kline, R. B. (2009). *Becoming a behavioral science researcher: A guide to producing research that matters.* New York, NY: Guilford Press.

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95*, 3-21. doi: 10.1037/0022-0663.95.1.3

Kwok, O., West, S. G., & Green, S. B. (2007). The impact of misspecifying the within-subject covariance structure in multiwave longitudinal multilevel models: A Monte Carlo study. *Multivariate Behavioral Research, 42*, 557-592. doi: 10.1080/00273170701540537

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293-323.

Lakeshore Learning. (1997). *Question of the day*. Carson, CA: Author.

Lara-Alecio, R., & Parker, R. I. (1994). A pedagogical model for transitional English bilingual classrooms. *Bilingual Research Journal, 18*(3/4), 119-133.

Li, D., & Nes, S. (2001). Using paired reading to help ESL students become fluent and accurate readers. *Reading Improvement, 38*, 50-61.

Linan-Thompson, S. (2010). Response to instruction, English language learners and disproportionate representation: The role of assessment. *Psicothema, 22*, 970-974.

Linan-Thompson, S., Cirino, P. T., & Vaughn, S. (2007). Determining English language learners' response to intervention: Questions and some answers. *Learning Disability Quarterly, 30*, 185-195.

Linan-Thompson, S., Vaughn, S., Hickman-Davis, P., & Kouzekanani, K. (2003). Effectiveness of supplemental reading instruction for second-grade English language learners with reading difficulties. *The Elementary School Journal, 103*, 221-238.

Linan-Thompson, S., Vaughn, S., Prater, K., & Cirino, P. T. (2006). The response to intervention of English language learners at risk for reading problems. *Journal of Learning Disabilities, 39*, 390-398.

Lindsey, K. A., Manis, F. R., & Bailey, C. E. (2003). Prediction of first-grade reading Spanish-speaking English-language learners. *Journal of Educational Psychology, 95*, 482-494. doi: 10.1037/0022-0663.95.3.482

Marston, D. B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children*. New York, NY: Guilford Press.

Mathes, P. G. (n.d.). The case for early intervention in reading. Retrieved from https://www.sraonline.com/download/EarlyInterventionsinReading/Research/case_for_eir.pdf

Mathes, P. G., & Torgesen, J. K. (2005a). *SRA early interventions in reading, level 1*. Columbus, OH: SRA/McGraw-Hill.

Mathes, P. G., & Torgesen, J. K. (2005b). *SRA early interventions in reading, level 2*. Columbus, OH: SRA/McGraw-Hill.

McMaster, K. L., Kung, S., Han, I., & Cao, M. (2008). Peer-assisted learning strategies: A "Tier 1" approach to promoting English learners' response to intervention. *Exceptional Children, 74*, 194-214.

Melton, E. K. (2009). *The best school districts in Texas for Latino students 2007* (Technical Report). College Station, TX: Texas Educational Excellence Project, Texas A&M University. Retrieved from http://perg.tamu.edu/PERG/The_Best_School_Districts_in_Texas_for_Latinos_2007_files/LatinoFullReport07.pdf

Muyskens, P., Betts, J., Lau, M. Y., & Marston, D. (2009). Predictive validity of curriculum-based measures in the reading assessment of students who are English language learners. *The California School Psychologist, 14*, 11-21.

Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2007). A longitudinal analysis of English language learners' word decoding and reading comprehension. *Reading and Writing, 20*, 691-719. doi: 10.1007/s11145-006-9045-7

Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2008). A cross-linguistic investigation of English language learners' reading comprehension in English and Spanish. *Scientific Studies of Reading, 12*, 351-371. doi: 10.1080/10888430802378526

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

*Nation's report card: Reading*. (2009). Washington, DC: National Assessment of Educational Progress, National Center for Education Statistics, U.S. Department of Education. Retrieved from http://nationsreportcard.gov/reading_2009/

Orosco, M. J., Almanza de Schonewise, E., de Onis, C., Klingner, J. K., & Hoover, J. J. (2008). Distinguishing between language acquisition and learning disabilities among English language learners: Background information. In J. K. Klingner, J. J. Hoover & L. M. Baca (Eds.), *Why do English language learners struggle with* *reading?: Distinguishing language acquisition from learning disabilities* (pp. 5-16). Thousand Oaks, CA: Corwin Press.

Palumbo, T. J., & Willcutt, J. R. (2006). Perspectives on fluency: English-language learners and students with dyslexia. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 159-178). Newark, DE: International Reading Association.

Perfetti, C. A. (1985). *Reading ability*. New York, NY: Oxford University Press.

Pikulski, J. J. (2006). Fluency: A developmental and language perspective. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 70-93). Newark, DE: International Reading Association.

Pikulski, J. J., & Chard, D. J. (2005). Fluency: Bridge between decoding and reading comprehension. *The Reading Teacher, 58,* 510-519. doi: 10.1598/RT.58.6.2

Planty, M., Hussar, W., Snyder, T., Kena, G., KewalRamani, A., Kemp, J., . . . Dinkes, R. (2009). *The condition of education 2009* (NCES 2009-081). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://nces.ed.gov/pubs2009/2009081.pdf

Pollard-Durodola, S., Mathes, P. G., Vaughn, S., Cardenas-Hagan, E., & Linan-Thompson, S. (2006). The role of oracy in developing comprehension in Spanish-speaking English language learners. *Topics in Language Disorders, 26,* 365-384.

Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology, 97,* 246-256. doi: 10.1037/0022-0663.97.2.246

*Quality counts: Perspectives on a population: English-language learners in American schools*. (2009). Bethesda, MD: Education Week & Editorial Projects in Education.

Rasinski, T. (2006). A brief history of reading fluency. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 4-23). Newark, DE: International Reading Association.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Reschly, A. L. (2010). Reading and school completion: Critical connections and Matthew effects. *Reading and Writing Quarterly, 26*, 67-90. doi: 10.1080/10573560903397023

Riches, C., & Genesee, F. (2006). Literacy: Crosslinguistic and crossmodal issues. In F. Genesee, K. Lindholm-Leary, W. M. Saunders & D. Christian (Eds.), *Educating English language learners: A synthesis of research evidence* (pp. 64-108). Cambridge, UK: Cambridge University Press.

Roberts, T., & Neal, H. (2004). Relationships among preschool English language learners' oral proficiency in English, instructional experience and literacy development. *Contemporary Educational Psychology, 29*, 293-311. doi: 10.1016/j.cedpsych.2003.08.001

Saenz, L. M., Fuchs, L. S., & Fuchs, D. (2005). Peer-assisted learning strategies for English language learners with learning disabilities. *Exceptional Children, 71*, 231-247.

Samuels, S. J. (2006). Toward a model of reading fluency. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 24-46). Newark, DE: International Reading Association.

Santoro, L. E., Jitendra, A. K., Starosta, K., & Sacks, G. (2006). Reading well with Read Well: Enhancing the reading performance of English language learners. *Remedial and Special Education*, *27*, 105-115.

SAS Institute. (2011). SAS (Version 9.3) [Computer software]. Cary, NC: Author.

Saunders, W. M., & O'Brien, G. (2006). Oral language. In F. Genesee, K. Lindholm-Leary, W. M. Saunders & D. Christian (Eds.), *Educating English language learners: A synthesis of research evidence* (pp. 14-63). Cambridge, UK: Cambridge University Press.

Schreiber, P. A. (1991). Understanding prosody's role in reading acquisition. *Theory into Practice*, *30*, 158-164.

Schwartz, G. (1978). Estimating the dimensions of a model. *Annals of Statistics, 6*, 461-464.

Shanahan, T., & Beck, I. (2006). Effective literacy teaching for English-language learners. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (pp. 415-488). Mahwah, NJ: Lawrence Erlbaum.

Shaw, R., & Shaw, D. (2002). *DIBELS Oral Reading Fluency-based indicators of third grade reading skills for Colorado State Assessment Program (CSAP)* (Technical Report). Eugene, OR: University of Oregon. Retrieved from https://dibels.uoregon.edu/docs/techreports/shaw_csap_technical_report.pdf

Shinn, M. R., Good, R. H., Knutson, N., Tilly, W. D., & Collins, V. L. (1992). Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review, 21,* 459-479.

Simmons, D. C., Coyne, M. D., Kwok, O., McDonagh, S. H., Harn, B. A., & Kame'enui, E. J. (2008). Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities*, *41*, 158-173. doi: 10.1177/0022219407313587

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford, UK: Oxford University Press.

Speece, D. L., & Ritchey, K. D. (2005). A longitudinal study of the development of oral reading fluency in young children at risk for reading failure. *Journal of Learning Disabilities*, *38*, 387-399.

Sprick, M., Howard, L. M., & Fidanque, A. (1998-2000). *Read well: Critical foundations in primary reading*. Longmont, CO: Sopris West.

Stage, S. A., & Jacobsen, M. D. (2001). Predicting student success on a state-mandated performance-based assessment using oral reading fluency. *School Psychology Review, 30*, 407-419.

Stage, S. A., Sheppard, J., Davidson, M. M., & Browning, M. M. (2001). Prediction of first-graders' growth in oral reading fluency using kindergarten letter fluency. *Journal of School Psychology, 39*, 225-237.

Swanson, H. L., Rosston, K., Gerber, M., & Solari, E. (2008). Influence of oral language and phonological awareness on children's bilingual reading. *Journal of School Psychology, 46*, 413-429. doi: 10.1016/j.jsp.2007.07.002

Texas Administrative Code. (1996). Chapter 89. Adaptations for Special Populations. Subchapter BB. Commissioner's Rules Concerning State Plan for Educating English Language Learners. §89.1225. Retrieved from http://ritter.tea.state.tx.us/rules/tac/chapter089/ch089bb.html

Texas Education Agency. (2008). Texas Academic Excellence Indicator System. Retrieved from the Texas Education Agency website: http://ritter.tea.state.tx.us/perfreport/aeis/

Texas Education Agency. (2010). Texas public school statistics (2009-10) pocket edition. Retreived from the Texas Education Agency website: http://ritter.tea.state.tx.us/perfreport/pocked/2010/pocked0910.pdf

Tindal, G., Marston, D., & Deno, S. L. (1983). *The reliability of direct and repeated measurement* (Research Report #84). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities.

Tong, F. (2006). *Oral English development and its impact on emergent reading achievement: A comparative study of transitional bilingual and structured*

*English immersion models* (Doctoral dissertation, Texas A&M University). Retrieved from http://hdl.handle.net/1969.1/ETD-TAMU-1114.

Tong, F., Irby, B. J., Lara-Alecio, R., & Mathes, P. G. (2008). English and Spanish acquisition by Hispanic second graders in developmental bilingual programs: A 3-year longitudinal randomized study. *Hispanic Journal of Behavioral Sciences, 30*, 500-529.

Tong, F., Irby, B. J., Lara-Alecio, R., Yoon, M., & Mathes, P. G. (2010). Hispanic English learners' responses to longitudinal English instructional intervention and the effect of gender: A multilevel analysis. *The Elementary School Journal, 110*, 542-566. doi: 10.1086/651195

Tong, F., Lara-Alecio, R., Irby, B., Mathes, P., & Kwok, O. (2008). Accelerating early academic oral English development in transitional bilingual and structured English immersion programs. *American Educational Research Journal, 45*, 1011-1044. doi: 10.3102/0002831208320790

Tong, F., Lara-Alecio, R., Irby, B. J., & Mathes, P. G. (2011). The effects of an instructional intervention on dual language development among first-grade Hispanic English-learning boys and girls: A two-year longitudinal study. *The Journal of Educational Research, 104*, 87-99. doi: 10.1080/00220670903567364

Torgesen, J. K., & Hudson, R. F. (2006). Reading fluency: Critical issues for struggling readers. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 130-158). Newark, DE: International Reading Association.

Torres-Guzmán, M. E. (2002). Dual language programs: Key features and results. In *NCBE Directions in Languge and Education: Vol. 14*. Washington, DC: National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs. Retrieved from http://www.ncela.gwu.edu/files/rcd/BE021969/14.pdf

Trevino, E. P., Guerrero, C., Lara-Alecio, R., Irby, B. J., Mathes, P. G., & Rodriguez, L. (2007). *Daily oral and written language: Science and language skills, 10 minutes at a time*. Paper presented at the meeting of the Texas Association for Bilingual Education, San Antonio, TX.

Vander Meer, C. D., Lentz, F. E., & Stollar, S. (2005). *The relationship between oral reading fluency and Ohio Proficiency Testing in reading* (Technical Report). Eugene, OR: University of Oregon. Retrieved from https://dibels.uoregon.edu/docs/techreports/ohio.pdf

Vaughn, S., Cirino, P., Linan-Thompson, S., Mathes, P. G., Carlson, C. D., Cardenas-Hagan, E., . . . Francis, D. J. (2006). Effectiveness of a Spanish intervention and an English intervention for English-language learners at risk for reading problems. *American Educational Research Journal, 43*, 449-487.

Vaughn, S., Mathes, P., Linan-Thompson, S., Cirino, P., Carlson, C., Pollard-Durodola, S., . . . Francis, D. (2006). Effectiveness of an English intervention for first-grade English language learners at risk for reading problems. *The Elementary School Journal*, *107*, 153-180. doi: 10.1086/510653

Ventriglia, L., & González, L. (2000). *Santillana intensive English*. Miami, FL: Santillana USA.

Wiley, H. I., & Deno, S. L. (2005). Oral reading and maze measures as predictors of success for English learners on a state standards assessment. *Remedial and Special Education, 26*, 207-214.

Wilson, J. (2005). *The relationship of Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Oral Reading Fluency to performance on Arizona Instrument to Measure Standards (AIMS)* (Research Brief). Tempe, AZ: Tempe School District No. 3. Retrieved from https://dibels.uoregon.edu/docs/techreports/arizona.pdf

Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading, 5*, 211-239.

Yaghoub Zadeh, Z., Farnia, F., & Geva, E. (2012). Toward modeling reading comprehension and reading fluency in English language learners. *Reading and Writing, 25*, 163-187. doi: 10.1007/s11145-010-9252-0

Yesil-Dagli, U. (2011). Predicting ELL students' beginning first grade English oral reading fluency from initial kindergarten vocabulary, letter naming, and phonological awareness skills. *Early Childhood Research Quarterly, 26*, 15-29. doi: 10.1016/j.ecresq.2010.06.001

## APPENDIX A

## LIST AND DESCRIPTION OF ESL STRATEGIES EMBEDDED IN INTERVENTION CURRICULUM

| ESL Strategy | Description |
|---|---|
| Academic language scaffolding | Teachers support academic language comprehension and acquisition using various strategies, including modeling, visuals, gestures, demonstrations, and interactive learning activities. |
| Advanced organizers | Before new information is presented, graphic organizers are used to help students connect their prior knowledge and experiences. |
| Bridging | Teachers help students understand the meanings of words, particularly action verbs and descriptive words related to color, shape, size, etc. Students must then identify or label those meanings in English. Bridging is directly correlated to meaning and understanding. |
| Communication games | Communication games are small-group learning activities designed to provide authentic opportunities and purposes for oral language practice. |
| Free voluntary reading | Teachers provide class time for students to read self-selected materials. |
| Interactive read-aloud | Teachers read books aloud with expression, different voices, and gestures, and involve the students in active listening through discussion (including eliciting students' predictions and checking for students' comprehension). |
| Language experience | Teachers engage students in a writing lesson that builds on their shared personal experiences (e.g., field trips, science demonstration, etc.), which helps students connect oral and written language. |
| Leveled questioning | Teachers differentiate questioning for each student's level of oral English proficiency. Teachers also adapt questioning by adding gestures, using visuals, or slowing rate of speech. |
| Manipulative and realia strategies | Real objects and/or concrete representation manipulatives are used to develop academic vocabulary. |
| Partner work and tutoring | Teachers pair 2 students and provide them with detailed instructions to carry out a specific learning task. Pairs demonstrate their learning to other pairs, in small groups, or to the entire class. |
| Preview and review | Teacher pre-teaches key vocabulary and concepts, teaches the lesson, and then reviews the key vocabulary and concepts. |
| Total physical response and music and movement | Teachers introduce progressively more difficult commands, using body movements to dramatize the meanings as they say the commands. Students respond by acting out the actions with the teacher. Teacher demonstration of the commands is gradually removed so that students move in response to the oral commands only. Students listen, acquire language, and show comprehension through movements, speaking commands or other language only when ready. |
| Word/story dramatization | Teachers use role-play and dramatization to explore and discuss new vocabulary words or to review stories that students have read or have heard read. |
| Word walls | Teachers display high-frequency words, words from literature, and/or content area vocabulary on a special area of classroom wall space. |

Adapted from Tong et al. (2010)

## APPENDIX B

## RESULTS OF PRELIMINARY 3-LEVEL MODELS

| | Preliminary Model 1 | Preliminary Model 2 |
|---|---|---|
| **Model description** | 3-level unconditional means model with school at Level 3 | 3-level unconditional means model with students cross-classified by Grade2 and Grade 3 classrooms at Level 3 |
| Sample size (Level 2, students) | 283 | 227 |
| Number of observations | 1470 | 1346 |
| ICC | .014 | .046 (Grade 2 classrooms) .075 (Grade 3 classrooms) |
| **Fixed effects** | | |
| Grand mean, $\gamma_{000}$ | 75.85 (1.97) | 82.02 (3.03) |
| **Variance components** | | |
| Level 1 (Within-person), $e_{tij}$ | 580.54*** (23.88) | 576.75*** (24.37) |
| Level 2 (between-person), $r_{0ij}$ | 641.63*** (68.30) | 452.26*** (55.23) |
| Level 3 (between-school), $u_{00j}$ | 16.85 (22.94) | |
| Level 3 (between Grade 2 Classrooms), $b_{00k10}$ | | 57.43 (47.34) |
| Level 3 (between Grade 3 Classrooms), $c_{000k2}$ | | 87.68 (57.93)~ |
| **Fit indices** | | |
| Deviance | 14057.00 | 12800.50 |
| AIC | 14065.00 | 12810.50 |
| BIC | 14057.00 | 12800.50 |

~ $p$ < .10, *** $p$ <.001