Clinical Forum

Bilingual Children's Language Abilities and Early Reading Outcomes in Head Start and Kindergarten

Carol Scheffner Hammer
Frank R. Lawrence
Adele W. Miccio
The Pennsylvania State University, University Park

U.S. statistics have consistently shown that children from homes in which a language other than English is the primary language are at risk for poor reading outcomes (Denton, West, & Walston, 2003; National Center for Education Statistics, 2003; Snow, Burns, & Griffin, 1998; U.S. Department of Education, 2000). For example, the results of the Head Start Families and Children Experiences Survey (FACES) demonstrated that Spanish-speaking Head Start and English-proficient bilingual children began kindergarten with language and literacy abilities that were below age expectations for monolingual children (U.S. Department of Health and Human Services [DHHS], 2003). Unfortunately, these discrepancies typically continue as children progress through the educational system. Reports have documented that Latino children read below the mainstream population in grades 4, 8, and 12 (U.S. Department of Education, 2000). As a result, there is a critical need to improve our understanding of factors that influence the literacy outcomes of children who are acquiring Spanish and English.

Oral language is one key factor that has been studied in many investigations of monolingual children's reading abilities, but has received relatively little attention in investigations of the reading development of young bilingual children living in the United States. Because of this fact, we review the research on monolingual children before turning our attention to work that has been conducted on bilingual children.

ABSTRACT: **Purpose:** The purpose of this study was to investigate the relationship between Head Start children's receptive language development and their kindergarten reading outcomes.
Method: Eighty-eight bilingual children who were eligible to attend Head Start for 2 years participated in the study. Growth curve models were used to examine the relationship between children's language abilities during 2 years in Head Start and end-of-kindergarten reading outcomes.
Results: The results revealed that children's English and Spanish receptive language abilities increased during Head Start, and children's early reading abilities in English were within the typical range of monolingual norms at the end of kindergarten. Children's early reading abilities in Spanish were nearly 1 *SD* below the test mean or lower. The results also showed that children's growth in their English and Spanish language abilities during Head Start predicted their early reading abilities in English and Spanish.
Implications: The findings imply that preschool programs are needed that target children's *growth* in language and not their performance measured at a particular point in time. Also, the results demonstrate the importance of early and regular evaluation of bilingual children's development in both languages in order to monitor children's growth in their two languages.

KEY WORDS: bilingual children, receptive language, reading development, Head Start

The Relationship Between Monolingual Children's Language and Early Reading Outcomes

Numerous research studies have shown that the language abilities of monolingual children are good predictors of reading outcomes (see Scarborough, 2001, for a review). The reasons for this relationship are complex but may be due to differences in children's world knowledge, language experiences, and access to lexical items when reading (Dickinson & McCabe, 2001). A meta-analysis of reading research identified 20 studies that found significant correlations between kindergartners' language abilities and reading outcomes (Scarborough, 2001). Researchers who study children with language impairment have found that this relationship between the components of language (i.e., vocabulary, syntax, as defined by decoding and reading comprehension) continues in early elementary school grades (cf. Catts, Fey, Zhang, & Tomblin, 1999; Joanisse, Manis, Keating, & Seidenberg, 2000; Tomblin, Zhang, Buckwalter, & Catts, 2000).

Studies that have investigated the relationship between monolingual preschoolers' language abilities and their early reading outcomes have focused on concurrent and predictive relationships. In general, studies have found that preschoolers' oral language abilities were positively related to their emergent reading outcomes in preschool. Specifically, Dickinson and McCabe (2001) found that preschoolers' receptive vocabulary correlated with their literacy and print concepts in preschool. Also, Storch and Whitehurst (2002) found that preschoolers' oral language skills (as defined as receptive and expressive vocabulary and narrative recall) were positively correlated with their early code-related skills (e.g., phonological awareness, print principles, and emergent writing).

Differing findings, however, have occurred when the relationship between preschoolers' language abilities and later early reading abilities have been investigated. Some studies have found positive predictive relationships between language and later reading abilities. For example, Dickinson and McCabe (2001) found that preschoolers' receptive vocabulary correlated with their print-related abilities in kindergarten. Similarly, Lonigan, Burgess, and Anthony (2000) showed that children's receptive and expressive vocabulary abilities predicted later letter knowledge. More recently, the National Institute of Child Health and Human Development (NICHD) Early Childcare Research Network (2005) found that 3-year-olds' comprehensive oral language skills predicted code-related abilities, as demonstrated by phonological awareness, at 54 months. When children were older, a relationship between 3-year-olds' language abilities and letter–word identification abilities in first grade was found; however, this relationship was mediated by the children's oral language abilities at 54 months. Additionally, children's oral language abilities at 54 months were found to be directly and indirectly related to children's first-grade letter–word identification abilities.

Contrary evidence concerning the relationship between early language and later reading abilities has been provided by other sources. Storch and Whitehurst (2002) found that the relationship between language and early literacy that they observed in preschool weakened considerably when investigating the relationship between preschool language abilities and children's code-related skills in early grades. Additionally, the Head Start FACES demonstrated that Head Start children's receptive vocabulary abilities at the end of Head Start and their gains in receptive vocabulary from fall to spring of the Head Start year did not predict their early decoding abilities at the end of kindergarten (U.S. DHHS, 2003). Because of differences in measures and designs, the studies have inconsistent findings. Therefore, additional research is needed to clarify the relationships among components of language and later reading abilities. As will be seen, research on bilingual children is also needed, as relatively little knowledge exists concerning bilingual children's language and reading outcomes.

The Relationship Between Bilingual Children's Language and Early Reading Outcomes

Few studies have been conducted on the relationship between the language and early reading outcomes of bilingual preschoolers living in the United States. Tabors, Páez, and López (2003) studied relationships between 4-year-olds' abilities in Spanish and English in the following areas: phonological awareness, expressive vocabulary, letter–word identification, memory for sentences, and concepts about print. Positive correlations were found between the children's Spanish and English in all areas but expressive vocabulary, where a negative correlation between the two languages was observed. Thus, in all but expressive vocabulary, children who had strong English abilities also had abilities in Spanish. With regard to vocabulary, children with higher vocabulary skills in English tended to have lower abilities in Spanish and vice versa. In most areas, however, children's abilities in one language were related to their abilities in the other.

No other studies on bilingual preschoolers' emerging language and early reading abilities were identified in the literature; therefore, an investigation by Lindsey, Manis, and Bailey (2003) is presented to provide background information for the study reported in this manuscript. Lindsey et al. studied the predictive relationships between language and reading abilities of bilingual kindergartners and first graders from low-income homes who were enrolled in an early transitional bilingual program. The researchers found that children's English expressive vocabulary was strongly associated with their English letter–word identification abilities, and Spanish expressive vocabulary predicted Spanish letter–word identification abilities. Similar to other research on bilingual children, this research showed cross-linguistic influence of children's phonological awareness abilities but not vocabulary abilities. That is, children's English vocabulary did not predict their Spanish letter–word identification abilities, and their Spanish vocabulary did not affect their English letter–word identification abilities. As a result, Lindsey et al. concluded that language-specific relationships occur between vocabulary and reading outcomes. Additional studies, however, that investigate the contribution of bilingual children's language abilities to their reading outcomes are clearly needed.

Considerations to be Made When Studying Bilingual Children

When studying bilingual children, the children's language experiences and their educational environment need to be considered. Researchers have argued that it is important to consider children's language experiences at the time of school entry because the findings from studies of bilingual children in nonschool settings may not apply to outcomes associated with changes in the language environment due to schooling (Butler & Hakuta, 2004; Genesee, 2004; Oller & Eilers, 2002). Therefore, it is hypothesized that children's outcomes may differ depending on whether they were exposed to two languages in the home from birth or they were exposed to Spanish at home and English at the time of school entry.

Additionally, the information about the educational environment must be documented so that results of studies can be interpreted appropriately. In particular, it is important to take into account the language of instruction and the type of educational program the children attend because the Spanish and English reading outcomes of children who receive instruction in Spanish may differ from those of children who receive instruction in English (August & Hakuta, 1997).

Purpose of the Investigation

The purpose of this study was to investigate bilingual Head Start children's receptive language development and reading outcomes. The following questions were examined:

- Are there differences between the overall receptive language abilities and early reading abilities of children who were exposed to Spanish and English at home from birth and those who were not exposed to English until age 3 when they attended Head Start?
- Do children's receptive language abilities in Spanish and English during Head Start predict their reading outcomes at the end of kindergarten?

METHOD

Participants

Eighty-eight bilingual children who attended Head Start programs in urban centers in Central Pennsylvania participated in this investigation. In order to participate in the study, children had to qualify financially for Head Start services for 2 years, have a mother who spoke the Puerto Rican dialect of Spanish, have no parent or teacher concerns about their development, score within the typical range on the Denver II (Frankenburg et al., 1990), and pass a hearing screening that was administered by Head Start nurses.

The children were divided into two groups based on information obtained from the children's mothers on home visits. Mothers were asked to report the ages at which children were spoken to and expected to communicate in Spanish and English. Children who were talked to in two languages in the home from birth were classified as having home English communication (HEC, n = 53). Children who were talked to in Spanish in the home from birth and were not expected to communicate in English until entry into Head Start at age 3 were considered as having school English communication (SEC, n = 35). It is possible that the children in the SEC group may have had some exposure to English before entrance into Head Start; however, children in this group were not expected to communicate in English on a regular basis until they entered Head Start (Hammer, Miccio, & Rodríguez, 2004; Kohnert, Bates, & Hernández, 1999).

At the beginning of the study, children in the HEC and SEC groups averaged 3;9 (years;months) in age, with standard deviations of 4.1 months and 4.4 months, respectively. Children attended Head Start classrooms in which English was the primary language of instruction. Children in the SEC group typically were placed in classrooms in which either the teacher or the classroom assistant spoke Spanish; however, English was the language of instruction, and use of Spanish in the classrooms was minimal. When children entered kindergarten, literacy in English was the targeted goal of the schools the children attended. The schools did not focus on Spanish language and literacy development.

Procedure

Children's receptive language abilities were assessed in the fall and spring of their 2 years in Head Start by trained examiners. Bilingual examiners who were native speakers of Puerto Rican Spanish assessed the children in Spanish. Native speakers of English tested the children in English. Children's receptive vocabularies were assessed in English using the Peabody Picture Vocabulary Test—III (PPVT–III; Dunn & Dunn, 1997) and in Spanish using the Test de Vocabulario Imágenes Peabody (TVIP; Dunn, Lugo, Padilla, & Dunn, 1986). Children's oral language comprehension was tested in English with the Receptive Language subtest of the Test of Early Language Development—3 (TELD–3; Hresko, Reid, & Hammill, 1999) and in Spanish with the Auditory Comprehension subtest of the Preschool Language Scale—3, Spanish version (PLS–3; Zimmerman, Steiner, & Pond, 1992).

Children's emergent reading abilities were tested in the spring of the children's kindergarten year using the Test of Early Reading Ability—2 (TERA–2; Reid, Hresko, & Hammill, 1991). The TERA–2 contains items that tap knowledge of contextual meaning, knowledge of the alphabet and its functions, and print knowledge. A comparable version in Spanish was not available. In addition, the Letter–Word Identification subtests of the Woodcock-Muñoz Language Proficiency Battery—Revised (WLPB–R; Woodcock & Muñoz-Sandoval, 1995) were administered in Spanish and English. The Letter–Word subtest assesses children's abilities to name letters and decode single words.

It should be noted that none of the tests used in the study provides normative information for bilingual children. The PPVT–III, TELD–3, and TERA–2 were developed for monolingual English-speaking populations. The TVIP contains normative data for monolingual Puerto Rican children, and the PLS–3 was based on a monolingual Spanish-speaking sample. The percentage of bilingual children included in the normative sample of the WLPB–R is unclear. These tests were used because no tests that have been standardized on bilingual populations were available at the time of the study. The tests that were employed were not used for diagnostic purposes but were used to describe changes in children's development. The internal consistency reliability coefficients for the tests used in this investigation were as follows: PPVT–III, r = .94; TVIP, r = .93; TELD–3, r = .91; PLS–3, r = .81; TERA–2, r = .91; and WLPB–R Letter–Word Identification, r = .96.

Analyses

Growth curve modeling, which has been used in a small but growing number of investigations on children's language and reading development (Compton, 2000; Foorman, Francis, Mehta, Schatschneider, & Fletcher, 1997; Hammer, Lawrence, & Miccio, in press; Pan, Rowe, Singer, & Snow, 2005; Speece, Ritchey, Cooper, Roth, & Schatschneider, 2004; Stage, Sheppard, Davidson, & Browning, 2001; Torgesen et al., 1999), was used to describe the children's language development. Growth curves allow the researcher to capture the shape and individual variation of development. Additionally, growth curve models are adaptable to circumstances where the participants could not all be measured at the same time. Growth curves are not used just to estimate the curve's fixed effects (i.e., effects that impact all individuals); they are also useful at parsing the unexplained variance into informative parcels—the part that occurs within a subject and the part that occurs between subjects (Lawrence & Hancock, 1998; Raudenbush, 2001)—thereby allowing us to obtain a better understanding of how participants change (See Appendix A for a more detailed discussion).

Modeling of children's development was accomplished in two stages. In the first stage, growth curves for the English and Spanish receptive language outcomes were fit using a linear mixed model (cf. Littell, Milliken, Stroup, & Wolfinger, 1996; Pinheiro & Bates, 2000; Singer, 1998). These growth curve models allowed us to determine the growth trajectories of the children's language development during Head Start. An important but unsettled issue with this type of model is the determination of statistical significance of the parameter estimates. The common approach has been to calculate a test statistic by dividing the parameter estimate by its standard error. The quotient is compared to a reference distribution, frequently the *t* distribution. At issue is what degrees of freedom should be applied to evaluate the statistic. The reason for the debate originates with the estimation method. For many modeling techniques, the degrees of freedom that are associated with a parameter estimate is clear. But for the maximum likelihood estimates, the denominator degrees of freedom comes from a penalized function of the residual and may vary with the parameter being estimated. The size of the penalty imposed for a particular parameter estimate is not a settled issue. Because we find the degree of freedom debate compelling, we chose an alternative method to determining a parameter's statistical significance. We used simulation via Markov Chain Monte Carlo (MCMC) methods to determine the distribution of our parameter estimates. Information on the shape of the distribution is used to compute confidence (credible) intervals around the estimates. We report those values for our fixed effects estimates instead of *p* values.

All confidence intervals shown in the Results section are 95% confidence intervals ($p \leq .05$), accordant with our stated Type I error rate. The estimates are considered significant if the displayed confidence interval does not contain zero, signifying that the true value is unlikely to be supportive of our null hypothesis.

The growth curve criterion measures were component scores that were formed from the two English language measures (i.e., PPVT–III and the Receptive Language subtest of the TELD–3) and the two Spanish language measures (i.e., TVIP and the Auditory Comprehension subtest of the PLS–3). These scores were created to capture the children's overall receptive language abilities by combining information about children's receptive vocabulary and language comprehension. Using components is a known and recommended method for variable reduction that is commonly used (e.g., Everitt & Dunn, 2001; Faraway, 2004; Fox, 1997; Harrell, 2001). The method was originally proposed by Pearson (1901), making it one of the oldest multivariate techniques.

The aim of creating components was to reduce the number of variables used in the analysis. The smaller set of variables or components retains most of the variation in the original set while ensuring that each of the new variables (components) is orthogonal to all others. For this project, the first principal component captured a large portion of the shared variance among the original measures. Hence, we were confident that one component for each set of language outcome measures would suffice for modeling child development in that language. The first component for the English outcome measures captured an average 97.5% of the variance in the original variables. The variance captured by the first component ranged from 96% to 99% over the four measurement occasions. For the Spanish outcome, the first component captured an average 95% of the original variance, with a range from 93% to 97%.

In the second stage, the estimated growth parameters that were calculated in the first stage were used to predict distal reading outcome scores. The models were constructed with the child's reading score at the end of kindergarten as the criterion variable and the growth model's estimated intercept and slope for each child as the predictor variables. The parameter estimates from this last linear model allowed us to make inferential statements concerning the contribution of early language abilities to later reading.

RESULTS

Children's English and Spanish Language Abilities

Table 1 displays the means and standard deviations for the children's English and Spanish language component scores during the children's 2 years in Head Start. As shown in the table, the children's average English scores were higher and the variation was larger among children in the HEC group than those in the SEC group. With regard to the children's Spanish abilities, different results were found. Children in the SEC group had higher scores, on average, but the HEC group displayed more variation in the component scores.

Children's Early Reading Outcomes

Table 2 displays the descriptive information for the distal reading outcomes according to bilingual group. The number of children with scores on the distal reading outcomes varied between 49 and 57, depending on the measure. In general, no significant differences were observed between the two groups on these measures. As the table indicates, the means for each of the two groups were within 1 *SD* of each other. Correlations among language and early reading outcomes are presented in Appendix B.

Table 1. Descriptive information for language component scores.

	Bilingual group			
	HEC		*SEC*	
Measurement occasion	M	SD	M	SD
English				
Time 1	26.17	13.77	17.21	8.82
Time 2	32.34	16.24	25.18	12.02
Time 3	44.35	15.63	36.20	13.60
Time 4	55.38	16.30	49.34	12.40
Spanish				
Time 1	21.65	7.52	27.15	7.04
Time 2	24.88	11.28	31.56	9.22
Time 3	32.30	10.90	39.27	6.14
Time 4	31.70	10.08	39.85	10.14

Note. HEC = home English communication; SEC = school English communication.

Table 2. Descriptive information for reading outcomes in kindergarten.

Bilingual group	*English LW-Raw*	*English LW-SS*	*Spanish LW-Raw*	*Spanish LW-SS*	*TERA Raw*	*TERA SS*
Means						
HEC	17.41	101.18	11.39	87.25	24.12	98.65
SEC	16.74	101.05	9.57	80.76	21.75	93.35
Standard deviations						
HEC	4.24	13.00	4.09	16.26	5.84	14.65
SEC	4.12	11.27	3.85	16.87	6.20	15.28

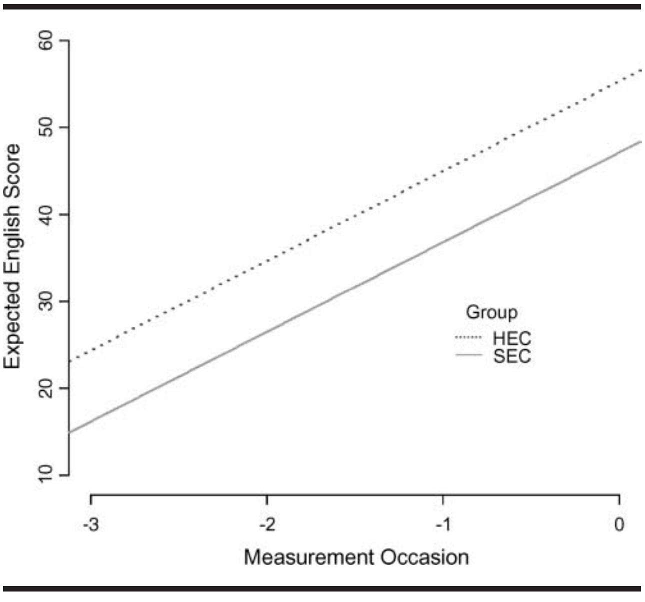
Note. LW-Raw = Letter–Word Identification raw score; LW-SS = Letter–Word Identification standard score; TERA = Test of Early Literacy Development–2.

Children's English Language Abilities and Early Reading Outcomes

The model used for the English component model of receptive language was a random slope growth model with the intercept placed at the last measurement occasion (i.e., the end of Head Start; see Figure 1). Coding time in this fashion provided positive estimates for linear rate-of-change. Use of the random slope model implied significant between variation in linear rates of change. Besides the intercept and slope, the model contained an effect for bilingual status.

Table 3 displays the confidence intervals for the parameter estimates from the English receptive score model. The results revealed that the children exhibited a strong linear rate-of-change ($\beta = 10.3$, $p < 0.05$) in their English receptive language abilities during their 2 years in Head Start. Furthermore, the model demonstrates that children in the SEC group scored, on average, approximately 8 units below their counterparts in the HEC group at the end of Head Start ($\beta = -7.68$, $p < 0.05$). The model also shows that the trajectories of the children in the SEC group were, on average, parallel to the HEC trajectories during the children's 2 years in Head Start (see Figure 1). In other words, children in both groups demonstrated positive rates of growth during their preschool years; however, the children in the HEC group scored consistently higher than those in the SEC group.

Figure 1. Growth trajectories for English receptive language component scores.



The second stage of the modeling predicted distal outcomes (i.e., reading scores at the end of kindergarten) using the parameter estimates from the English growth model. Table 4 shows the results from that modeling effort. The parameter labels in the table show the effect of the expected scores for HEC and SEC at the end of Head Start and the effect of the linear rate-of-change on each of the six reading outcomes at the end of the children's kindergarten year. Specifically, the lines in Tables 4 and 6 labeled "HEC" determine if the children's English receptive language scores at the end of Head Start predict their reading outcomes at the end of kindergarten. The lines labeled "SEC" tell whether children in the SEC group performed significantly different from those in the HEC group. The "slope" lines indicate if the linear rate-of-change in the children's receptive language abilities during Head Start significantly impacted the children's later reading outcomes. The intercept is at the end of Head Start, and the slope is the linear rate-of-change in child language. The slope is not interacted with the grouping variable; therefore, the slope effect is not group specific. Rather, the slope reflects the average linear rate-of-change in language ability during the Head Start years.

Because the statistical tests shown in this table are related, the tolerance for Type I error was adjusted accordingly. Therefore, we compared the calculated p values listed in the table to an adjusted Type I error rate of 0.008. The results revealed that the children's English language scores at the end of Head Start did not predict their early reading abilities at the end of kindergarten, with the exception of the standard scores of the Spanish version of the Letter–Word Identification subtest of the WLPB–R. The linear rate-of-change of the children's English receptive language abilities during Head Start, however, did predict the children's early reading abilities. The coefficients associated with the linear rate-of-change

Table 3. English component trajectory.

Confidence interval/pparameter*	*Lower limit 2.5%*	*Estimate 50%*	*Upper limit 97.5%*
(Intercept)	50.73	55.04	59.54
Time	9.41	10.30	11.22
Bilingual status: SEC	–13.68	–7.68	–1.34

*If the confidence interval (i.e., range between the lower and upper limits) does not contain 0, the level of significance is ≤.05.

Table 4. Parameter estimates for English component growth parameters predicting reading outcomes.

Parameter	*Estimate*	*Standard error*	T *value*	Pr(>|t|)
English Letter–Word Identification raw scores				
HEC	–0.075	0.097	–0.773	0.445
SEC	0.016	0.170	0.094	0.926
Slope	2.079	0.536	3.875	0.000*
English Letter–Word Identification standard scores				
HEC	–0.773	0.290	–2.668	0.012
SEC	–0.160	0.508	–0.315	0.755
Slope	13.926	1.606	8.672	0.000*
Spanish Letter–Word Identification raw scores				
HEC	–0.143	0.086	–1.661	0.104
SEC	0.139	0.160	0.867	0.391
Slope	1.877	0.477	3.939	0.000*
Spanish Letter–Word Identification standard scores				
HEC	–1.122	0.339	–3.310	0.002*
SEC	0.586	0.633	0.926	0.360
Slope	14.534	1.880	7.730	0.000*
TERA–2 raw scores				
HEC	–0.110	0.144	–0.762	0.451
SEC	0.185	0.251	0.737	0.466
Slope	2.919	0.800	3.648	0.001*
TERA–2 standard scores				
HEC	–0.911	0.364	–2.507	0.017
SEC	0.428	0.832	0.678	0.503
Slope	14.419	2.015	7.157	0.000*

*$p \leq .008$, the adjusted Type I error rate.

during Head Start were positive for all outcomes, indicating that children who demonstrated higher rates of learning during Head Start, regardless of bilingual status, performed better on all future reading tasks, including the TERA–2 and the English and Spanish Letter–Word Identification subtests of the WLPB–R. The children's bilingual status, that is, being in the HEC or SEC group, did not affect the predictive relationship between their language growth and early reading outcomes.

Children's Spanish Language Abilities and Early Reading Outcomes

The parameter estimates for the Spanish language component model are provided in Table 5. The model was a random slope model with the intercept, once again, located at the last measurement occasion (see Figure 2). The linear rate-of-change of children's

Table 5. Spanish component trajectory.

Confidence interval/parameter*	*2.5% Lower limit*	*50% Estimate*	*97.5% Upper limit*
(Intercept)	31.05	33.93	36.96
Time	3.12	4.11	5.00
Bilingual status: SEC	2.61	6.46	10.36

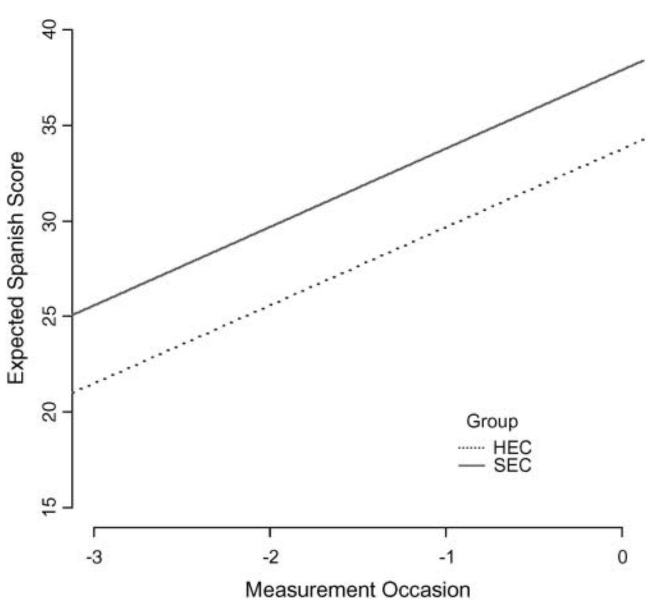
*If the confidence interval (i.e., between the lower and upper limits) does not contain 0, the level of significance is ≤.01.

Table 6. Parameter estimates for Spanish component growth parameters predicting reading outcomes.

Parameter	*Estimate*	*Standard error*	T *value*	Pr(>|t|)
English Letter–Word Identification raw scores				
HEC	–1.115	0.276	–4.045	0.000*
SEC	–0.099	0.209	–0.474	0.639
Slope	13.432	2.309	5.816	0.000*
English Letter–Word Identification standard scores				
HEC	–5.998	0.758	–7.912	0.000*
SEC	–0.002	0.574	–0.004	0.997
Slope	74.107	6.351	11.668	0.000*
Spanish Letter–Word Identification raw scores				
HEC	–0.693	0.262	–2.647	0.011
SEC	–0.278	0.183	–1.517	0.137
Slope	8.499	2.195	3.872	0.000*
Spanish Letter–Word Identification standard scores				
HEC	–5.450	1.060	–5.144	0.000*
SEC	–1.037	0.741	–1.398	0.170
Slope	66.344	8.885	7.467	0.000*
TERA–2 raw scores				
HEC	–1.430	0.403	–3.545	0.001*
SEC	–0.354	0.302	–1.172	0.249
Slope	17.668	3.380	5.227	0.000*
TERA–2 standard scores				
HEC	–6.212	0.948	–6.549	0.000*
SEC	–0.775	0.710	–1.092	0.283
Slope	75.240	7.947	9.468	0.000*

*$p \leq .008$, the adjusted Type I error rate.

Figure 2. Growth trajectories for Spanish receptive language component scores.



Spanish receptive language abilities was positive (β = 4.1, $p < 0.05$) over the 2-year period, with children in the SEC group having higher scores than children in the HEC group at the end of Head Start and over the 2-year period (β = 6.6, $p < 0.05$).

Table 6 shows the results for the second stage modeling using the Spanish language model parameter estimates as predictor variables for the children's reading outcomes. As in the English component models, the linear rate-of-change from the Spanish growth curve models was the best predictor of future reading performance. Higher linear rates of change in children's Spanish language abilities during Head Start resulted in higher reading scores at the end of kindergarten, as measured by the TERA–2 and the English and Spanish Letter–Word Identification subtests of the WLPB–R. The Spanish models, however, differed from the English models in that the children's Spanish language abilities at the end of Head Start were useful in predicting all distal reading outcomes except the raw scores on the Spanish Letter–Word Identification subtest. In each of these circumstances, higher Spanish language scores at the end of Head Start indicated lower reading scores at the end of kindergarten. There was no evidence to indicate that the ability of the children's Spanish scores at the end of Head Start to predict future reading ability was moderated by the bilingual status.

DISCUSSION

This study investigated the predictive relationships among bilingual children's receptive language abilities in English and Spanish during Head Start and children's emergent reading outcomes at the end of kindergarten. Specifically, the abilities of children who had been exposed to English in the home before entry into Head Start (HEC) and who had not been expected to communicate in English until they entered Head Start (SEC) were examined.

Results revealed that children's English and Spanish receptive language abilities increased throughout their 2 years in Head Start, with no differences being observed between the two groups in terms of their rates of change. Overall, children in the HEC group had higher English skills, and children in the SEC group had higher Spanish skills. This is not surprising given that the children in the HEC group communicated in English on a consistent basis before attending Head Start. Also, maternal reports of language usage in the home revealed that children in the HEC group were exposed to English more commonly than were children in the SEC group.

The early English reading abilities of children in the HEC and SEC groups were within 1 *SD* of the test means of the TERA–2 and the English Letter–Word Identification subtest of the WLPB–R, meaning that the children were performing within the expectations set for monolingual children at the end of kindergarten. Children's letter–word identification abilities in Spanish were not as well developed as their English abilities. Although this result is disappointing, it is not unexpected because the children were not receiving literacy instruction in Spanish in the schools. Some children did receive support for their Spanish literacy in the home; however, both children and parents most likely understood that the schools placed emphasis on English outcomes.

This investigation also demonstrated that *growth* in children's English receptive language abilities during Head Start, as opposed to the level of English they had achieved by the end of Head Start, positively predicted the children's emergent reading abilities in English and the children's ability to identify letters and words in English. This was the case regardless of the level of the children's prior exposure to English. These English early reading results complement the findings of Dickinson and McCabe (2001) and the NICHD Early Child Care Research Network (2005), who found a relationship between language and reading outcomes of monolingual kindergarten children, and Lindsey et al. (2003), who found significant relationships between bilingual children's English expressive vocabulary and English letter–word identification abilities. Our finding, however, is inconsistent with the results of the Head Start FACES, which did not identify a positive impact of the gains made by bilingual children in their English language abilities on children's early reading outcomes (U.S. DHHS, 2003). There are two possible reasons for this. First, our analyses took into account children's rates of change, whereas FACES examined gains in children's scores. Second, the children in our study attended Head Start for 2 years, which allowed us to look at their development during this time, whereas FACES examined children's development over a 1-year period. Additionally, the majority of children in Head Start attend the program for 1 year. The children in our study participated in Head Start for 2 years. It may be that attendance in Head Start for 2 years has more beneficial effects on children's development than does 1 year.

Similar to the English language results, children's growth in their Spanish language abilities predicted their early Spanish reading abilities, as measured by the Spanish Letter–Word Identification subtest of the WLPB–R. Thus, growth in children's understanding of Spanish had a positive impact on their Spanish outcomes. This is true whether they came from homes where they were exposed to English before Head Start (HEC) or from homes where they were exposed only to Spanish (SEC). Contrary to the English language results, children's Spanish comprehension at the end of Head Start had a negative impact on the children's early Spanish reading abilities, as measured by their standard scores, but not raw scores, on the Spanish Letter–Word Identification subtest. This may be because children's Spanish receptive language abilities at the end of Head Start were not sufficient to support their early Spanish reading abilities in comparison to the monolingual Spanish-speaking population upon which the test was normed.

This investigation also found that *changes* in children's English language abilities during Head Start predicted their ability to identify letters and words in Spanish and English, and *growth* in their Spanish language abilities predicted their early Spanish and English reading abilities. This is a key finding that indicates that growth in *either* Spanish or English language development during the preschool years results in positive reading outcomes in kindergarten.

This result, however, contradicts the findings of Lindsey et al. (2003), who concluded that relationships between expressive vocabulary and early reading abilities were language specific. Our study differs from Lindsey et al.'s in at least three ways. First, Lindsey and her colleagues investigated predictive relationships at older ages (i.e., prediction of first-grade reading outcomes from kindergarten abilities). As discussed in the introduction of this article, research on monolingual children has shown that the relationship between language and literacy abilities differs at various ages. Second, we investigated *growth* in children's development as well as the impact of children's language abilities at a particular point in time on later reading abilities, whereas Lindsey et al. focused on a particular point in time. Use of growth curves allowed us to study the rate of change in children's language abilities over a 2-year period, which provides a broader picture of children's language development. Cross-sectional data capture a snapshot of children's abilities and do not reflect children's growth in an area. Third, Lindsey et al. studied relationships between children's expressive vocabularies and reading outcomes, whereas we used a component score that included children's receptive vocabulary and language comprehension. It may be that a more global measure of language reflects an underlying proficiency with language that then impacts children's reading abilities in both languages.

This investigation also demonstrated that children's English language comprehension abilities at the end of Head Start positively impacted children's standard scores on their Spanish letter–word identification abilities and that children's Spanish language comprehension scores at the end of Head Start negatively predicted their emergent reading scores and English letter–word identification abilities. As mentioned previously, the end of Head Start scores represent a discrete point in children's development and do not reflect the growth that is occurring over the 2 years in Head Start. For example, a child who scores low on an assessment instrument may be developing language at an accelerated rate. A discrete measurement taken at the end of Head Start does not reflect the growth in language abilities that the child is experiencing. Therefore, we assert that results that take into account 2 years of information about children's language development, and in particular, rate of change, are more meaningful.

Implications for Research and Practice

The findings of this investigation have several implications for the early education of, and research on, bilingual children from low-income families. First, children's English language abilities developed in a positive direction during the preschool years regardless of whether they were exposed to English before or after entrance into Head Start. This indicates that children experienced positive growth in their English language abilities regardless of their early home language environment. Growth occurred between the spring and fall testing points, which included the summer months when children did not attend the English-speaking Head Start programs and when children spent more time at home and in their neighborhoods, during which time their exposure to Spanish may have been greater. This result, in tandem with the finding that growth in Spanish and English language supported children's early reading outcomes, should minimize the concern of those who believe that lack of knowledge of English before entry into Head Start is detrimental to children.

Second, attendance in an English-speaking preschool and amount of exposure to Spanish in the home did not appear to support children's Spanish language and literacy abilities. With regard to children's language abilities, although children's raw scores on Spanish language measures increased, examination of standard scores revealed that the children's Spanish comprehension abilities and receptive vocabulary were not keeping up with monolingual Spanish norms (Hammer, Lawrence, & Miccio, in press). With regard to their Spanish literacy abilities, children's early ability to identify letters and words were 14 to 19 points below their English abilities in this area, on average. Once again, this implies that children's exposure to Spanish literacy at home and school was not sufficient to fully support children's Spanish early reading abilities to the level of monolinguals. If the goal of an educational program is to develop children's abilities in Spanish and English, then increased support of children's Spanish literacy abilities is needed during the preschool years. We assert that the goal of bilingualism is desirable given the cognitive, cultural, and economic benefits of being bilingual (Bialystok, 2001).

Third, the results demonstrate that if bilingual children's language growth is progressing well in *either* Spanish or English during the preschool years, positive early English and Spanish reading outcomes result in kindergarten. This result has significant implications for educational programming. Currently, efforts are commonly made to push Spanish-speaking children to acquire English as quickly as possible. Many Head Start and preschool programs immerse children in English-language classrooms upon entry into the preschool. The results of this longitudinal study indicate, however, that it is the rate of language change that is important, not the immediate acquisition of English. Therefore, preschool programs are needed that target children's progress in acquiring language and not their performance as measured by a score on a test at a particular time. It may be possible to develop high-quality language interventions that change the developmental trajectories of bilingual children during the preschool years. Targeting children's language development is critical given the links between language and literacy that have been shown in the research literature. If the trajectories of children's language development are changed, then a positive impact on children's early reading outcomes should occur. This, in turn, would allow children to build on their early reading abilities as they progress through school. Research has demonstrated that reading deficits in early grades are extremely difficult to overcome and place children at great risk for school failure and school dropout (cf. Snow et al., 1998; Stanovich, 2000). High-quality preschool programs that target children's language acquisition may ultimately be less expensive and less intensive than remedial programs that are implemented later in children's academic careers once significant deficits in literacy have developed.

Fourth, these results necessitate that bilingual children's language abilities be evaluated early and monitored regularly. When concerns arise about language abilities, it is common practice to give bilingual children time in an educational program before making a referral. This is appropriate for many children; however, if concerns continue and a referral is made, it is likely that insufficient data on children's language development has been collected up until this point. This lack of information may make it difficult for speech-language pathologists (SLPs) to make an appropriate diagnostic decision. Therefore, educational programs need to collect baseline data on children's abilities *in both languages* upon school entry *and* continue to collect data periodically in order to determine if growth is occurring in both languages. Our results demonstrate that assessment at a particular point in time does not provide sufficient information about children's language development, and in fact may not provide accurate information. Recall, for example, that children's English language abilities at the end of Head Start did not predict their early reading outcomes in both languages (with the exception of standard scores on the Spanish Letter–Word Identification subtest), but language growth during Head Start did. Therefore, information that is gathered over multiple time points is needed to determine whether growth in language is occurring. This complements arguments made by proponents of dynamic assessment (cf. Lidz & Peña, 1996), who argue that data that are collected at a single point do not supply the practitioner with meaningful information about the process of language development. Without information collected over time, it is difficult to determine if a bilingual child's development is progressing normally, at a faster rate of development or, most importantly for SLPs, at a slower rate of development. Regular monitoring of bilingual children's language development will provide SLPs with valuable information that will assist them in determining whether or not children require intervention.

Finally, given the fact that growth in children's English and Spanish language abilities had a positive impact on children's English and Spanish reading abilities, we advocate for additional research in this area. In particular, longitudinal studies are needed that investigate children's developmental trajectories and their relationships with later outcomes. Cross-sectional studies that examine children's abilities at particular points in time may provide insufficient information about bilingual children's language development and the relationship between language and reading outcomes because the process of development is not taken into account. Additionally, prevention studies are needed that implement high-quality programs and investigate whether the magnitude of children's developmental trajectories are affected by the intervention. Typically, studies assess the language abilities of children pre- and posttreatment. This may not be sufficient to understand the developmental phenomena that are occurring as a result of the program that is implemented. In particular, studies involving bilingual and dual language programs are needed to determine which programs work for which children, as it is unlikely that one model will address the needs of all bilingual children. Such investigations will provide great insights as to which programs maximize bilingual children's language and reading growth and later outcomes.

ACKNOWLEDGMENTS

This study was supported by a grant from the National Institutes of Health, National Institute of Child Health and Human Development, and the United States Department of Education, Institute of Education Sciences (5-R01-HD-39496-05). The authors wish to thank the parents and children who participated in the investigation and the staff of the Head Start programs and kindergartens for their support and assistance with the project. In addition, the authors thank the project staff and the graduate and undergraduate research assistants from The Pennsylvania State Department of Communication Sciences and Disorders for their assistance with the project.

REFERENCES

August, D., & Hakuta, K. (Eds.). (1997). *Improving schooling for language-minority children: A research agenda*. Washington, DC: National Academy Press.

Bialystok, E. (2001). *Bilingualism in development*. New York: Cambridge University Press.

Butler, Y., & Hakuta, K. (2004). Bilingual and second language acquisition. In T. Bhatia & W. Ritchie (Eds.), *The handbook of bilingualism* (pp. 114–144). Malden, MA: Blackwell.

Catts, H., Fey, M., Zhang, X., & Tomblin, J. B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading, 3,* 331–361.

Compton, D. L. (2000). Modeling the growth of decoding skills in first grade children. *Scientific Studies of Reading, 4,* 219–259.

Denton, K., West, J., & Watson, J. (2003). *Reading—Young children's achievement and classroom experiences* (Rep. No. NCES2003-070). Washington, DC: U.S. Department of Education National Center for Education Statistics.

Dickinson, D., & McCabe, A. (2001). Bringing it all together: The multiple origins, skills, and environmental supports of early literacy. *Learning Disabilities Research and Practice, 16,* 186–202.

Dunn, L. W., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test—III*. Circle Pines, MN: AGS.

Dunn, L. W., Lugo, D. E., Padilla, E. R., & Dunn, L. M. (1986). *Test de vocabulario en imágenes Peabody*. Circle Pines, MN: AGS.

Everitt, B. S., & Dunn, G. (2001). *Applied multivariate data analysis* (2nd ed.). London: Arnold.

Faraway, J. J. (2004). *Linear models with R*. Boca Raton, FL: Chapman & Hall.

Foorman, B. R., Francis, D., Mehta, P., Schatschneider, C., & Fletcher, J. (1997). Early interventions for children with reading disabilities. *Scientific Studies of Reading, 1,* 255–276.

Fox, J. (1997). *Applied regression, linear models, and related methods*. Thousand Oaks, CA: Sage.

Frankenburg, W. K., Dodds, J., Archer, P., et al. (1990). Denver—II: Screening manual. Denver, CO: Denver Developmental Materials.

Genesee, F. (2004). What do we know about bilingual education for majority-language students? In T. Bhatia & W. Ritchie (Eds.), *The handbook of bilingualism* (pp. 547–576). Malden, MA: Blackwell.

Hammer, C. S., Lawrence, F. R., & Miccio, A. W. (in press). Exposure to English before and after entry into Head Start: Bilingual children's receptive language growth in Spanish and English. *International Journal of Bilingual Education and Bilingualism*.

Hammer, C. S., Miccio, A. W., & Rodríguez, B. (2004). Bilingual language acquisition and the child socialization process. In B. Goldstein (Ed.), *Bilingual language development and disorders in Spanish–English speakers* (pp. 21–52). Baltimore: Brookes.

Harrell, F. E. (2001). *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis*. New York: Springer.

Hresko, W., Reid, D. K., & Hammill, D. (1999). *Test of Early Language Development—3*. Austin, TX: Pro-Ed.

Joanisse, M., Manis, F., Keating, P., & Seidenberg, M. (2000). Language deficits in dyslexic children: Speech perception, phonology, and morphology. *Journal of Experimental Child Psychology, 77,* 30–60.

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

Lawrence, F. R., & Hancock, G. R. (1998). Assessing change over time using latent growth modeling. *Measurement and Evaluation in Counseling and Development, 30*(4), 211–224.

Lidz, C., & Peña, E. (1996). Dynamic assessment: The model, its relevance as a nonbiased approach, and its application to Latino American preschool children. *Language, Speech, and Hearing Services in Schools, 27,* 367–372.

Lindsey, K., Manis, F., & Bailey, C. (2003). Prediction of first-grade reading in Spanish-speaking English-language learners. *Journal of Educational Psychology, 95,* 482–494.

Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). *SAS system for mixed models*. Cary, NC: SAS Institute.

Lonigan, C., Burgess, S., & Anthony, J. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology, 26,* 596–613.

Muthen, L. K., & Muthen, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling Journal, 9,* 599–620.

National Center for Education Statistics. (2003). *Status and trends in the education of Hispanics* (Rep. No. NCES 2003-008). Washington, DC: U.S. Department of Education.

National Institute of Child Health and Human Development Early Child Care Research Network. (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology, 41,* 428–442.

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

Pan, B., Rowe, M., Singer, J., & Snow, C. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development, 76,* 763–782.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine, 2,* 559–572.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-plus*. New York: Springer.

Raudenbush, S. W. (2001). Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual Review of Psychology, 52,* 501–525.

Reid, D., Hresko, W., & Hammill, D. (1991). *Test of Early Reading Ability—2*. Austin, TX: Pro-Ed.

Scarborough, H. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory and practice. In S. Neuman & D. Dickinson (Eds.), *Handbook of family literacy* (pp. 97–110). New York: Guilford Press.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 24,* 323–355.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Speece, D., Ritchey, K., Cooper, D., Roth, F., & Schatschneider, C. (2004). Growth in early reading skills from kindergarten to third grade. *Contemporary Educational Psychology, 29,* 312–332.

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

Stanovich, K. (2000). *Progress in understanding reading*. New York: Guilford Press.

Storch, S., & Whitehurst, R. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38,* 934–947.

Tabors, P. O., Páez, M., & López, L. (2003). Dual language abilities of bilingual four-year olds: Initial findings from the early childhood study of language and literacy development of Spanish-speaking children. *NABE Journal of Research and Practice, 1,* 70–91.

Tomblin, B., Zhang, X., Buckwalter, P., & Catts, H. (2000). The association of reading disability, behavioral disorders, and language impairment among second-grade children. *Journal of Child Psychology and Psychiatry, 41,* 473–482.

Torgesen, J., Wagner, R., Rashotte, C., Lindamood, P., Rose, E., Conway, T., & Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities. *Journal of Educational Psychology, 91,* 579–593.

U.S. Department of Education. (2000). *The condition of education 2000*. Washington, DC: U.S. Government Printing Office.

U.S. Department of Health and Human Services. (2003). *Head Start FACES 2000: A whole-child perspective on program performance*. Washington, DC: Author.

Woodcock, R. W., & Muñoz-Sandoval, A. F. (1995). *Woodcock Language Proficiency Battery—Revised: Spanish Form*. Itasca, IL: Riverside.

Zimmerman, I., Steiner, V., & Pond, R. (1992). *Preschool Language Scale—3* (Spanish Edition). San Antonio, TX: The Psychological Corporation.

Received October 15, 2005
Revision received May 19, 2006
Accepted January 26, 2007
DOI: 10.1044/0161-1461(2007/025)

Contact author: Carol Scheffner Hammer, The Pennsylvania State University, Department of Communication Sciences and Disorders, 110 Moore Building, University Park, PA 26802. E-mail: cjh22@psu.edu.

APPENDIX A. GROWTH CURVE MODELING

Representing and measuring change is a fundamental concern to all who study and attempt to understand child development. Unfortunately, the study of change is not as transparent as we might hope. To demonstrate that change has occurred, the researcher is required to use a longitudinal research design.

Regardless of the type of investigation (i.e., true, quasi-, or nonexperimental), the application of the longitudinal design to the study of change offers several challenges with regard to data analysis. These challenges arise because the responses to the variables of interest are not independent, as is required for most statistical procedures; instead, the responses are unarguably correlated. Further complicating the analysis are the research questions. Because the questions posed by researchers are varied, the methods applied to the data must also be multifaceted.

Various procedures exist for the analysis of longitudinal data. Their purpose is to quantify change. Among those procedures are repeated measures analysis of variance, autoregressive time-series models, and latent variable growth models. The different procedures make differing assumptions about the data and may yield quite different results when they are applied to the same data. Choosing the appropriate method for the analysis requires the researcher to understand the nature of the data, the specific research question, and the assumptions made by the various analytic techniques.

For many developmental research questions, the researcher would like to summarize the sample data. To be complete, the summary would contain an expected trajectory for each participant as well as report the variability in trajectories. Besides the summary, the research questions tend to focus on identifying causal factors. Therefore, the modeling scheme should be flexible enough to allow the researcher to test hypotheses about the effects on intercepts (initial scores) and slopes (rates of change) produced by differences in exogenous variables. For those researchers most interested in this type of question, the latent variable growth curve model provides answers.

The latent variable growth curve model provides a means for analyzing repeated observations made over a finite time period. Some of the strengths of the latent variable growth curve model include the capacity to test different forms for growth, to examine the effects of time-varying covariates, to summarize the observations in the form of an expected growth trajectory, and to quantify the amount of variation about that expected trajectory. Limitations include sample size, scale of measurement for the response, and reliability of the measures. Generally, the sample size required to execute the growth curve model is dependent on the number of within-subject observations as well as the number of subjects. When the number of observations within subjects is small (e.g., three or less), it is difficult to conclude that the growth form is other than linear. A large number of representative subjects is required to accurately assess the variability that exists in the growth form over individuals (viz. Muthen & Muthen, 2002). Finally, the reliability of the measures used will influence the models' ability to replicate the data.

Growth curve analyses generate "true scores," which are preferable to observed scores. (Note: Observed scores are reported in analyses such as repeated measures of analysis of variance). Observed scores are fallible. That is, each observed score is subject to measurement error. The statistical interpretation of measurement error is that it is random and non-zero for an individual observation but that it balances itself out across many observations. Thus, we say that on average, measurement error is zero, knowing that in a specific instance, it is rarely zero. Because each observation is measured with some amount of error that is random for any single observation, the notion put forth by statisticians is that replicate observations will tend to have a certain amount of error but that the average over the observations [replications] will represent a score without error, a true score. Hence, the mean is considered a "true" score, whereas individual observations used to compute the mean each contain some amount of error.

Statisticians assess the precision of a parameter estimate in terms of sampling variation. Sampling variation occurs because each sample represents a subgroup of the population; that is, it is not the population but the data from the sample that are used to infer characteristics of the population. Because the sample is not the population, there is always some discrepancy that exists between the members of the sample and those in the entire population. The sampling variation captures the variability. In our illustration using the mean, we recognize that the mean of the sample is an estimate of the mean (or the parameter mu) in the population. We use the standard error of the mean to communicate the variation that exists in that estimate due to the fact that we are using a sample taken from the population, not the population, to estimate the mean value.

Sampling variation is a measure of the variability in a parameter estimate that would be found across an infinite number of resamplings from the same population. Growth curves contain many parameter estimates, but two are of primary interest—the intercept and the slope. The *intercept* conveys information about the expected initial score for the phenomena under investigation; the *slope* is the expected rate-of-change. Both of these parameter estimates exhibit sampling variation. If i indexes the individual and j indexes the time of measurement, then sampling variation in the slope estimate is sigma_sq epsilon_i/CSST_i where epsilon is the within-subject residual and CSST_i is the corrected sums of squares for the time component in the model.So in the mixed effects growth model

$$\mathrm{y\_ij} = \mathrm{pi\_0i} + \mathrm{pi\_1i(time)\_ij} + \mathrm{r\_ij}$$

$$\mathrm{pi\_0i} = \mathrm{gamma\_00} + \mathrm{u\_0i}$$

$$\mathrm{pi\_1i} = \mathrm{gamma\_10} + \mathrm{u\_1i}$$

signifying that the response (*y*) for individual *i* at time *j* is a function of the individual's expected intercept score (pi_0i) and the individual's expected rate of change (pi_1i) plus some unexplained residual (r_ij). Further, the individual expected value at intercept (time = 0) is a function of a grand mean (gamma_00), which is the average intercept for all in the sample plus some unexplained component attributable to that individual (u_0i). The slope interpretation is constructed in a similar manner. The notion of the true score is apparent in the parameter estimates gamma_00 and gamma_10. These are the expected values for intercept and slope gathered using all data in the sample. These parameter estimates represent our best judgment of the population values. Hence, they represent the true scores on those aspects of the model.

When a model is designated as "random intercept," the term, u_0i, appears in the formula for pi_0i. The term, u_1i, does not appear in the formula for pi_1i. The implication is that between-subject variation exists at time = 0, but the individual trajectories are essentially parallel. When u_1i appears in the formula for pi_1i, the model is designated as a "random slope" model. Unless specifically stated to the contrary, it is assumed that a random slope model also has a random intercept; that is, it uses the formula for the intercept (pi_0i) that contains the random component, u_0i.

APPENDIX B. CORRELATIONS BETWEEN LANGUAGE AND EMERGENT READING OUTCOMES

Table A displays all pairwise correlations for the English language component scores and the children's reading outcomes. The correlation among the English component scores is fairly stable and significant ($r \sim 0.7$, $p < 0.05$). Correlations among each of the English component scores and the reading outcomes are small and are not statistically different from zero.

Table B shows the correlations for the Spanish component scores and all reading outcomes. The correlations among the Spanish component scores range from approximately 0.4 to approximately 0.6. The correlations are significantly different from zero ($p < 0.05$) and positive. In general, the Spanish component scores do not exhibit a relationship with the distal outcomes. One exception is the Spanish measure at occasion 3. For that particular measurement, the correlations are negative, moderate (cf. Cohen, 1969), and significantly different from zero ($p < 0.05$), with two exceptions. Those two exceptions are the letter–word Spanish raw and standard score.

Table A. Correlations among English language components and reading outcomes.

	English Time 1	*English Time 2*	*English Time 3*	*English Time 4*	*TERA Raw*	*TERA SS*	*English LW-Raw*	*English LW-SS*	*English LW-Raw*	*English LW-SS*
English Time 1	****	0.775	0.710	0.723	0.207	0.078	0.152	0.098	–0.044	–0.096
English Time 2	<0.001	****	0.685	0.670	0.149	0.078	0.119	0.093	–0.074	–0.100
English Time 3	<0.001	<0.001	****	0.774	0.095	–0.002	0.123	0.085	0.148	0.086
English Time 4	<0.001	<0.001	<0.001	****	0.190	0.071	0.208	0.114	–0.029	–0.104
TERA Raw	0.213	0.373	0.569	0.254	****	0.867	0.516	0.555	0.320	0.355
TERA SS	0.643	0.643	0.992	0.673	<0.001	****	0.511	0.721	0.289	0.470
English LW-Raw	0.369	0.483	0.467	0.216	0.001	0.001	****	0.880	0.647	0.587
English LW-SS	0.565	0.583	0.615	0.501	<0.001	<0.001	<0.001	****	0.505	0.617
Spanish LS-Raw	0.772	0.631	0.331	0.852	0.050	0.079	<0.001	0.001	****	0.927
Spanish LS-SS	0.531	0.511	0.575	0.498	0.029	0.003	<0.001	<0.001	<0.001	****

Note. Upper diagonal part contains correlation coefficient estimates. Lower diagonal part contains corresponding *p* values. English = English language component score; TERA = Test of Early Literacy Development–2; LW-Raw = Letter–Word Identification raw score; LW-SS = Letter–Word Identification standard score.

Table B. Correlations among Spanish language components and English reading outcomes.

	Spanish Time 1	*Spanish Time 2*	*Spanish Time 3*	*Spanish Time 4*	*TERA Raw*	*TERA SS*	*English LW-Raw*	*English LW-SS*	*English LW-Raw*	*English LW-SS*
Spanish Time 1	****	0.631	0.451	0.613	–0.230	–0.399	–0.237	–0.331	–0.101	–0.193
Spanish Time 2	<0.001	****	0.416	0.630	–0.154	–0.266	–0.231	–0.313	–0.241	–0.322
Spanish Time 3	<0.001	0.001	****	0.415	–0.325	–0.475	–0.383	–0.461	–0.196	–0.287
Spanish Time 4	<0.001	<0.001	0.001	****	–0.181	–0.297	–0.091	–0.181	–0.149	–0.232
TERA Raw	0.171	0.363	0.050	0.284	****	0.866	0.498	0.544	0.253	0.305
TERA SS	0.014	0.111	0.003	0.074	<0.001	****	0.498	0.715	0.245	0.448
English LW-Raw	0.164	0.175	0.021	0.598	0.002	0.002	****	0.877	0.636	0.561
English LW-SS	0.049	0.063	0.005	0.292	0.001	<0.001	<0.001	****	0.492	0.606
Spanish LS-Raw	0.513	0.115	0.203	0.335	0.131	0.144	<0.001	0.002	****	0.928
Spanish LS-SS	0.209	0.033	0.059	0.129	0.066	0.005	<0.001	<0.001	<0.001	****

Note. Upper diagonal part contains correlation coefficient estimates. Lower diagonal part contains corresponding *p* values. Spanish = Spanish language component score; TERA = Test of Early Literacy Development–2; LW-Raw = Letter–Word Identification raw score; LW-SS = Letter–Word Identification standard score.