Applied Psycholinguistics **34** (2013), 673–696
doi:10.1017/S0142716411000919

Oral language skills of Spanish-speaking English language learners: The impact of high-quality native language exposure

PERLA B. GÁMEZ
Harvard Graduate School of Education

SUSAN C. LEVINE
University of Chicago

Received: September 7, 2010 Accepted for publication: June 30, 2011

ADDRESS FOR CORRESPONDENCE
Perla B. Gámez, Harvard Graduate School of Education, Larsen 316, 14 Appian Way, Cambridge, MA 02138. E-mail: perla@uchicago.edu

ABSTRACT
This study examined the relation between young English language learners' (ELL) native oral language skills and their language input in transitional bilingual education kindergarten classrooms. Spanish-speaking ELLs' ($n = 101$) Spanish expressive language skills were assessed using the memory for sentences and picture vocabulary subtests of the Woodcock Language Proficiency Battery—Revised. Samples of transitional bilingual education teachers' ($n = 21$) speech were recorded and coded for syntactic complexity and vocabulary usage. Results revealed considerable variation in ELLs' language scores, with overall performance below the normative sample. There was also wide variation in teachers' speech across classrooms. Hierarchical linear modeling revealed that gains in ELLs' expressive language skills were positively related to the diversity of teachers' vocabulary and teachers' syntactic complexity. These findings suggest that the quality of teachers' language input, not just the quantity of their input, plays a significant role in the language learning trajectories of ELLs.

A growing body of research provides evidence of a relation between language input and young children's language development. In particular, the variation in young monolingual children's oral language skills has been linked to the amount and type of input to which they are exposed in their monolingual environments (e.g., Hart & Risely, 1995; Hoff, 2003; Hoff-Ginsberg, 1998; Hoff & Naigles, 2002, Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2007; Pan, Rowe, Snow, & Singer, 2005). In light of current demographic shifts in the United States (see Capps et al., 2005), it is important to also examine the

input factors that impact the development of children who are negotiating more than one language in language-mixed environments.

A number of factors impact our motivation to understand the input features associated with robust oral language skills in dual-language learning children, particularly children who are English language learners (ELLs). First, there are approximately 5.3 million children learning English as a second language in US schools (NCELA, 2010). These ELL children are particularly concentrated in the early elementary grades, with over 44% in pre-kindergarten through third grade (Kindler, 2002), and over 70% from Spanish speaking homes (Fry & Gonzales, 2008). Moreover, ELLs constitute the fastest growing student population in the United States: from 1995 to 2005, ELL enrollment increased by more than 60%, whereas enrollment for the total student population increased by less than 3% (NCELA, 2006). Second, many ELLs have poor language and literacy outcomes (for a review, see August & Shanahan, 2006). In particular, ELLs exhibit lower than expected native (first language [L1]) and English (second language [L2]) oral language skills on average when compared to monolingual norms, although there are wide variations in skill levels (e.g., Páez, Tabors, & Lopez, 2007). The low L1 skills of many ELL children are of particular concern because early oral language skills are associated with later reading comprehension, which is often cited as the skill that matters most for academic success (National Early Literacy Panel [NELP], 2008; Snow, Burns, & Griffin, 1998).

The fact that some ELLs may have low L1 skills despite receiving significant amounts of exposure to their native language at home (Capps et al., 2005) and in some cases at school (Kindler, 2002), raises the possibility that it is the *quality* rather than the *quantity* of L1 exposure that impacts their native language development as well as their subsequent L2 learning and school success. Yet, previous research with the ELL population has focused mainly on the effects of differing amounts of L1 and L2 exposure in the home (e.g., Oller & Eilers, 2002) and at school (see review in Lindholm-Leary & Borsato, 2006), paying minimal attention to the quality of the language input this population is exposed to at home or in their language-mixed classrooms. In order to expand the existing knowledge base, our study examines the quality of the L1 language input in ELLs' language-mixed transitional bilingual education (TBE) classrooms. At the early grades, TBE classrooms provide young Spanish-speaking ELLs with instruction in Spanish, their native language, as well as in English (see Genesee, 1999, for a description of different bilingual program models). We specifically examine the Spanish syntactic complexity and vocabulary usage that teachers address to their Spanish-speaking ELLs in TBE classrooms and the relation of this input to the development of student's Spanish oral language skills.

ELLS' ORAL LANGUAGE DEVELOPMENT: WITHIN- AND BETWEEN-GROUP DIFFERENCES

A review of research on ELLs revealed that only about 25% of studies have examined ELLs' oral language development, with the vast majority focusing instead on their literacy achievement (Saunders & O'Brien, 2007). Many studies that have examined ELLs' oral language development report generally low

performance on standardized measures of L1 and L2 proficiency (Hammer, Davison, Lawrence, & Miccio, 2009; Mancilla-Martinez & Lesaux, 2011; Miccio, Tabors, Páez, Hammer, & Wagstaff, 2005; Páez et al., 2007). In particular, Páez et al. (2007) report that at the start of preschool, Spanish-speaking ELLs, on average, performed more than 2 *SD* below the national mean of monolingual Spanish speakers on standardized measures of expressive language, including sentence memory and vocabulary knowledge. Mancilla-Martinez and Lesaux (2011) replicated these findings with a subsample from Páez et al.'s (2007) study and showed that by middle school (i.e., seventh grade), their sample continued to perform poorly (>2 *SD* below the national mean) on both of these measures in Spanish. Children in the Páez ELL sample, who had been educated in English-only classrooms since school entry, scored relatively higher in English (∼1 *SD* below the national mean of monolingual English speakers).

The low oral L1 performance of many ELLs is especially noteworthy given evidence showing a link between children's early oral language skills and their later literacy outcomes (NELP, 2008; Snow et al., 1998). Most relevant to the present study are studies examining within language relations of oral language and reading skills for ELLs attending bilingual programs that provide L1 instruction. In one such longitudinal investigation of Spanish-speaking ELLs enrolled in a TBE program, the results showed that Spanish expressive language (i.e., vocabulary knowledge, sentence memory) in kindergarten was a strong predictor of Spanish reading comprehension by second grade (Manis, Lindsey, & Bailey, 2004). A follow-up study with the same Spanish-speaking ELL group revealed that native oral language skills (i.e., vocabulary, listening comprehension) in third grade predicted their Spanish reading comprehension in sixth grade (Nakamoto, Lindsey, & Manis, 2008).

In addition to within-language effects, cross-linguistic influences of L1 on L2 skills have also been documented (Genesee, Lindholm, Saunders, & Christian, 2005). First, in the studies cited above, expressive language skills were consistently correlated between the L1 and L2, and this was seen as early as kindergarten (see also Lindsey, Manis, & Bailey, 2003). Second, L1 and L2 reading comprehension also tend to be positively correlated (see also Lindsey et al., 2003; Proctor, August, Snow, & Barr, 2010). Third, there is some evidence that L1 oral language skills predict L2 reading comprehension (August & Shanahan, 2006), at least for children who receive L1 literacy instruction (August et al., 2006). For instance, the study by Manis et al. (2004) demonstrated a cross-linguistic transfer effect from L1 expressive language skills to L2 reading comprehension in second grade, with developing L2 skills mediating the L1 contribution. By early adolescence, however, the predictive value of oral language (i.e., vocabulary and listening comprehension) on reading comprehension was much stronger within language than across the two languages, likely because of the L2 literacy instruction children received once they transitioned into mainstream, English-only elementary school classrooms (Nakamoto et al., 2008). In fact, when literacy instruction occurs only in the L2, cross-language effects of L1 oral language skills and reading-related skills in the L2 are generally not found (Gottardo & Mueller, 2009; Mancilla-Martinez & Lesaux, 2010). This is in contrast to the findings of Proctor, Carlo, August, and Snow (2005) that Spanish–English biliterate fourth

graders show a positive relation between L1 vocabulary and English reading comprehension.

Findings showing a relation between L1 and L2 oral language and reading skills, as well as the potential cross-linguistic influences of L1 oral language skills on L2 reading outcomes, highlight the need to investigate the factors that promote the development of ELLs' early oral language skills in their native language. It is important to note that ELLs' demographic information may, in part, explain their low oral language skills; ELLs disproportionately live in poverty and are born to parents with low education and literacy rates (Capps et al., 2005; Fry & Gonzales, 2008; Hernandez, Denton, & Macartney, 2008). However, it is worth noting that even within this group, which carries risk factors associated with low academic success, there are wide individual differences, raising the possibility that input factors explain some of this variation.

ACCOUNTING FOR VARIABILITY IN LANGUAGE DEVELOPMENT

Monolingual environments

A growing body of research on monolingual English speakers implicates variations in linguistic input at home and at school as contributing to individual variation in children's language development (Hoff, 2006). The bulk of this research has investigated language input in the home environment, at the very early stages of language development. Although Furrow, Nelson, and Benedict (1979) found a positive relation between mothers' shorter mean length of utterances and their 18-month-olds' development of syntax, most research suggests that young children whose input consists of longer utterances evidence more advanced levels of syntactic development than those exposed to shorter utterances (e.g., Harkness, 1977, Hoff-Ginsberg, 1998). Huttenlocher et al. (2002) found a significant relation between the proportion of multiclause sentences mothers used and the production of multiclause sentences in 3- and 4-year-olds. In terms of vocabulary development, research has shown that exposure to longer mean length of utterances, richer vocabulary, and larger amounts of speech is positively related to young children's vocabulary size (e.g., Bornstein, Haynes, & Painter, 1998; Hart & Risley, 1995; Hoff & Naigles, 2002; Huttenlocher et al., 1991; Pan et al., 2005).

Little research has directly examined the specific linguistic features of teachers' speech in the classroom on school-aged children's language skills, specifically, syntax (Huttenlocher et al., 2002) and vocabulary skills (Bowers & Vasilyeva, 2011; Dickinson & Porche, 2011). Generally, in classroom studies where the focus is on the linguistic environment, the level of analysis has involved broad measures of language use, for example, amount of talk (i.e., teacher–student interactions by minute: Connor, Morrison, & Slominski, 2006), amount of instruction (i.e., in teacher-managed vs. child-managed instruction: Connor, Morrison, & Underwood, 2007), type of interaction style (i.e., didactic vs. cognitively demanding talk: Dickinson & Smith, 1991; Smith & Dickinson, 1994) or instructional moves made by the teacher (e.g., modeling: see review in Lawrence & Snow, 2011). Moreover, like the studies of young children's home environments, the few linguistically detailed classroom studies have been situated within monolingual environments.

For example, although the classroom sample in Bowers and Vasilyeva's (2011) study included ELLs, the study centered on mainstream, English-only classrooms and preschool children's English language development.

Existing studies of teacher input in monolingual, mainstream classrooms show substantial variation in the type and amount of teacher talk. Moreover, these studies show that this variation, in part, accounts for differences in children's language skills, mirroring the findings from the studies of parent input at home. Specifically, in terms of syntactic development, Huttenlocher et al. (2002) found that the proportion of multiclause utterances used by preschool teachers was positively related to growth in their preschool students' syntactic comprehension. In terms of vocabulary development, Dickinson and Porche (2011) found that preschool teachers' use of sophisticated vocabulary predicted children's later reading comprehension in fourth grade, with effects mediated by children's kindergarten vocabulary and literacy skills. Bowers and Vasilyeva (2011) also showed that the diversity of preschool teachers' words was positively related to their monolingual children's English vocabulary growth. They also found a positive association between the number of English word tokens preschool teachers produced and preschool ELLs' English vocabulary growth. However, they found that the mean number of words per utterance (i.e., a measure of structural complexity) was negatively associated with ELLs' English vocabulary growth, and not significantly associated with English-monolinguals' vocabulary growth.

In sum, the existing literature, which has primarily focused on monolingual children, indicates that the quality of the input (i.e., syntactic complexity, vocabulary diversity, and sophistication) explains significant amounts of variation in children's language development. Much less is known about how characteristics of input relate to the language development in ELL children who are acquiring multiple languages. Given the current demographic shifts, there is a pressing need to also study the relation of input to the native language development of ELL children.

Language-mixed environments

A relatively smaller, but growing body of literature also shows that the amount of early home input is related to ELLs' native language development (De Houwer, 2007; Oller & Eilers, 2002; Pearson, Fernandez, Lewedeg, & Oller, 1997). As Pearson and colleagues (e.g., Oller & Pearson, 2002; Pearson et al., 1997) explain, the amount of exposure ELLs receive in their native language at home varies substantially. Moreover, some ELL children may receive about equal amounts of exposure to two languages whereas others receive more exposure to one language than the other. Further, it has been documented that the relative amount of exposure in each language is positively related to children's language skills in that language. In one study, Pearson et al. (1997) found that for dual-language learning toddlers, the more exposure children had to a language, the larger their vocabulary size was in that language. Therefore, a factor that may explain the observed individual differences in L1 performance within the ELL group is minimal exposure to the L1.

Research in the area of language maintenance and shift among immigrant generations in the United States reveals that immigrants tend to eventually make a shift away from the native language toward English dominance (e.g., Rumbaut, Massey, & Bean, 2006; Veltman, 1983; Zentella, 1997). This shift toward English across immigrant generations is important for the present study given that the majority of children in US schools who have limited-English proficiency are US-born children of immigrants (Capps et al., 2005). As Kohnert and Pham (2010; and Kohnert, 2008) argue, factors such as participation in English-only schools, which may limit the opportunities and motivation to use the L1 outside of the home, and the low social status of minority languages, such as Spanish in the United States, may contribute to an early plateau or regression of L1 skills. Therefore, for continued L1 development, ELLs require both home support and early instruction in that language at school (see Duursma et al., 2007; Winsler, Díaz, Espinosa, & Rodríguez, 1999).

However, and in contrast to the literature base on the language development of monolingual English-speaking children, very few studies have attempted to study the quality of the input to which ELLs are exposed in their language-mixed classrooms. As noted, the majority of studies have instead focused on studying the quantity of language input on L2 development (see review in Lindholm-Leary & Borsato, 2006). In one study by Bowers and Vasilyeva (2011) that measured the structural complexity of teacher speech, low levels of structural complexity, albeit in English, were found to be beneficial for preschool ELLs; but this contrasts with the bulk of research on monolinguals, which shows a positive relation between young children's language skills and the complexity of their speech input (e.g., Hoff-Ginsberg, 1998; Huttenlocher et al., 2002). This discrepancy warrants further examination of the features of language input that are associated with ELLs' language development.

THE PRESENT STUDY

In the present study, we examine the variation in ELL children's language skills and variation in their language input in TBE classrooms, a common type of bilingual program in the early grades (Genessee, 1999; Kindler, 2002; Moughamian, Rivera, & Francis, 2009). Specifically, we conducted extensive observations in 21 TBE kindergarten classrooms and analyzed audio recordings of classroom teachers' speech to assess their syntactic complexity and vocabulary usage. We also documented ELLs' native, expressive language skills at the beginning and end of their kindergarten year using two subtests of the Spanish version of the Woodcock Language Proficiency Battery—Revised (WLPB-R; Woodcock & Munoz-Sandoval, 1995). The WLPB-R's memory for sentences and picture vocabulary subtests have been previously used to evaluate ELLs' expressive language skills and children's scores have been shown to predict their reading skills (e.g., Manis et al., 2004). We relied on hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002) to examine the impact of teachers' Spanish language input on the growth of their ELL students' Spanish oral language skills over the school year.

Table 1. *English language learners' home language input*

	Mean	*SE*	*n*
Language spoken to the child by the *mother*	1.31	0.06	81
Language spoken to the child by the *father*	1.60	0.09	76
Language spoken to the child by *other adults*	1.54	0.09	75
Language spoken to the child by *older siblings*	2.14	0.13	54
Language spoken to the child by *younger siblings*	1.73	0.10	58
Language spoken to the child by *friends*	2.29	0.11	80

METHOD

Participants

Students. One hundred one ELL kindergarteners (54 girls, 47 boys) were included in the present study. The mean age of children in our sample was 5.64 years ($SE = 0.04$ years) at pretest and 6.12 years ($SE = 0.04$ years) at posttest. All children were Latino and native Spanish speakers, classified as limited English proficient according to school administered language proficiency tests (e.g., pre-Language Assessment Scale; Duncan & DeAvila, 2000). Based on these scores, the children received language support services in the form of native language instruction, with 0.5 hr per school day devoted to English instruction. According to parental report at the end of the school year, children in this sample had limited experience speaking English outside of school, and spoke Spanish for 5.10 years ($SE = 0.13$, number of responses = 73) and English for 1.07 years ($SE = 0.12$, number of responses = 73). The amount of time spent in English corresponds to the time that had elapsed since their enrollment in TBE classrooms.

To obtain a measure of our sample's home language exposure, parents were asked to fill out the DELLSS Project Parent Questionnaire in Spanish. As in previous studies (i.e., Duursma et al., 2007), parent responses to the questions about language use and exposure at home were used to gauge Spanish and English input at home. Each question is answered on a 5-point scale where low values indicate a preference for Spanish, high values indicate a preference for English, and intermediate values indicate equal preference for the two languages. As can be seen in Table 1, children in our sample are primarily spoken to in Spanish in the home environment by parents, other adults, siblings, and friends.

Teachers. Twenty-one teachers (20 females, 1 male) in classrooms serving the 101 participating children were included in this study. Teachers were asked to fill out a survey containing questions related to their education, teaching experience, and demographic information. Two teachers did not return the survey; therefore, detailed information is provided for 19 of the 21 teachers. According to the survey report, the mean age of teachers in this sample was 39.47 years (range = 24–65 years, $SE = 4.41$). All were Spanish–English bilingual Latino/as and had taught for a mean of 12.29 years (range = 1.00–32.00 years, $SE = 1.81$).

All teachers reported holding a bachelor's degree with 12 also holding a master's degree. Nine teachers reported having a regular certificate for teaching in Illinois, 3 reported having a Transitional Bilingual Certificate only, and 7 reported having both.

Classrooms. The participating children and teachers were situated within 21 half- and full-day kindergarten classrooms with one teacher per classroom. Class sizes ranged from 16 to 29 children, with a mean of 21 children per classroom (mean number of girls = 10.52, mean number of boys = 10.62). The total number of children across all classrooms was 444 (221 girls, 223 boys).

All classrooms were labeled as early-TBE by principals and teachers of each school. As has been documented in prior research, the native language is the primary language used for instruction during the early grades in early-TBE programs (e.g., Lindsey et al., 2003). English language development (ELD) instruction is also included daily in TBE classrooms (generally 30–45 min/day). As noted, the structure of ELD instruction can vary (Genesee, 1999; Moughamian et al., 2009) such that some TBE classrooms include a separate block of ELD devoted specifically to instruction of English oral language, and others integrate ELD with other instruction. The intended structure of ELD for this sample was a separate block of instruction.

However, two full-day classroom observations revealed that approximately half of our classrooms included an ELD block ($n = 11$), whereas the other half did not ($n = 10$). In classrooms that did not include a separate block of ELD instruction, "morning meeting time" was designated as a time when teachers incorporated some oral language instruction in English. However, and as we confirm in the results section of this paper, Spanish was primarily spoken during morning meeting time, even in the nonseparate-block classrooms. Given the differences in the structure of ELD across our sample of classrooms (separate English block vs. nonseparate English block), this variable was included as a control in our HLM models. Our decision to control for differences in program structure was guided by the fact that ELD structure varies across TBE classrooms (Genesee, 1999) and has been shown to be associated with ELLs' language and literacy outcomes, albeit in English, on the WLPB-R (Saunders, Foorman, & Carlson, 2006). Specifically, these studies show higher English oral language scores for children in which ELD instruction is delivered in a separate instruction block.

Schools. The participating classrooms were recruited from 12 urban, public elementary schools in the midwestern United States. According to the state's interactive report card, all schools served a primarily Latino student population (mean = 94.70% Hispanic, $SD = 3.16$; range = 89.70%–99.20%). The average percentage of low-income students was high at 89.32% ($SD = 11.44$), ranging from 70.30% to 99.50%. At the time of data collection for this study, the mean number of students meeting or exceeding state assessment standards in reading was 57.78% ($SD = 9.01$; range across schools = 40.70%–70.60%).

Measures

Oral language assessments. Children's expressive language was assessed using the picture vocabulary and memory for sentences subtests of the WLPB-R, Spanish version (Woodcock, 1991). The WLPB-R is a standardized assessment consisting of a series of subtests that measure language and literacy skills. The specific details of these subtests can be found in the WLBP-R examiner's manual (1991), and only a brief description is provided here. In the memory for sentences subtest, children are asked to listen to and repeat sentences exactly. Although this subtest taps children's short-term verbal memory and word knowledge, it also targets syntactic skills as the test items increase in syntactic complexity, such that initial items are simple sentences and later items include multiple clauses (Francis et al., 2006). The picture vocabulary subtest measures children's word knowledge. Beginning test items measure receptive ability, but the majority of items (and all later items) measure expressive vocabulary.

Both subtests yield raw scores that can be converted to age-based standard scores and *W* scores. The standard scores are norm-referenced scores that are useful for describing children's relative standing within their age groups (mean = 100, $SD = 15$). In contrast, the *W* score is useful for reporting growth in the measured skills (Woodcock, 1991). The *W* scale is a transformation of the Rasch model based on item response theory (Jaffe, 2009), and *W* scores are on an equal-interval scale that takes variations in item difficulties into account. For 6-year-olds, the mean *W* score for the norming sample was 479 ($SD = 15.4$) for memory for sentences and 475 ($SD = 12.3$) for picture vocabulary.

The Spanish version of the WLPB-R was normed on 3,911 monolingual Spanish-speakers from the United States, Costa Rica, Mexico, Peru, Puerto Rico, and Spain. Thus, when interpreting WLPB-R score results from our ELL sample, we are making comparisons to Spanish monolingual norms. We rely on the WLPB-R because there are no standardized assessments normed on bilingual populations. It is important to note that although comparing our sample to monolingual norms is not optimal, it can inform us about how ELLs' languages develop within an educational system in which such standardized assessments carry great weight.

Classroom observations. Live classroom observations were conducted and detailed context notes were taken. During observation, teachers' speech was audiorecorded with an Olympus digital voice recorder (WS-300M) with an omnidirectional tie-clip microphone and videotaped using a Canon 12X Optical Zoom camcorder.

Procedure

Student sampling. Given that the unit of analysis was the teacher, we randomly sampled only a few children from each classroom to serve as representatives of classroom performance (see Raudenbush & Bryk, 2002). To do so, parental consent forms were distributed to all children in each participating classroom. Using an Internet-based random number generator (http://www.random.org; accessed October 2008), of the 238 returned signed consent forms, 111 were selected at

random to participate in the study. Attrition was low with less than 10% of the original sample either transferring schools ($n = 3$), refusing to participate ($n = 2$) or not completing posttest assessments ($n = 5$). The final sample was distributed in the following way: 17 classrooms = 5 children, and 4 classrooms = 4 children.[1]

Oral language assessments. The WLPB-R measures were given to children at two different time points, at the beginning (September to October) and end (May to June) of the 2007–2008 academic year. The assessments were conducted on a one-on-one basis at the children's school, usually in the hallway. Following the WLPB-R testing rules, individual subtests were administered in a fixed order: memory for sentences and picture vocabulary.

Each student's raw score on the WLBP-R memory for sentences and picture vocabulary subtests from the fall and spring administrations were converted to age-based standard scores and *W* scores using the Woodcock Compuscore and Program Profiles software program. The *W* scores on each subtest were used to compute spring minus fall difference scores to measure the gains exhibited over the school year.

Classroom observations. Observations, during which teacher speech was audiorecorded, were conducted at the middle (Observation I: February and early March; third quarter) and end (Observation II: late April and May; fourth quarter) of the academic year. Teachers' speech was also videotaped, but only during the second observation; during the first observation we were unable to videotape because of restrictions placed on us by our Internal Review Board.

Given that we were primarily interested in native oral language development, we focused our observations on a period of the school day during which teachers' offered specific oral language instruction in Spanish. In this sample, teachers and experimenters designated morning meeting time as providing specific Spanish oral language instruction. Morning meeting time was a daily classroom activity that occurred at the very beginning of the school day during which time children came together as a whole group and sat on a rug, akin to "circle time" in preschool classrooms. This period of the school day consisted of morning preparation activities and rituals (attendance, homework collection, etc.) and calendar activities, such as reviewing the days of the week, months, and weather, and counting aloud. Spanish oral language instruction focused on learning and rehearsing vocabulary, letter names, and sounds, and exercising speaking and listening skills in talking about the day's agenda and previous homework assignments, and so forth. By using this period of the school day to assess teachers' language input, we reduced the random variation in input that may occur around different contexts such as during literacy instruction. Past research has shown that whole group activities provide reliable indicators of teachers' input (e.g., Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006).

The number of minutes spent in morning meeting time varied across teachers, but was not significantly different across the two observation time points: $t\,(20) = -1.06$, $p = .300$; Observation I mean = 24.73, $SD = 10.75$; Observation II mean = 26.73, $SD = 9.95$). Thus, we computed the variable minutes of instruction as an average of the two time points for each teacher, and we used this average time as

a control in our main analyses. The start and end times were based on the whole group activity and did not include any independent-student or "center" work. That is, the start time began when teachers directed their speech to the entire class and ended when the whole group disbanded, for example, when teachers called the last students to their independent seats or center stations.

Speech transcription and coding. For transcription purposes, a modified version of the CHAT conventions of the CHILDES system (MacWhinney, 2007) was used. Conversational turn-taking and pauses were used to identify utterances; however, short word-searching pauses were not counted as separate utterances. To ensure transcription consistency across the two observations, only the speech samples obtained from the digital audio recording device were used in the present study; the video images were not used.

Once teachers' speech was transcribed, it was coded for language as (a) Spanish, (b) English, or (c) code-mixed. The frequency (freq) command was then used in CLAN (MacWhinney, 2007) to extract the number of words (i.e., tokens) and diversity of words (i.e., unique types) teachers used. In deciding what constituted a word type and token, we referenced the previous study of teacher input with ELL children (Bowers & Vasilyeva, 2011), and similarly excluded sounds (e.g., "baah," "beh"), letters of the alphabet (e.g., c@l, e@l), and interjections (e.g., "uhoh"). We also did not include proper names (i.e., roll call). Syntactic complexity was coded at the utterance level by assigning each utterance a code of either simple or complex. To do so, each utterance was first coded as zero, one clause, or multiclause. In accordance with Vasilyeva, Waterfull, and Huttenlocher (2008), zero-clause utterances do not contain a verb and contain either a noun (book), a pronoun (you), a proper noun/name (Ms. Gonzalez), a noun phrase (my book), a preposition (down), a prepositional phrase (in my desk), an adjective (big), an adjective phrase (very good), or an interjection (thank you). In addition, we coded utterances that consisted solely of letter names (e.g., "a") and sounds (e.g., ah) as zero clauses. One-clause utterances were coded as such if they contained one clause, even if the subjects were only implied (i.e., verb phrase). Utterances that contained more than one clause were coded as multiclause utterances. In determining the number of clauses, serial verb constructions (e.g., go get it), modals (e.g., going to do it), emerging modals (e.g., going to, have to), and tags (e.g., isn't it?) were not treated as forming separate clauses, whereas utterances containing verbs that take sentential complements (i.e., let, want, need, have) were coded as separate clauses when a subject was included (e.g., I want you to listen). In line with previous studies of teachers' syntactic complexity (i.e., Huttenlocher et al., 2002), those utterances containing only one clause were then coded as simple utterances and utterances containing multiple clauses were coded as complex. Therefore, utterances coded as complex included coordinated/compound sentences in which two or more independent clauses were joined with a coordinating conjunction (e.g., and, but), as well as complex sentences that contained dependent/subordinate clauses.

Given our focus on children's native, Spanish language development, we analyzed only teachers' speech in Spanish. Relatedly, we chose to exclude all code-mixed utterances from our analyses, as they were relatively infrequent in our transcripts ($<1\%$ of the total number of utterances). Thus, each teacher

received the following language input scores: the number of Spanish (a) word types, (b) word tokens, (c) complex utterances, and (d) "other" utterances. Teachers' syntactic complexity was further transformed into a proportion score by dividing the number of Spanish complex utterances by the total number of Spanish utterances.

RESULTS

Our primary research question addressed the relation between student vocabulary and syntactic gains in relation to particular aspects of their teachers' language input. To answer this question, we used HLM (version 6.08; Raudenbush & Bryk, 2002) as it accounts for the nesting of students within classrooms and because it allows us to control for other variables known to be associated with skill development. Specifically, given prior research that shows the benefits of blocked English language instruction on children's oral language skills (i.e., Saunders & O'Brien, 2006), we controlled for ELD program structure (i.e., blocked vs. integrated) in our models to account for differences in the way in which English oral language development instruction was carried out across our classrooms. We also included the variable minutes of instruction to account for the variation in the number of minutes of instruction observed across classrooms. To further clarify any observed impact of teacher input on student gains, we also examined the relation between student initial scores and teacher language input. As a precursor to these HLM analyses, we examined descriptive statistics for the variables we measured. We also conducted analyses of teachers' input as a guide to building our final models for the HLM analyses.

Descriptive statistics and preliminary analyses

ELL language skills. We first examined children's standard scores at both time points to situate our sample's relative standing in relation to national norms. Although not typically seen with standard scores, at least for monolingual children, we found significant growth in children's expressive language skills in Spanish. That is, children's mean standard scores on memory for sentences was higher at posttest (mean = 87.41, $SD = 9.62$) than pretest (mean = 81.18, $SD = 11.68$), $t\,(100) = -5.841$, $p < .001$, confidence interval (CI) = −8.343, 4.113. Similarly, children's mean standard score on picture vocabulary was significantly higher at posttest (mean = 89.93, $SD = 25.0$) than pretest (mean = 74.78, $SD = 18.03$), $t\,(100) = -8.495$, $p < .001$, CI = −18.686, −11.611. It is important to note that there was variability within our sample in standard score gains, with some children exhibiting substantial growth from pre- to posttest, and others exhibiting minimal or no gain, or even a decrease in scores (mean $\text{growth}_{\text{Memory for Sentences}} = 6.23$, $SD = 10.71$; mean $\text{growth}_{\text{Picture Vocabulary}} = 15.15$, $SD = 17.92$). Despite the gains over the school year, however, the mean performance level at both time points was below the national population mean of 100.

To evaluate our sample's actual progress on both skills measured, we also examined students' *W* scores at the beginning and end of the school year. As Table 2 shows, there was wide variation in scores across our sample. In addition, Pearson correlations showed modest, but significant associations between

Table 2. *Descriptives*

Variable Name	*N*	Mean	*SD*	Min	Max
Level 1 Statistics					
Expressive language difference score	101	13.25	9.88	−10.00	43.00
Memory for sentences difference score	101	9.31	10.91	−13.00	46.00
Picture vocabulary difference score	101	17.20	14.16	−11.00	47.00
Pretest *W* scores					
Memory for sentences	101	455.92	11.96	423.00	484.00
Picture vocabulary	101	447.58	15.20	408.00	491.00
Posttest *W* scores					
Memory for sentences	101	465.23	9.90	444.00	491.00
Picture vocabulary	101	464.78	20.55	430.00	506.00
Level 2 Statistics					
Spanish types	21	322.52	101.75	95.00	525.00
Spanish tokens	21	1130.00	477.97	211.00	2054.50
Spanish complex utterances	21	32.50	21.40	5.50	89.50
Spanish "other" utterances	21	220.62	68.18	122.50	350.00
Spanish total utterances	21	253.12	80.24	128.00	383.50
Program structure	21	0.52	0.51	0.00	1.00
Minutes of instruction	21	25.73	9.42	12.70	47.23

children's *W* scores on memory for sentences and picture vocabulary at pretest ($r = .382$, $p < .001$) and at posttest ($r = .502$, $p < .001$).

In addition, we examined the factorability of the two subtests. First, the two subtests were significantly correlated with each other at pretest ($r = .382$, $p < .001$) and posttest ($r = .502$, $p < .001$). Second, the Kaiser–Meyer–Olkin measure of sampling adequacy was .500 and Bartlett's test of sphericity was significant: pretest, χ^2 (1) = 15.559, $p < .001$; posttest: χ^2 (1) = 28.564, $p < .001$. Given these overall indicators, a principal components factor analysis of the two subtests was conducted and we found that children's *W* scores on the two subtests of the WLBP-R loaded onto one factor, explaining 83.1% of the variance at pretest and 86.7% at posttest. Thus, the WLPB-R subtests were reduced to one variable: expressive language. To do so, we averaged the children's *W* scores at each time point.

Teacher input characteristics. Next, we examined our teachers' language input characteristics. We found that across both observations, the majority of teachers used primarily Spanish for instruction during morning meeting time. Sixteen teachers used Spanish in over 75% of their speech, whereas the remaining 5 teachers used approximately equal amounts of English and Spanish (mean English utterances = 57%).

Our analyses revealed substantial variation across classrooms in terms of teachers' Spanish language input (see Table 2 for a summary of descriptive statistics). Of note, paired samples *t* tests were used to compare teachers' input measures between Observation I and Observation II and revealed no significant differences for any of our Spanish input measures; that is, in terms of teachers' complex utterances, $t\,(20) = -1.074$, $p = .296$, CI $= -10.788, 3.454$, "other" utterances, $t\,(20) = 0.264$, $p = .795$, CI $= -30.953, 39.905$; total utterances, $t\,(20) = 0.043$, $p = .966$, CI $= -38.749, 40.368$; word types, $t\,(20) = -0.967$, $p = .345$, CI $= -52.027, 19.074$; and word tokens, $t\,(20) = -0.693$, $p = .496$, CI $= -242.155, 125.107$. Similarly, the amount of instructional time spent at the two observations did not significantly differ, $t\,(20) = -1.064$, $p = .300$, CI $= -5.922, 1.922$. Given the lack of significant differences between the two observations, we computed teachers' speech scores as an average of the two time points.

Further, Pearson correlations revealed a positive association between teachers' complex utterances and the number of total utterances ($r = .655$, $p = 0.001$, $n = 21$), underscoring the need to control for number of utterances in our models. To do so, and avoid the endogeneity between two variables, we controlled for teachers' "other" utterances in our main analyses. There was also a very high correlation between teachers' types and tokens ($r = .976$, $p < .001$, $n = 21$), and thus, in subsequent analyses, we use only teachers' word types (i.e., diversity of words) to reduce collinearity.

HLM analyses: Relation between teacher input and student gains

Description of models fitted. We fit two-level HLMs to examine ELLs' expressive language, where the *W* score difference score was the Level 1 outcome variable and Level 2 variables included classroom-level predictors. We first examined the extent of variation at Level 1 (child) and Level 2 (classroom) using students' difference *W* score as the outcome variable in a fully unconditional model. Results showed significant variability in posttest scores among students within classrooms, $\chi^2\,(20) = 96.56$, $p < .001$; deviance $= 721.76$. We then calculated the intraclass coefficient (ICC) to be able to assess the percentage of variance to be explained at the classroom level. The ICC was 0.44, revealing that a significant proportion of the variance in scores could potentially be explained by teachers' input at the classroom level. Thus, we proceeded to build full, two-level HLMs by adding our teacher/classroom-level predictors at Level 2.

Given that our main teacher input variables were situated at both the utterance and word level, we built the following two separate models: utterance-level and word-level models. Separating the models in this way allowed us to maintain an acceptable ratio between the number of predictors and observations (Field, 2005). Specifically, in the utterance-level model we added teachers' number of Spanish complex utterances as the main predictor, along with the number of Spanish "other" utterances. In the word-level model, we added teachers' number of Spanish types as the main predictor and included two control variables to both models: minutes of instruction and ELD program structure. All predictors were grand mean centered for all models, except for ELD program structure given its dichotomous nature. See Appendix A for an example model.

Table 3. *Results of hierarchical linear modeling on English language learners' expressive language gains: Teacher utterance-level input characteristics*

Final Estimation of Fixed Effects	Coefficient	*SE*	*T* Ratio	*df*
Intercept, γ_{000}	13.13**	1.26	10.402	16
Spanish complex utterances, γ_{001}	0.22*	0.07	3.118	16
Spanish other utterances, γ_{010}	−0.01	0.02	−0.393	16
ELD program structure, γ_{020}	−2.49	2.69	−0.926	16
Minutes of instruction, γ_{030}	−0.16	0.16	−0.986	16
Deviance = 712.48				

Note: ELD, English language development.
*$p < .05$. **$p < .01$.

Table 4. *Results of hierarchical linear modeling on English language learners' expressive language gains: Teacher word-level input characteristics*

Final Estimation of Fixed Effects	Coefficient	*SE*	*T* ratio	*df*
Intercept, γ_{000}	13.14**	1.41	9.297	17
Spanish word types, γ_{001}	0.03*	0.02	2.117	17
ELD program structure, γ_{010}	−1.15	2.86	−0.402	17
Minutes of instruction, γ_{020}	−0.19	0.17	−1.126	17
Deviance = 717.13				

Note: ELD, English language development.
*$p < .05$. **$p < .01$.

Utterance-level teacher input. As Table 3 shows, the results of an utterance-level model showed that when controlling for Spanish "other" utterances, ELD program structure, and minutes of instruction, there was a significant and positive relation between our main teacher input predictor Spanish complex utterances and ELLs' difference *W* score (see Table 3 for a summary of results). This model explained 25% of the variance between classrooms; however, there still remained significant variation among class mean scores to be explained, χ^2 (16) = 55.35, $p < .05$. To maximize practical interpretability, we interpret our results in terms of effect sizes (Cohen, 1988).[2] That is, while controlling for all other variables, an increase of 1 *SD* in the average teacher's number of complex utterances would be associated with a change of 0.477 *SD* of achievement gain, which Cohen (1988) called a medium effect. In other words, for the average teacher who increases the number of Spanish complex utterances from 33 to about 54 (i.e., 1 *SD*), it is expected that there would be a gain of about 4 points on the *W* scale for expressive language.

Word-level teacher input. As Table 4 shows, the results of our word-level model revealed that, controlling for ELD program structure, and minutes of instruction,

there was a significant and positive relation between our main teacher predictor Spanish types and ELLs' difference *W* score. This model explained 25% of the variance between classrooms; however, there still remained significant variation among class mean scores to be explained, χ^2 (17) = 77.29, $p < .05$. Again, we report effect sizes in the form of coefficients ($d = 0.309$). That is, while controlling for all other variables, for the average teacher, an increase from about 322 to about 423 types in Spanish would result in a gain of about three points on the *W* scale for expressive language. It should be noted that an HLM analysis using Spanish tokens was also conducted with very similar results obtained, coefficient = 0.007, $SE = 0.003$; t (17) = 2.456, $p = .025$.

Follow-up analyses: Relation between teacher input and student initial scores

As noted, to further clarify the role of input in children's development of expressive language skills, we tested whether teacher input was related to ELLs' initial scores. If there is not a significant relation between teacher input and student's initial scores, then this would support the view that teacher input is a critical factor in *growth* of skills. To investigate this issue, we fit two-level models where students' fall pretest score was the Level 1 outcome variable and Level 2 (teacher/classroom) variables included only our main teacher input predictors (word-level: Spanish types; utterance-level: Spanish multiclause utterances, Spanish "other" utterances). Results of an unconditional model showed significant variability at the child-level, χ^2 (20) = 34.69, $p < .05$; the ICC was 0.12. The results of two full HLMs showed no significant association between children's expressive language scores at pretest and our two main teacher input predictors: Spanish complex utterances: coefficient = −0.157, $SE = 0.141$, t (18) = −1.112, p value = .281, and Spanish types: coefficient = −0.030, $SE = 0.027$, t (19) = −1.100, p value = .286.

DISCUSSION

The number of ELLs in US schools is increasing rapidly, thereby changing the linguistic diversity of the US student population. Despite the well-documented relation between later literacy outcomes and early oral language skills (e.g., NELP, 2008; Snow et al., 1998), as well as the cross-linguistic relationship between L1 and L2 skills (e.g., Manis et al., 2004; Proctor, August, Carlo, & Snow, 2006), few studies have examined the input factors impacting the early oral language skills of ELLs in their native language. The present study explored the characteristics of the Spanish that teachers provided to Spanish-speaking students in TBE classrooms, and examined the relation of these characteristics to the students' native Spanish oral language skills. To do so, we measured ELLs' memory for sentences and vocabulary knowledge (i.e., expressive language) at the beginning and end of their kindergarten year. We also observed naturalistic classroom practices, at the middle of the school year, in 21 TBE classrooms and coded teachers' speech to obtain measures of their syntactic complexity and vocabulary use.

ELLs' native language skills and language input in TBE

We found that the children in our sample exhibited substantial gains in Spanish over the kindergarten school year. That is, scores were higher at posttest than pretest even in comparison to the normative sample. This progress is in line with previous research with ELLs, which reports rapid growth when ELLs are at low levels of proficiency in their first (e.g., Mancilla-Martinez & Lesaux, 2011) and second language (see Hakuta, Butler, & Witt, 2000; Medina & Escamilla, 1992). However, despite significant gains in L1 skills across these kindergarten classrooms, results showed that our sample arrived at school and ended the kindergarten year with Spanish language levels well below monolingual norms, mirroring the results of previous studies with dual language learners (Hammer et al., 2009; Miccio et al., 2005; Páez et al., 2007).

It is worth noting that, unlike the ELLs in the present study, the dual language learners in these previous studies were enrolled in English-speaking classrooms, which likely minimized their exposure to the L1, and in turn, their opportunities to use and learn the L1 (Kohnert, 2008; Kohnert & Pham, 2010). In addition, as mentioned, the pattern of low performance among ELLs, at least in vocabulary, appears to be a property of dual language learning that has been termed "distributed characteristic" (Oller & Pearson, 2002). Without translation equivalents in a dual language learners' two languages, words may be distributed across the two languages, making words available in only one of the languages and not the other (i.e., "singlet" vocabulary). According to this view, it would be expected that the language scores of ELLs would differ from that of monolinguals', especially in the case of vocabulary. In addition, our samples' lower than average L1 performance may be partly explained by demographic information. That is, given that the participating schools in the present study served a high percentage of students from low socioeconomic status backgrounds, our sample likely carried significant risk factors for low academic performance, including living in poverty (Fry & Gonzales, 2008) and being born to parents with low education and literacy rates (Capps et al., 2005). Nevertheless, that our ELL sample exhibited low L1 performance, despite their extensive exposure to the L1 in school and at home, indicates that without intensive supports that are substantially greater than those that are currently provided, ELLs may continue to lack the foundation for academic oral language skills that are required for school success.

With this in mind, our results also document wide variation in teachers' language use in TBE classrooms. That is, there was substantial variation in individual teacher's Spanish syntactic complexity and vocabulary usage and individual teachers were consistent in these input measures across our two observations. However, it is worth noting that our two teacher observation time points were relatively close (third and fourth quarters of the school year), and this may have decreased our ability to detect changes in teacher input. If we had observed the teachers at the beginning and end of the school year differences might have emerged.

Relation between language input and variability in language skills

The primary aim of the present study was to examine whether teacher's language input was related to ELLs' gains in native oral language skills, specifically,

expressive language, over the kindergarten year. We found that child gains in expressive language were, indeed, significantly and positively related to particular characteristics of teachers' speech, that is, teachers' Spanish word types and complex utterances. These results were found when controlling for teachers' total amount of talk—their Spanish utterances—and other differences across classrooms, such as the structure of the bilingual program and the length of time spent on specific oral language instruction. Of note, the level of children's language skills at the start of the school year was not significantly related to teacher speech measures, which adds some support to the hypothesis that teacher language characteristics are influencing the gains in children's language over the school year.

The results from the present study are novel in that they extend findings of language input effects to language-mixed classrooms. Our findings on the impact of teachers' vocabulary are in line with the extensive literature on vocabulary input at home, in particular, studies showing effects on young children's vocabulary over time (Huttenlocher et al., 1991; Pan et al., 2005), and the literature on vocabulary input at school showing effects on preschoolers' vocabulary (Dickinson & Porche, 2011). The findings on the impact of teachers' syntactic complexity also replicate studies of young children's home environments that show a positive association between caregivers' structurally complex speech and children's vocabulary development (Hoff-Ginsberg, 1998), and studies in mainstream monolingual-English classrooms, for example, Huttenlocher et al.'s (2002) that show a relation between teachers' syntactic complexity and children's syntactic skill.

The significant relation between teachers' syntactic complexity and their ELL students' native oral language skills contrasts with Bowers and Vasilyeva's (2011) finding that the structural complexity of teacher speech did not benefit monolingual preschoolers' vocabulary skills in their first language. Perhaps the difference in results between the two studies was due to differences in how syntactic complexity was operationalized. Although we coded teachers' use of multiclause utterances in a manner similar to Huttenlocher et al. (2002; i.e., multiclause utterances), Bowers and Vasilyeva (2011) measured structural complexity in terms of the mean number of words per utterance. As the authors pointed out, although the mean number of words per utterance is positively correlated with the use of multiclause sentences, given that the number of words increases as the number of clauses increase, it is only a rough measure of structural complexity. It is possible that some teachers with a high mean number of words per utterance embellished their sentences with extra lexical items, yet constructed simple sentences, instead of multiclause sentences. In contrast, the teachers in our study who scored high on syntactic complexity were rated as such because they used a greater number of complex utterances by way of combining multiple clauses. Moreover, consistent with the assumptions of the syntactic bootstrapping hypothesis (Landau & Gletiman, 1985), the stringing together of multiple clauses may be what benefits children's language learning, such that children can use the syntactic structures that surround words to learn the meanings of those words. That is, complex sentences may provide children with multiple sources of information about how words are arranged in sentences, as well as with multiple clues as to the meanings of words.

Limitations and future directions

The present study has some limitations. Specifically, the participant sample used in this study was relatively small compared to the sample sizes generally used in research focusing on broad program comparisons for ELL children (i.e., Ramirez, Pasta, Yuen, Billings, & Ramey, 1991). However, the present study provides findings from a larger sample than is the norm for studies using detailed linguistic analyses (see review in Ellis, 2009). On a related note, given our interest in the children's native language development, we focused on a sample of teachers' Spanish input. As noted, we disregarded teacher input in English, which amounted to approximately half of the total amount of speech for five teachers and much less for other teachers. In the present study, we were unable to examine teachers' English input because so few teachers used English during the time sampled (and half did not include a separate block of ELDinstruction). Future studies would do well to examine the effect of dual language input on ELLs' language development when both English and the native language are used for instruction. That type of study may best be targeted to later grades, when early TBE programs may include more English instruction, or to other kinds of early bilingual program models in which the intended curriculum calls for more equal exposure to the L1 and L2 (e.g., two-way or dual immersion programs).

Further, we focused on instruction during morning meeting time, a time designated by teachers to provide specific oral language instruction (i.e., listening, speaking, learning and reviewing vocabulary, etc.). We focused on this period of the school day to obtain a context when teachers addressed the entire class and when we would reliably sample teachers' naturalistic language use and not other forms of input such as book reading. Studying teachers' speech during literacy instruction, however, may have revealed more sophisticated vocabulary and discussion about words and text meaning, as has been found for monolingual children in school and at home (Dickinson & Smith, 1991; Weizman & Snow, 2001). In fact, given the complex language patterns children are exposed to during reading (e.g., Hoff-Ginsberg, 1991), in future studies, we plan to extend our research to examine teachers' syntactic complexity, and use of sophisticated vocabulary during literacy instruction. In doing so, we plan to also focus on teachers' analytic talk about books, as this type of interactive book reading has been shown to be predictive of monolingual children's language skills (Dickinson & Porche, 2011; Dickinson & Smith, 1991, 1994; Wasik & Bond, 2001).

Finally, the correlational data used in the present study does not allow us to make a causal connection between teacher talk and growth in children's language skills over the school year. To make a casual claim, experimental studies that manipulate the frequency and complexity of talk are needed. For example, ELLs would need to be randomly assigned to treatment groups with differing language input complexity.

Despite these limitations, our findings are consistent with the possibility that classroom teachers may be able to foster the oral language skills of ELLs by increasing the diversity of words they use and the structural complexity of their speech. That is, although we also found an effect of teachers' increased use of Spanish word tokens, our findings implicate the quality of the input as playing

a significant role in the language learning trajectories of ELL children. It should be noted that although we highlight teacher input in the present study, we do not make the assumption that more teacher talk at the expense of student talk leads to optimal learning. Instead, we speculate that a more favorable learning situation is one in which teachers provide their students with a sufficient amount of high-quality speech, while also giving students ample opportunities to speak, including opportunities for peer-to-peer talk (see Mashburn, Justice, & Pianta, 2009). The findings of the present study, coupled with evidence that L1 oral language skills predict L1 as well as L2 reading comprehension, suggest that a promising way to increase ELLs' school success is to provide them with early linguistic environments that promote the development of their L1 skills.

APPENDIX A: EXAMPLE OF UTTERANCE-LEVEL HLM FOR MEMORY FOR SENTENCES

Level 1 model:

$$Y_{ij} = \beta_{oj} + r_{ij},$$

where β_{oj} is the mean difference score in classroom j and r_{ij} is the individual level random effect.

Level 2 model:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \times (\text{Spanish multiclause utterances}) + \gamma_{02} \times (\text{Spanish "other" utterances}) + \gamma_{03} \times (\text{program structure}) + \gamma_{04} \times (\text{minutes of instruction}) + u_{0j},$$

where β_{oj} is the mean difference score in classroom j, γ_{00} is the classroom-specific intercept, γ_{01} is the relation between Spanish multiclause utterances and the class mean difference in classroom j, γ_{02} is the relation between Spanish "other" utterances and the class mean in classroom j, γ_{03} is the relation between program structure and the class mean in classroom j, γ_{04} is the relation between minutes of instruction and the class mean difference in classroom j, and u_{0j} is the classroom level random effect.

ACKNOWLEDGMENTS

The authors thank David J. Francis for his help during the planning stage of this project; Janellen Huttenlocher, and Susan Goldin-Meadow for their feedback throughout the completion of this study; Diana Pineda and Elizabeth Fein for their help with data collection; Nonie K. Lesaux for her comments on earlier manuscript drafts; and the anonymous reviewers for their suggestions to improve the manuscript. We are also grateful to the participating families and teachers.

NOTES

1. One student who scored 2.5 *SD* above the sample's picture vocabulary difference mean (standard score mean = 17.55; *SD* = 14.53) was excluded from all analyses, leaving 101 participating children.

2. In line with previous studies of teacher input (i.e., Klibanoff et al., 2006), effect sizes were calculated in the following way: we divided the main teacher input coefficient by the standard deviation of the *W*-score gain (0.22×9.88) and multiplied this number by the standard deviation of the main teacher predictor (0.002×21.40).

REFERENCES

August, D., Snow, C., Carlo, M., Proctor, P., Rolla de San Francisco, A., Duursma, E., et al. (2006). Literacy development in elementary school second-language learners. *Topics in Language Disorders, 26*, 351–364.

August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth*. Mahwah, NJ: Erlbaum.

Bornstein, M. H., Haynes, O. M., & Painter, K. M. (1998). Sources of child vocabulary competence: A multivariate model. *Journal of Child Language, 25*, 367–394.

Bowers, E., & Vasilyeva, M. (2011). The relation between teacher input and lexical growth of preschoolers. *Applied Psycholinguistics, 32*, 221–241.

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: Urban Institute.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instruction and children's literacy skill growth. *Journal of Educational Psychology, 98*, 665–689.

Connor, C. M., Morrison, F. J., & Underwood, P. (2007). A second chance in second grade? The independent and cumulative impact of first and second grade reading instruction and students' letter-word reading skill growth. *Scientific Studies of Reading, 11*, 199–233.

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

Dickinson, D. K., & Porche, M. V. (2011). Relationship between language experiences in preschool classrooms and children's kindergarten and fourth grade language and reading abilities. *Child Development, 82*, 870–886.

Dickinson, D. K., & Smith, M. W. (1991). Preschool talk: Patterns of teacher–child interaction in early childhood classrooms. *Journal of Research in Childhood Education, 6*, 20–29.

Dickinson, D., & Smith, M. W. (1994). Long-term effects of preschool teachers' book readings on low-income children's vocabulary and story comprehension. *Reading Research Quarterly, 29*, 104–122.

Duncan, S., & DeAvila, E. (2000). *Prelanguage Assessment Scale*. Monterey, CA: CTB/McGraw–Hill.

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

Ellis, R. (2009). *The study of second language acquisition* (2nd ed.). New York: Oxford University Press.

Field, A. (2005). *Discovering statistics using SPSS*. Thousand Oaks, CA: Sage.

Furrow, D., Nelson, K., & Benedict, H. (1979). Mothers' speech to children and syntactic development: Some simple relationships. *Journal of Child Language, 6*, 423–442.

Francis, D., Snow, C., August, D., Carlson, C., Miller, J., & Iglesias, A. (2006). Measures of reading comprehension: A latent variable analysis of the diagnostic assessment of reading comprehension. *Scientific Studies of Reading, 10*, 301–322.

Fry, R., & Gonzales, F. (2008). *One-in-five and growing fast: A profile of Hispanic public school students*. Washington, DC: Pew Hispanic Center.

Genesee, F. (1999). *Program alternatives for linguistically diverse students* (Educational Practice Report No. 1). Washington, DC: Center for Research on Education, Diversity & Excellence.

Genesee, F., Lindholm-Leary, K., Saunders, W., & Christian, D. (2005). *Educating English language learners: A synthesis of research evidence*. New York: Cambridge University Press.

Gottardo, A., & Mueller, J. (2009). Are first and second language factors related in predicting L2 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.

Hakuta, K., Goto Butler, Y., & Witt, D. (2000). *How long does it take English Learners to attain proficiency?* Linguistic Minority Research Institute Policy Report 2000-1, University of California.

Hammer, C. S., Davison, M. D., Lawrence, F. R., & Miccio, A. W. (2009). The effect of home language on bilingual children's vocabulary and emergent literacy development during Head Start and kindergarten. *Scientific Studies of Reading, 13*, 99–121.

Harkness, S. (1977). Aspects of the social environment and first language acquisition in rural Africa. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 309–318). Cambridge: Cambridge University Press.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experiences of young American children*. Baltimore, MD: Paul H. Brookes.

Hernandez, D. J., Denton, N. A., & Macartney, S. E. (2008). Children in immigrant families: Looking into America's future. *Social Policy Report, 22*, 3–22.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*, 1368–1378.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26*, 55–88.

Hoff-Ginsberg, E. (1998). The relation of birth order and socioeconomic status to children's language experience and language development. *Applied Psycholinguistics, 19*, 603–629.

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development, 73*, 418–433.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*, 236–248.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology, 45*, 337–374.

Huttenlocher, J., Vasilyeva, M., Waterfall, H., Vevea, J., & Hedges, L. (2007). The varieties of speech to young children. *Developmental Psychology, 43*, 1062–1083.

Jaffe, L. E. (2009). *Development, interpretation, and application of the W score and the relative proficiency index* (Woodcock–Johnson III Assessment Service Bulletin No. 11). Rolling Meadows, IL: Riverside.

Kindler, A. L. (2002). *Survey of the states' limited English proficient students and available educational programs and services: 2000–2001 summary report*. Washington, DC: National Clearinghouse for English Language Acquisition.

Klibanoff, R., Levine, S. C., & Huttenlocher, J., Vasilyeva, M., & Hedges, L. (2006). Preschool children's mathematical knowledge: The effect of teacher "math talk." *Developmental Psychology, 42*, 59–69.

Kohnert, K. (2008). Second language acquisition: Success factors in sequential bilingualism. *ASHA Leader, 13*, 10–13.

Kohnert, K., & Pham, G. (2010). The process of acquiring a first and second language. In M. Shatz & L. Wilkinson (Eds.), *Preparing to educate English language learners*. New York: Guilford Press.

Landau, B., & Gleitman, L. (1985). *Language and experience*. Cambridge, MA: Harvard University Press.

Lawrence, J., & Snow, C. (2011). Oral discourse and reading. In M. Kamil, P. D. Pearson, E. B. Moje, & P. Afflerbach (Eds.), *Handbook of reading research* (Vol. 4). Mahwah, NJ: Erlbaum.

Lindholm-Leary, K., & Borsato, G. (2006). Academic achievement. In F. Genesee (Ed.), *Educating English language learners: A synthesis of research evidence*. New York: Cambridge University Press.

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

MacWhinney, B. (2007). *The CHILDES Project: Tools for analyzing talk*. Mahwah, NJ: Erlbaum.

Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority learners. *Journal of Educational Psychology, 102*, 701–711.

Mancilla-Martinez, J., & Lesaux, N. K. (2011). The gap between Spanish-speakers' word reading and word knowledge: A longitudinal study. *Child Development, 82*, 1544–1560.

Manis, F. R., Lindsey, K. A., & Bailey, C. E. (2004). Development of reading in grades K–2 in Spanish-speaking English-language learners. *Learning Disabilities Research & Practice, 19*, 214–224.

Mashburn, A. J., Justice, L. M., Downer, J. T., & Pianta, R. C. (2009). Peer effects on children's language development during pre-kindergarten. *Child Development, 80*, 686–702.

Medina, M., Jr., & Escamilla, K. (1992). English acquisition by fluent- and limited-Spanish-proficient Mexican Americans in a 3-year maintenance bilingual program. *Hispanic Journal of Behavioral Sciences, 14*, 252–267.

Miccio, A. W., Patton, T. O., Páez, M. M., Hammer, C. S., & Wagstaff, D. A. (2005). *Vocabulary development in Spanish-speaking Head Start children of Puerto Rican descent. ISB4: Proceedings of the 4th International Symposium on Bilingualism*. Somerville, MA: Cascadilla Press.

Moughamian, A. C., Rivera, M. O., & Francis, D. J. (2009). *Instructional models and strategies for teaching English language learners*. Portsmouth, NH: Center on Instruction.

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.

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.

NCELA. US Department of Education. (2006). *The growing numbers of limited English proficient students.* Office of English Language Acquisition, Language Enhancement and Academic Achievement for Limited English Proficient Students (OELA). Retrieved November 2006 from http://www.ncela.gwu.edu/stats/2_nation.htm

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

Oller, D. K., & Pearson, B. Z. (2002). Assessing the effects of bilingualism: A background. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 3–21). Clevedon: Multilingual Matters.

Páez, M. M., Tabors, P. O., & López, L. M. (2007). Dual language and literacy development of Spanish speaking preschool children. *Journal of Applied Developmental Psychology, 28*, 85–102.

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

Pearson, B. Z., Fernandez, S., Lewedag, V., & Oller, D. K. (1997). Input factors in lexical learning of bilingual infants. *Applied Psycholinguistics, 18*, 41–58.

Proctor, C. P., August, D., Carlo, M. S., & Snow, C. E. (2006). The intriguing role of Spanish vocabulary knowledge in predicting English reading comprehension. *Journal of Educational Psychology, 98*, 159–169.

Proctor, C. P., August, D., Snow, C., & Barr, C. D. (2010). The interdependence continuum: A perspective on the nature of Spanish–English bilingual reading comprehension. *Bilingual Research Journal, 33*, 5–20.

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.

Ramirez, J. D., Pasta, D. J., Yuen, S., Billings, D. K., & Ramey, D. R. (1991). *Final report: Longitudinal study of structural immersion strategy, early-exit, and late-exit transitional bilingual education programs for language minority children* (Report to the US Department of Education). San Mateo, CA: Aguirre International.

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

Rumbaut, R. G., Massey, D. S., & Bean, F. D. (2006). Linguistic life expectancies: Immigrant language retention in southern California. *Population Development and Review, 32*, 447–460.

Saunders, W., Foorman, B., & Carlson, C. (2006). Is a separate block of time for oral English language development in programs for English learners needed? *Elementary School Journal, 107*, 181–198.

Saunders, W., & O'Brien, G. (2006). Oral language. In F. Genesee (Ed.), *Educating English language learners: A synthesis of research evidence*. New York: Cambridge University Press.

Smith, M., & Dickinson, D. (1994). Describing oral language opportunities and environments in Head Start and other preschool classrooms. *Early Childhood Research Quarterly, 9*, 345–366.

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

Vasilyeva, M., Waterfull, H., & Huttenlocher, J. (2008). Emergence of syntax: Commonalities and differences across children. *Developmental Science, 11*, 84–97.

Veltman, C. (1983). *Language shift in the United States*. Amsterdam: Mouton.

Wasik, B., & Bond, M. A. (2001). Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of Educational Psychology, 93*, 243–250.

Weizman, Z. O., & Snow, C. E. (2001). Lexical output as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology, 37*, 265–279.

Winsler, A., Diaz, R. M., Espinosa, L., & Rodriguez, J. L. (1999). When learning a second language does not mean losing the first: Bilingual language development in low-income, Spanish-speaking children attending bilingual preschool. *Child Development, 70*, 349–362.

Woodcock, R. W. (1991). *Woodcock Language Proficiency Battery—Revised, English and Spanish forms: Examiner's manual*. Itasca, IL: Riverside.

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

Zentella, A. C. (1997). *Growing up bilingual: Puerto Rican children in New York*. Oxford: Blackwell.