



Use of the home language in preschool classrooms and first- and second-language development among dual-language learners

Helen H. Raikes^{a,*}, Lisa White^b, Sheridan Green^c, Margaret Burchinal^d, Kirsten Kainz^b, Diane Horm^e, Gary Bingham^f, Alan Cobo-Lewis^g, Lisa St. Clair^h, Daryl Greenfield^b, Jan Esteraich^a

^a University of Nebraska-Lincoln, Lincoln, NE, United States

^b American Institutes for Research, San Mateo, CA, United States

^c Clayton Early Learning, Denver, CO, United States

^d Frank Porter Graham Child Development Institute, University of North Carolina, Chapel Hill, NC, United States

^e University of Oklahoma, Tulsa, OK, United States

^f Georgia State University, Atlanta, GA, United States

^g University of Maine, Orono, ME, United States

^h Omaha Program Evaluation Services, Omaha, NE, United States

ARTICLE INFO

Article history:

Received 17 April 2016

Received in revised form 19 February 2018

Accepted 20 June 2018

Keywords:

Dual language learners

Head Start

Teachers

Home language use

English

Language outcomes

ABSTRACT

The number of dual-language learners (DLLs) from low-income backgrounds who attend early education programs in the U.S. is rapidly increasing, leading to a need for research that examines the effects of classroom practices, including whether teachers speak English or the home language, on DLL children's school readiness. This issue was examined in Educare, an expanded Early Head Start/Head Start program for low-income children that ensures children receive high-quality early education and care. Analyses of 1961 DLL Educare children across 16 sites were conducted to compare the acquisition of language skills in English and Spanish among children in classrooms in which teachers used English and little or no Spanish, English with some Spanish, and both English and Spanish. Findings indicated that DLL children in all groups showed gains in language skills in both English and Spanish, but that DLL children from classrooms with both English and Spanish instruction had significantly higher Spanish auditory comprehension scores than other children. Findings from this study have implications for practice, highlighting the value of all 3 types of classrooms for English-language growth and the additional value of English/Spanish instruction for Spanish language growth.

© 2018 Elsevier Inc. All rights reserved.

The number of dual language learners (DLLs) in the United States has more than doubled in the past 30 years (Federal Interagency Forum on Child and Family Statistics, 2015), with estimates of over one quarter of all children under age 5 learning a language other than English at home (Matthews, 2011). Children who are DLLs start learning a second language while still developing their first, making their language-learning trajectories and their developmental needs, distinct from those of monolingual children (Espinosa, 2013a). Unfortunately, DLL children live in poverty at a higher rate than other children in the United States (Administration for Children and Families, 2008), which comes with a number of associated risks, including low academic achievement (National

Task Force on Early Childhood Education for Hispanics, 2007). The overwhelming majority of DLLs in the United States come from Spanish-speaking homes (Barrueco, López, Ong, & Lozano, 2012) and already perform lower than their monolingual peers at kindergarten entry (Espinosa, 2013b). This gap continues to widen during formal schooling (Rumberger & Tran, 2010), making it critical to support the unique developmental needs of Hispanic DLL children from low-income families from an early age.

High-quality early care and education (ECE) programs like Head Start or state pre-kindergarten may potentially buffer against these risks as they demonstrate large impacts, particularly for Spanish-English-speaking DLL children from low-income backgrounds (Administration for Children and Families, 2010; Authors, 2015; Currie & Thomas, 1996; Gormley, 2008; Gormley, Gayer, Phillips, & Dawson, 2005; Yoshikawa et al., 2013). These are encouraging findings given the increasing number of DLLs attending public

* Corresponding author.

E-mail address: hraikes2@unl.edu (H.H. Raikes).

early childhood education programs, now comprising nearly a third of all children attending such programs (*Administration for Children and Families*, 2013). These statistics highlight the need to understand the elements of ECE programs that best meet the needs of young Spanish–English DLLs from low-income families.

Building children's language and literacy skills in prekindergarten appears fundamental to ensuring school success for all children and may be key to reducing the achievement gap between DLL children and their monolingual peers if programs have larger impacts for DLL children (August & Shanahan, 2006). Rigorous quasi-experimental studies suggest that high-quality ECE improves language and literacy skills for low-income children in general, and has even larger impacts for Spanish–English DLLs from low-income families (*Administration for Children and Families*, 2010; Currie & Thomas, 1996; Gormley et al., 2005; Gormley, 2008; Yoshikawa et al., 2013). Some, albeit mixed, evidence suggests that incorporating the DLL child's home language and culture in the ECE classrooms increases ECE impacts (Buysse, Peisner-Feinberg, Pérez, Hammer, & Knowles, 2014; Hammer et al., 2014; Winsler, Burchinal et al., 2014). However, there remains a shortage of consistent evidence to demonstrate how varying levels of the home language in ECE classrooms affects young DLL children's language development in both languages. Differences in the extent to which the home language was incorporated in the classroom in most prior studies is believed to be one factor that accounts for the mixed findings (Hammer et al., 2014). The current study examines this issue in a large ECE program that serves many DLL preschoolers to test whether exposure to the home language is related to higher levels of language skills in both the home language, Spanish, and the second language, English.

1. Theoretical frameworks for dual-language learning

The use of the home language in ECE classrooms is believed to be especially important in promoting the development of young DLLs, from theoretical and practitioner perspectives (Castro, Espinosa, & Pérez, 2011; Halle et al., 2014). Recommendations from practitioners suggest that it is particularly important for young DLLs in ECE classrooms to experience environments in which the home language is valued, to ultimately help DLL children feel comfortable, respected, safe, and eager to learn (*National Association for the Education of Young Children*, 1995, 2009). Home language use in the classroom is posited to provide culturally and linguistically sensitive environments for young DLLs, which may be particularly important for their learning and development, and ability to benefit from experiences in the early childhood classroom (Collier & Thomas, 2004; Espinosa & Magruder, 2015).

Theoretical support for incorporating the home language in early childhood classrooms is provided to a large extent by cross-language transfer theory, which suggests that skills learned in a child's native language (L1) will transfer to the learning of a second language (L2) (Cummins, 1979). Studies predicated on this theory provide evidence that a strong base in the child's first language fosters the development of a second language (August & Shanahan, 2006; Dickinson, McCabe, Clark-Chiarelli, & Wolf, 2004; López & Greenfield, 2004; Winsler, Kim, & Richard, 2014). Such claims are often used to justify the employment of children's home language in the ECE classroom, with the idea that strengthening, supporting, and fostering first-language development for DLL children will create a solid foundation for DLLs' language development in English (Goldenberg, Hicks, & Lit, 2013). However, policy expectations remain mixed—many policies support bilingual development and recommend exposing DLL children to instruction in the home language (*Administration for Children and Families*, 2017), while other policies and associated practices support English-only instruction (Hoff et al., 2012). These debates are often complicated by other,

more practical constraints, such as limited resources (e.g. shortage of bilingual staff; availability of adequate dual language tools and curricula) and a lack of targeted professional development initiatives to prepare the workforce for working with linguistically diverse learners, making dual-language instruction difficult to implement. Studies with large samples, relatively at scale, as is true for the current study, are useful for examining opposing expectations, potentially leading to greater clarity in policy-making and expectations for practice, as well as verifying or refuting theories.

2. Effects of language use on DLL children's outcomes

There is empirical evidence to suggest that language development for DLLs in both L1 and L2 is largely determined by both the quality and amount of input the child receives in both languages (Hammer, Komaroff et al., 2012; Hoff, 2006; Hoff et al., 2012). Some studies focus particularly on parent/caregiver use of language in L1 and L2, and demonstrate that the more input a child receives in a given language is related to higher proficiency in that language (Hammer, Komaroff et al., 2012; Paradis, Nicoladis, Crago, & Genesee, 2011; Pearson, Fernandez, Lewedeg, & Oller, 1997). These studies examining home language use provide support for similar effects of teacher language use and prompt further investigation of language contexts in early childhood classrooms.

Studies have empirically examined the language environment in the classroom as it relates to DLL children's outcomes; however, findings from the research literature are mixed (Buysse et al., 2014). Several scholars have conducted randomized control trials to understand the impact of specific instructional formats for DLLs on children's outcomes in L1 and L2 (e.g., Bernhard et al., 2006; Durán, Roseth, & Hoffman, 2010; Farver, Lonigan, & Eppe, 2009; Madera & Gerber, 2008; Méndez, Crais, Castro, & Kainz, 2015). Combining English- and Spanish-learning opportunities in targeted group instruction relates to higher levels of vocabulary skills in both the home language and English among DLL preschoolers (Collier & Thomas, 2004; Farver et al., 2009; Méndez et al., 2015) and to improved literacy scores in both languages in first grade (Tong, Lara-Alecio, Irby, & Mathes, 2011), compared to English-only instruction.

Other experimental or quasi-experimental designs provide support for particular dual language approaches (e.g., Barnett, Yarosz, Thomas, Jung, & Blanco, 2007; Restrepo et al., 2010). In one study, for example, Barnett and colleagues found that Hispanic children in a two-way immersion program in English and Spanish made comparable gains in English vocabulary skills compared to those in English-only instruction (i.e., no differences in English vocabulary between the two groups), but they made greater gains in Spanish vocabulary and phonological awareness (in both English and Spanish) than the English-only group (Barnett et al., 2007). Taken together, these findings highlight both concurrent and longitudinal associations on the positive effects of combining the home language with English instruction to support L1 and L2 development and, importantly, that the increased use of L1 does not impede language development in L2.

Other studies on this topic have focused particularly on examining the percentage of time teachers spend speaking L1 and L2 in the classroom, irrespective of a specific program model (e.g., Burchinal, Field, López, Howes, & Pianta, 2012; Chang et al., 2007), which affords an examination of typical language use in the classroom (compared to testing a specific DLL approach/technique). Generally, higher levels of skills in the home language are related to greater instruction time in the home language (Burchinal et al., 2012; Collins, 2014) and higher levels of skills in English are related to greater instruction time in English (Collins, 2014; Solari, Landry, Zucker, & Crawford, 2011). These trends align with results from

some of the experimental studies described above (e.g., Barnett et al., 2007; Méndez et al., 2015) and with the idea that greater input in one language is related to enhanced development in that language (Hammer, Komaroff et al., 2012; Hoff, 2006; Hoff et al., 2012).

However, these findings are not always consistent, adding to concerns frequently expressed by policymakers that an emphasis on Spanish instruction may actually undermine English acquisition for DLL children (Hoff et al., 2012). For example, Vitiello, Downer, and Williford (2011) found that the amount of Spanish instruction during classroom observations was negatively associated with children's English language proficiency, which is in contrast to studies finding support for use of the home language to promote English learning. Similar effects have been found in the reverse direction; Solari et al. (2011), for example, reported that amount of English instruction during book reading and oral language activities was negatively associated with Spanish language outcomes. Although primarily correlational, such findings or views highlight a concern that children who receive continued support in Spanish (L1) in preschool may enter primary school with limited English (L2) language skills, which may make it difficult for them to fully benefit from classroom instruction (Shatz & Wilkinson, 2010). Even other investigators have found that children in classrooms with an increased use of L1 instruction (in one case, exclusive Spanish instruction in the first year of a dual-language intervention), did not differ in English receptive vocabulary, compared to children who experienced English-only instruction (Durán et al., 2010), concluding that Spanish instruction does not harm English acquisition. However, the authors do not explicitly imply that Spanish instruction *promotes* the acquisition of English. These discrepant findings make it challenging to conclude how L1 and L2 use in the classroom may influence both first- and second-language learning, leaving room for more exploratory studies to examine such questions, such as the current study.

3. Limitations of existing literature

In a comprehensive review on the effects of early education programs on DLL children's academic outcomes, Buysse et al. (2014) reported that although there was some support for using the home language to educate DLLs, results were inconsistent and provide a lack of cohesive evidence on the best approach of L1 and L2 use in the classroom. In general, only a small number of studies empirically examined these questions, and those that have varied greatly in their definition and assessment of DLLs (including choice and language of assessment), research design, and measurement/implementation of language of instruction in the classroom.

For example, some studies have only assessed children in one language (e.g., Chang et al., 2007; Restrepo et al., 2010), making them unable to examine impact of L1 and L2 language use on children's outcomes in *both* languages, which is critical for examining DLL development (Espinosa, 2013b). Others are limited by small sample sizes of children and classrooms, making interpretation difficult and limiting generalizability. Additionally, existing studies differ greatly in their measurement and operationalization of classroom language use—some studies only examined the effects of a few specific activities (often in the context of a small intervention on a particular concept, such as vocabulary) with some level of the home language included alongside typical instruction in English (e.g., Farver et al., 2009; Matera & Gerber, 2008; Méndez et al., 2015; Restrepo et al., 2010), while others captured typical classroom language use in L1 and L2 across the day (Burchinal et al., 2012; Solari et al., 2011; Vitiello et al., 2011), as they relate to children's language development in the first and second language. Last, classrooms in existing studies varied to a great extent on overall *quality* of classrooms, which is an important variable to consider, or control for,

when examining effect of L1 and L2 language use (Burchinal et al., 2012).

Researchers and policymakers, for these many reasons, have recognized the need for more research on this topic and have identified this as one of the most important gaps of knowledge in the field of early childhood (Buysse et al., 2014; National Task Force on Early Childhood Education for Hispanics, 2007). Although a growing research base documents the importance of examining DLL children's home and classroom experiences in relation to their L1 and L2 language development, early experiences with both languages vary greatly among DLLs, which contributes to difficulty understanding and differentiating classroom effects (Hammer et al., 2014). In addition, mixed findings on the effects of classroom home language use on children's L2 development makes drawing firm hypotheses difficult. Factors contributing to mixed or inconsistent findings include variability in percent of language exposure, limitations of assessment tools for DLLs, timing of children's first exposure to second language, SES, and family acculturation, among other unstudied factors (Hammer et al., 2014). In this study, we examine variability in language exposure in the context of Educare, a high-quality birth-5 education program serving children from diverse, low-income backgrounds, using a large sample, which allows for controlling a number of factors that might contribute to inconsistent findings as described above.

4. This study

The purpose of the current study was to examine the extent to which home language use in ECE classrooms related to the acquisition of Spanish and English language skills for Spanish–English DLLs from low-income families attending Educare Head Start/Early Head Start programs. Specifically, research questions were as follows:

1. Does the level of Spanish language exposure in ECE classrooms relate to maintaining and/or enhancing the skills in the home language for Spanish–English DLLs?
2. Does the level of Spanish language exposure in ECE classrooms relate to the acquisition of the second language, English?

5. Method

5.1. Study sites

This study uses data from the Educare Early Learning Network (ELN; Educare Learning Network, 2015) implementation study, with data from Spanish–English DLL children attending 16 of the 21 sites between 2008 and 2016. All sites were Head Start programs, served children at poverty level, provided birth-to-kindergarten, full-day, full-year center-based ECE, and shared a goal to use at least some amount of the home language in classrooms with Spanish–English DLL children.

Educare programs meet the Head Start and Early Head Start Program Performance Standards (Administration for Children and Families, 2016), but also include a variety of enhancements recommended, but not required, by Head Start (U.S. Department of Health and Human Services, 2012), including lead teachers with bachelor's degrees, intensive ongoing embedded professional development, and family support staff with limited caseloads and bachelor's-level training (Yazzejian, Bryant, & Kennell, 2013; see also <http://www.educarenschools.org/locations/index.php>). The ELN also intentionally uses common cross-site data (including the data used in this paper) in an on-going implementation study to inform classroom practices, as well as to monitor progress of the ELN in reducing the achievement gap for low-income children. More details on the Educare model and implementation study can be found in published reports (e.g., Guss, Norris, Horm, Monroe, & Wolfe, 2013;

Table 1
Educare classroom characteristics by classroom language groups.

	Classroom language									Sig. of F or χ^2
	English with no or little Spanish			English with some Spanish			English and Spanish			
	N	Mean or%	SD	N	Mean or%	SD	N	Mean or%	SD	
Prop. of class that is DLL	195	0.42	0.26	87	0.62	0.27	105	0.73	0.23	***
Number DLLs in classroom	195	6.51	4.27	89	9.30	5.00	105	11.03	4.55	***
How teacher uses Spanish										
Instruction and/or teaching	160	8%		85	49%		107	80%		***
Classroom management	160	8%		85	59%		107	86%		***
Social interactions	160	19%		85	84%		107	96%		***
Classroom quality										
ECERS-R total	163	5.32	0.79	76	5.81	0.72	92	5.78	0.59	***
ITERS-R total	20	5.52	0.89	10	5.81	0.44	14	5.99	0.40	
Number of classrooms										
Southeast	196	0%		89	0%		105	28%		***
Midwest	196	33%		89	39%		105	43%		
South central	196	47%		89	30%		105	11%		***
West or southwest	196	7%		89	12%		105	13%		
Northwest	196	5%		89	17%		105	5%		**
All regions	196	50%		89	23%		105	27%		

Note: *p*-Value associated with language exposure comparison based on F or χ^2 test is reported.

Each classroom was counted separately for each year it was an Educare classroom.

The questions about the purpose of using Spanish in the classrooms were added after the first year of the study so the data are missing for earliest classrooms.

** *p* < 0.01.

*** *p* < 0.001.

Stein, Freel, Hanson, Pacchiano, & Eiland-Williford, 2013; Yazejian et al., 2013).

The Educare implementation study has documented that classroom quality tends to be relatively high and that children benefit from attending. For example, average cross-site Environmental Rating Scales scores were in the good quality range (ECERS-R *M* = 5.7, *SD* = 0.7 and ITERS *M* = 5.9, *SD* = 0.3) and the Classroom Assessment Scoring System (CLASS) PreK ratings were in moderately high to high quality range (Emotional Support *M* = 6.2, *SD* = 0.7; Classroom Management *M* = 5.5, *SD* = 0.9, and Instructional Support *M* = 3.4, *SD* = 1.1). A randomized control trial reported significantly higher outcomes at age 2 years for Educare than for control group children (Yazejian et al., 2017). A correlational study documented that children who entered earlier and stayed in Educare longer had higher receptive language skills, with significant effects sizes for all children and larger effect sizes for DLL children than for English-only children (Yazejian, Bryant, Freel, Burchinal, & the Educare Learning Network Investigative Team, 2015).

The ELN collected common cross-site semi-annual assessments of children's language skills in both languages and annual observations of the amount of Spanish and English used in the classroom together with classroom quality and interviews with parents. Data for the current study were secondary and had been collected annually in the Educare implementation study over multiple years, from 2008 to 2016. Altogether, the classrooms in the 16 sites were from various parts of the U.S., including 9% from the Southeastern U.S., 31% from Midwestern states, 47% from South Central areas, 7% from Western/Southwestern U.S., and 6% from the Northwestern U.S. Some sites had DLL children in every classroom every year, whereas other sites had only one classroom with DLL children. Counting each classroom each year as a separate classroom, there were 4 sites with fewer than 10 classrooms with DLL children, 4 sites with 10–20 classrooms with DLL children, 5 sites with 20–50 classrooms with DLL children, and 3 with more than 50 classrooms.

5.2. Participants

The sample for the study consisted of 1961 Spanish–English DLL children over 18 months of age who attended Educare during the

years 2008 through 2016. All children with any language assessments were included. See Tables 1 and 2 for descriptions of the classrooms and participants.

Children entered Educare at ages ranging from infancy to the year before kindergarten, with 2.90 years as the mean age of entry (*SD* = 1.28). Not surprisingly, with many children entering at 3 years or older, they left on average 2 years later. The median number of repeated assessments of children was 2, and children remained in Educare for a mean of 1.99 years (*SD* = 1.15) (see Table 1). These DLL children were almost all Hispanic (98%). Two-thirds of the parents in the sample were born in Mexico, another 7% were born in Central America, 2% South America, and 2% were from Caribbean islands. Half of the children were male (52%), and all the families met Head Start criteria regarding income. On average, mothers had not graduated from high school, but were not typically teen-age or single mothers.

5.3. Measures

All children in Educare schools are assessed regularly through the ELN implementation study. Language skills are measured at two and three years of age for children who enter as infants and toddlers. Language skills are measured in the spring of each preschool year for all children and in the fall of the first preschool year for children who enter as preschoolers. Due to the fact that children enter Educare at varying ages, not all instruments were administered in all ages or in all years, so the number of child assessments and classrooms represented differs slightly depending on which outcome is the focus of analysis and on when the ELN adopted use of the measure. We report the number of assessments separately for each outcome in Table 3.

5.4. Language skills

Expressive and receptive language skills in Spanish were measured using the Preschool Language Scale–Spanish Version 4 (PLS-4; Zimmerman, Steiner, & Pond, 2002), and receptive language skills in English were measured using the Peabody Picture Vocabulary Test Version 4 (PPVT-4; Dunn & Dunn, 2007). The PPVT

Table 2
Educare child and family characteristics by classroom language groups.

	Classroom language									Sig. of F or χ^2
	English with no or little Spanish			English with some Spanish			English and Spanish			
	N	Mean or%	SD	N	Mean or%	SD	N	Mean or%	SD	
Child's Educare participation										
Age child enrolled (years)	878	2.92	1.25	447	2.93	1.34	644	2.83	1.30	**
Age child exited (years) ^a	789	4.98	0.72	366	5.07	0.67	519	5.10	0.61	
Time in program (years) ^a	789	1.93	1.06	366	2.04	1.27	519	2.07	1.19	
Child characteristics										
Child gender: male	878	53%	0.83	447	53%	0.82	644	50%	0.78	**
Health	826	4.12		424	4.13		622	4.25		
Hispanic	878	98%		447	98%		644	97%		
IEP/ISFP	878	10%		447	12%		644	14%		
Spanish is language child hears most at home	828	72%		424	71%		624	74%		
Spanish was child's first language ^b	831	82%		424	79%		624	83%		
Family characteristics										
Parent born in USA ^b	831	21%		424	24%	0.43	624	25%		
Food insecurity ^b	831	50%		424	52%		624	49%		
Mother depressive symptoms ^b	828	0.21		424	0.20		622	0.20		
Mother age at child's birth ^b	844	26.53	6.54	430	26.29	6.47	617	26.78	6.54	***
Mother education in years ^b	849	11.07	2.10	429	11.31	2.14	634	11.72	2.35	
Stressful life events ^b	824	2.50	1.89	422	2.59	1.95	619	2.60	1.93	
Single Parent ^b	852	37%		421	35%		632	34%		

Note: *p*-Value associated with language exposure comparison based on F or χ^2 test is reported.

^a 119 children were still enrolled at the final assessment point so the sample size for exit age and time in program is smaller than for enrollment age.

^b Not all families completed the family survey so sample sizes are slightly smaller for these variable.

** *p* < 0.01.

*** *p* < 0.001.

Table 3
Repeated assessments of language outcomes in Spanish and English.

		Assessment count	Classroom language								
			English with no or little Spanish			English with Some Spanish			English and Spanish		
			N	Mean	SD	N	Mean	SD	N	Mean	SD
Preschool language scale 4 – Spanish											
Auditory Comprehension standard score	1	238	93.93	16.86	123	92.24	16.18	205	99.46	18.52	
	2	142	94.08	16.82	83	95.06	15.03	148	99.72	17.75	
	3	69	96.20	16.32	43	91.86	15.26	73	103.1	15.36	
Expressive Communication standard score	1	239	96.82	17.20	121	97.54	14.94	205	98.09	20.10	
	2	142	95.61	18.86	81	96.05	14.92	147	96.71	21.94	
	3	69	94.52	19.18	43	92.09	17.45	72	96.93	22.67	
Age in months when tested on PLS	1	239	44.81	12.49	123	44.54	14.01	205	45.56	13.05	
	2	142	51.49	9.72	83	52.86	11.04	148	52.03	10.96	
	3	69	54.74	8.70	43	58.33	8.16	73	54.64	10.25	
Peabody Picture Vocabulary Test – English											
PPVT standard score	1	844	81.10	14.69	464	82.19	12.78	609	82.53	14.81	
	2	709	82.27	14.21	396	83.14	13.23	528	84.21	14.41	
	3	428	83.55	14.05	213	86.85	13.36	238	88.00	13.83	
	4	295	84.57	13.87	143	90.47	14.07	141	89.65	12.99	
	5	82	87.39	13.77	49	90.47	13.92	54	92.63	12.06	
Age in months when tested on PPVT	1	844	48.39	9.20	464	48.56	9.12	609	49.04	9.50	
	2	709	52.81	8.08	396	52.70	8.01	528	52.80	8.34	
	3	428	56.38	7.18	213	56.82	7.02	238	55.65	6.82	
	4	295	59.43	5.35	143	58.64	5.64	141	60.72	3.26	
	5	83	60.08	3.98	49	59.84	3.85	54	60.72	3.26	

Note: Descriptive statistics were not provided for the 94 children with a fourth assessment and 2 children with a fifth assessment on the PLS-4 and 81 children with sixth assessments and 11 children with sixth assessments on the PPVT.

was administered from the onset of the Educare Implementation Study in fall of 2007, and the PLS-4 was added in 2010. Prior to that time, the Test de Vocabulario en Imágenes Peabody (TVIP; Dunn, Lugo, Padilla, & Dunn, 1986) was administered, but concerns about its ability to assess language skills led to the switch to the PLS, due to its stronger psychometric properties. Only the PLS-4 was administered at 2 years of age, whereas the PLS-4 and PPVT were administered at 3 years of age and in the subsequent preschool assessments. These measures, selected by the ELN, were chosen to describe children's language progress in both Spanish and English.

The PLS-4 with standardized scores for Spanish Auditory Comprehension (AC) and Expressive Communication (EC) was judged by the ELN as a good measure for Spanish language progress. The PLS-4 administration requires an hour or more to administer in each language, so the PPVT-4 was selected as a low-burden way to assess children's English-language skills. Moreover, while both English and Spanish versions of the PLS tests were available, in addition to the burden to children of administering the assessment in two languages, the equivalence in measurement between the two versions is not well established.

5.4.1. Spanish PLS-4

The Preschool Language Scale-Version 4 (PLS-4) Spanish Auditory Comprehension (AC) and Spanish Expressive Communication (EC; Zimmerman et al., 2002) scales were administered. The PLS-4 norms include 1188 Spanish-speaking children living in the United States who were between 3.0 and 5 years 11 months. The PLS-4 reports (Zimmerman et al., 2002) test–retest reliability of 0.77–0.86, and split half internal consistency of 0.80–0.90. Validity was reported (Zimmerman et al., 2002) as showing good sensitivity (AC: 0.83–0.91; EC 0.91–0.92) and moderate specificity (AC: 0.57–0.55; EC: 0.77–0.61).

5.4.2. Peabody picture vocabulary test

The Peabody Picture Vocabulary Test 4th Edition (PPVT-4; Dunn & Dunn, 2007) assessed children's receptive vocabulary skills. The PPVT-4 was developed to examine the receptive vocabulary development of adults and children as young as 2.5 years old. The PPVT reports strong reliability in terms of internal consistency reliability coefficients (0.95–0.97 for the 2- to 6-year-old age range) and strong test–retest reliabilities (0.91–0.94) (Dunn & Dunn, 2007).

The data collectors attended a central training. The training provided an overview of the measures and then practice administering each measure. Subsequently, each data collector submitted a videotape demonstrating administration of all measures with non-study children. The trainers reviewed this tape, and certified the data collector if they met all criteria in the common standard defined by the administration procedures outlined in user manuals for each assessment. This certification procedure was repeated annually.

5.4.3. Demographic information

Information about demographic characteristics was collected during interviews with parents. Demographics included child gender, whether the child had an IEP/ISFP, maternal education, maternal age, and the maternal partner status was collected by programs at enrollment and annually through an interview. In addition, the parent interview asked about whether the parent was born in the United States, whether Spanish was the child's first language, language child most frequently hears at home, child health and several social risk factors such as maternal depressive symptoms, life stressors, and food insecurity (note, these are presented to describe the sample, but are not included in the analysis model due to considerable missing data and confounding of variables). Children's time in Educare was determined by calculating the difference between the reported enrollment date and the date of the assessment.

5.5. Classroom measures

Each Educare classroom was observed annually, assessing classroom quality using the Early Childhood Environmental Rating Scale-Revised (ECERS-R; Harms, Clifford, & Cryer, 2005) or the Infant Toddler Environment Rating Scale-Revised (ITERS-R; Harms, Cryer, & Clifford, 2006) and rating the use of Spanish in the classroom.

5.5.1. Home language use in classrooms

Data collectors recorded the languages used in the classroom during the annual classroom quality observations. They noted languages used at the completion of their 3-h classroom observation period. Data collectors responded to the question, "Estimate how this classroom can be classified (do not count display or books—consider spoken language) (1) mono-lingual English with little or no use of another language (0–10% other language); (2) mono-lingual English with occasional use of another language in greetings, songs, etc. (11–25% other); (3) dual-lingual English with use of another language less than half of the time (other 26–50%

other); (4) dual-lingual English with use of another language more than half of the time (other 51–75%); (5) monolingual non-English language with occasional use of English (other 76–90%); and (6) mono-lingual non-English language with little or no use of English (other 91–100%)." No classrooms were in the final category, and the groups representing 25% or more dual-linguistic classrooms were combined due to limited sample sizes in each of those groups in order to obtain three groups with similar sample sizes.¹ The three classroom language groups included: (1) no home language – English with little or no Spanish (i.e., used Spanish 0–10% of the time); (2) limited home language – English with some Spanish (i.e., reported using Spanish 11–25% of the time); and (3) substantial home language—English and Spanish (i.e., reported using Spanish 26–90% of the time).

5.5.2. Classroom quality

The Early Childhood Environment Rating Scales (ECERS-R; Harms et al., 2005) and the Infant Toddler Environment Rating Scale-Revised (ITERS-R; Harms et al., 2006) were the measures of classroom quality used in the study. The ECERS-R consists of 43 items. The scale's authors reported an interclass correlation for the total score of 0.915 and the Cronbach's alpha for the total scale is 0.92 (Harms et al., 2005). ECERS-R scores have been associated with children's social-emotional and cognitive outcomes (Peisner-Feinberg et al., 2001), language development (Burchinal, Peisner-Feinberg, Bryant, & Clifford, 2000) and kindergarten readiness (Peisner-Feinberg et al., 1999). ITERS-R scale reliability was established on 90 observations with two paired observations each in 45 group settings. Interclass correlation was 0.92 for 39 items and 32 child-related items. The scale had a Cronbach's alpha of 0.93 for all items and 0.92 for 32 child-related items. Rating scale scores have been associated with children's developmental outcomes (Burchinal, Roberts, Nabors, & Bryant, 1996; Peisner-Feinberg et al., 1999), structural measures of classroom quality (Phillipsen, Burchinal, Howes, & Cryer, 1998), and teacher salaries and total program costs (Whitebook, Howes, & Phillips, 1989).

The ERS classroom observers were trained by a small team of observers (referred to as "gold standard observers"), who in turn were trained by the scale authors. ERS classroom observers were members of local research teams. The majority of research teams were led by faculty at nearby universities. Thus, many of the observers were M.S. or Ph.D. graduate students. Some were full-time employees hired for the local Educare evaluations. Nearly all of these evaluation employees had a B.A., but many had a M.S. and some, a Ph.D. Local site ERS observers were required to complete several practice sessions prior to reliability testing. Next, they had to demonstrate reliability at a minimum of 85% within one point across all items (omitting the staff items) with the consensus score from the two separate classroom observations conducted with the "gold standard observer." Each data collector had to be recertified annually by demonstrating at least 85% agreement with the "gold standard observer" for one observation. Instructions for scoring language use in the classroom were reviewed as part of the ERS training.

5.6. Analyses

First, descriptive statistics described the Educare classrooms with Spanish–English DLLs and those children and their families. Second, multilevel analyses compared language skills over time in Spanish (L1) and English (L2) for children in classrooms that varied

¹ There were a few bilingual classrooms that strived to provide 50% Spanish and 50% English and categorizing them was difficult, so creating a single category of 25% or more was thought to provide the most accurate grouping.

in terms of whether and how much Spanish was used. Three-level cross-classified linear models (Beretvas, 2011; Goldstein, 1994) accounted for children changing classrooms over time in Educare by estimating a random intercept for every classroom and a random intercept for every child, allowing each to vary in terms of level and variability over time. Preliminary models also included a random-effects slope for time, retained in analyses of the PPVT, but not the PLS, due to fewer repeated assessments per child on this measure. All variables were grand-mean centered, thus main effects were estimated at the average level of all predictors, including time. The models included site, classroom quality, gender, IEP/ISFP status, maternal age, education, and maternal partner status as covariates. The analyses focused on asking whether children differed in terms of their level or rate of change over time in their language skills in either Spanish or English when their teachers spoke only English and little or no Spanish, English with some Spanish, or a substantial amount of both English and Spanish. This three-level model is

$$\text{Level 1 : } Y_{ijk} = B_{0ij} + B_{1ij} \text{ time-in-Educare}_{ij} + e_{ijk}$$

$$\text{Level 2 : } (\text{child's intercept}) B_{0ij} = \delta_{00j} + \delta_{01j} \text{Male}_{ij} + \delta_{03j} \text{IEP}_{ij} + \delta_{04j} \text{MaternalAge}_{ij} + \delta_{05j} \text{MaternalEd}_{ij} + \delta_{06j} \text{Singleparent}_{ij} + \zeta_{0j} \\ (\text{child's slope}) B_{1ij} = \delta_{10j} + \zeta_{1j}$$

$$\text{Level 3 : } (\text{classroom intercept}) \delta_{00j} = \pi_{00} + \pi_{01} \text{Site}_{ij} + \pi_{02} \text{ClassroomLanguageGroup}_{ij} + \pi_{03} \text{Classroom Quality(ERS)}_{ij} + \delta_{0j} \\ (\text{classroom slope}) \delta_{01j} = \pi_{10} + \pi_{11} \text{ClassroomLanguageGroup}_{ij} + \delta_{1j}$$

where Y_{ijk} is the score of the i -th child on the j -th occasion in the k -th classroom and the random error terms include: e_{ijk} as the level 1 residual, δ_{0j} as the individual child's intercept residual, δ_{1j} as the individual child's slope residual (estimated only for the PPVT), and ζ_k as the individual classroom's intercept residual. Focus in these analyses was on the classroom-level parameters describing the differences in mean scores among the three classroom language groups (π_{02}) and in rates of change in scores among the three classroom language groups (π_{11}).

6. Results

6.1. Descriptive analyses

Descriptive information is in Tables 1–3. The three classroom types differed in the average proportion ($F(2,386) = 54.4, p < 0.001$) and number of DLL children ($F(2,388) = 36.1, p < 0.001$). The classrooms that were *English with little or no Spanish* reported 4.8 DLL children (42% of children in the classroom) on average; those using *English with some Spanish* averaged 9.3 DLL children (62%), while the classrooms with both *English and Spanish* averaged 11.0 DLL children (73%). The three groups of classrooms also differed in whether they used Spanish in instruction ($\chi^2(2,352) = 146.5, p < 0.001$), classroom management ($\chi^2(2,352) = 170.8, p < 0.01$), and social interactions ($\chi^2(2,352) = 188.2, p < 0.001$). Not surprisingly, classrooms with substantial use of both English and Spanish typically reported using Spanish for all three functions (80%–96% on average), while only about half of classrooms with some Spanish reported using Spanish for instruction and classroom management (although over three-fourths of the English with some Spanish classrooms did report using Spanish for social interactions; see Table 1). Very few of the classrooms with limited or no Spanish reported using Spanish for any of the three functions.

There were also differences in the proportion of classrooms with no, some, or substantial Spanish in terms of the region of the country and classroom quality. In the Southeast ($\chi^2(2,392) = 83.4, p < 0.001$), proportionately more classrooms used substantial Spanish, but in the South Central ($\chi^2(2,392) = 40.2, p < 0.001$) and Northwest ($\chi^2(2,392) = 13.8, p < 0.01$) proportionately fewer classrooms used substantial Spanish. The preschool classrooms with DLL

children in which little or no Spanish was spoken were rated as lower quality than the classrooms with some or substantial Spanish ($F(2,328) = 17.6, p < 0.001$). Table 1 shows that all classroom quality measures averaged scores greater than 5.

Next, we examined the characteristics of the DLL children and their families based on whether the children were in the classrooms with no, limited, or substantial use of Spanish. As reported in Table 2, the DLL children in the three language exposure groups did not differ in the age they entered ($M = 2.8$ – 2.9 years-of age) and length of time they stayed in Educare ($M = 1.9$ – 2.1 years). They also did not differ in whether parents reported Spanish was most frequently spoken at home in a large proportion of the homes (Proportion = 0.71–0.74), whether Spanish was their child's first language (Proportion = 0.79–0.83), whether the parent immigrated to the US (Proportion = 0.75–0.79) or whether the child was Hispanic (Proportion = 0.97–0.98). There were no differences in sex or disability status. About half of the children in each group were

male, and about 10–14% in each group had an IEP. Parents rated the children in the substantial Spanish group as significantly healthier than children in the other groups ($F(2,1869) = 5.24, p < 0.01$). The families did not differ across the three groups in terms of maternal age ($M = 26.3$ – 26.8), maternal depression (Proportion = 0.20–0.21), single partner status (Proportion = 0.34–0.37), and food insecurity (Proportion = 0.49–0.52), but did differ on maternal education ($F(2,1909) = 16.20, p < 0.001$). Mothers of children in the substantial Spanish classrooms had more years of education ($M = 11.7$) than did mothers of children in the no Spanish ($M = 11.1$) or limited Spanish ($M = 11.3$) classrooms.

Finally, Table 3 shows the means, standard deviations, and sample sizes for the longitudinal assessments of English and Spanish language skills. The PPVT was collected starting in the spring of 2008, and the PLS was added in 2010, and thus collected on all children enrolled in that year and all subsequent years. Accordingly, there are more children with PPVT scores and more repeated assessments on the PPVT than on the PLS.

6.2. Hierarchical linear model analyses

Results of the HLM analyses are shown in Table 4. The model tested whether the language exposure groups differed on the intercept (estimated at about 4.5 years-of-age) or on the rate of change over time. The model included covariates for which group differences were detected (location, ERS quality, maternal education, child health) and other standard demographics (gender, disability status, maternal age, and parent partner status). The bottom of the table lists the random-effect variance, reflecting the amount of variability among classrooms (labeled class intercept), children's intercepts (labeled child intercept), children's slopes for the PPVT, and the level 1 residual. The top of the table displays the statistical tests and the classroom language group adjusted means and estimates slopes over time. Several follow-up analyses were conducted to determine whether findings appear to be sensitive to the inclusion of other factors, including immigration status of the parent and whether Spanish was the child's first language and whether differences in sample sizes for PLS-4 and PPVT account for differences in findings.

Table 4
HLM analyses of Spanish and English Language outcomes by classroom language groups.

	PLS4 Spanish Auditory Comprehension		PLS4 Spanish Expressive Communication		PPVT English Receptive Language	
	F	Adjusted mean (seB (se))	F	Adjusted mean (seB (se))	F	Adjusted mean (seB (se))
Fixed effects						
Classroom language	F (2,590) = 3.22*		F(2,582) = 0.93		F(2,893) = 1.38	
English		97.95 ^{ab} (1.28)		98.62 (1.44)		85.51 (1.36)
Some Spanish		96.28 ^a (1.27)		97.59 (1.44)		85.89 (1.39)
Both English and Spanish		99.65 ^b (1.24)		99.70 (1.40)		85.05 (1.40)
Time	F (1,590) = 11.81**		F(1,582) = 0.21		F(1,893) = 279.0***	
Time × class language	F (2,590) = 1.56		F(2,582) = 0.85		F(2,893) = 1.25	
English only		2.35 (0.79)		0.04 (0.90)		4.70*** (.39)
Some Spanish		0.85 (0.86)		−1.17 (1.00)		5.21*** (.40)
English and Spanish		2.47 (0.79)		0.28 (0.91)		5.43*** (.41)
Random effects						
	Variance (se)		Variance (se)		Variance (se)	
Class intercept	12.48* (5.44)		16.80** (6.46)		3.06* (1.342)	
Child intercept	188.07*** (13.85)		212.99*** (15.94)		128.10*** (5.94)	
Child slope					43.84*** (4.25)	
Cov (int, slope)					5.61* (3.40)	
Residual	49.52*** (2.93)		69.17*** (4.00)		26.23*** (0.86)	

Note: Group means with different superscripts are significantly different from one another. Means within columns with the same or no superscript are not reliability different from one another. So findings indicate the children in Some English group scored on average lower than the children in the Both English and Spanish group, but children in both of those groups were not reliably different than the children in English group.

Model asks whether the three language exposure groups showed different intercepts or slopes describing change over time, and included as covariates – site, ERS classroom quality total score, child sex, whether the child had an IEP/ISFP, maternal education, maternal age, and whether the parent was single.

* $p < 0.10$.

* $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$.

6.2.1. PLS-4 Spanish Auditory Comprehension

The results of analyses of the PLS-4 Spanish Auditory Comprehension standard scores are listed in the first set of columns in Table 4. These results describe the analysis of 629 PLS-4 assessments of DLL children in 70 classrooms in which only English was spoken, 40 classrooms in which English and some Spanish was spoken, and 52 classrooms in which both English and Spanish were spoken. As identified by superscripts in the table, the three classroom language type groups differed significantly in the level of PLS-4 Spanish Auditory Comprehension ($F(2,590) = 3.22$, $p < 0.05$), with higher scores in the classrooms using both English and Spanish ($Adjusted M = 99.7$) than in the classrooms using English and some Spanish ($Adjusted M = 96.3$). On average, children showed gains in the Spanish Auditory Comprehension standard scores with each additional year in Educare, but the rate of change was not reliably different across the three classroom language groups. The average rate of change per year ranged from 0.85 in the Limited Spanish group to 2.47 in the substantial Spanish group.

6.2.2. PLS-4 Spanish Expressive Communication

The results of the analysis of PLS-4 Spanish Expressive Communication standard scores are listed in the middle set of columns in Table 4. The classroom language type groups did not reliably differ in overall level ($F(2,582) = 0.93$, $p > 0.05$) or in rate of change over time ($F(2,582) = 0.85$, $p > 0.05$) in their Spanish expressive skills as measured in standard scores (note: no change means that children are showing normative rates of growth relative to the norming population). On average, children's scores were close to the norming average, with group means that ranged from 97.6 to 99.7, and did not show gains in their standard scores over time (i.e., maintained their relative position compared to the norming population).

6.2.3. PPVT-4

The results of the analysis of PPVT-4 English Receptive Language standard scores are listed in the final set of columns in Table 4. These results describe the analysis of 3612 assessments of DLL children ($n = 1820$) with 1–7 repeated assessments in 247

classrooms in which only English was spoken, 118 classrooms in which English and some Spanish was spoken, and 126 classrooms in which both English and Spanish were spoken. Overall, children showed substantial gains over time ($F(1,893) = 279$, $p < 0.001$). Classroom language type groups did not reliably differ in overall level ($F(2,893) = 1.38$, $p > 0.05$) or in rate of change over time ($F(2,893) = 1.25$, $p > 0.05$). On average, children's scores were about 85 on this standardized measure with a mean of 100 and standard deviation of 15, but showed gains of about 5 points per year they were enrolled in Educare.

6.3. Follow-up analyses

Robustness analyses were conducted. First, whether the parent immigrated to the United States and whether Spanish was the child's first language were included as covariates. Missing data due to failure to collect parent interviews reduced the sample size and generalizability of these analyses so they were regarded as robustness analyses rather than primary analyses. Analyses yielded the same pattern of findings. The PLS Auditory Comprehension in Spanish was lower, on average, in classrooms using English and some Spanish than in classrooms using English and Spanish ($t(587) = 2.21$, $p = 0.03$), but rates of change in scores were not reliably different across language groups ($F(2,587) = 1.63$, $p = 0.20$). The PLS Expressive Communication in Spanish did not show classroom language type differences ($F(2,584) = 0.95$, $p = 0.39$) or classroom language x time interactions ($F(2,584) = 0.81$, $p = 0.45$). The PPVT Receptive Language Score in English did not show classroom language type differences ($F(2,2208) = 1.87$, $p = 0.17$) or classroom language x time interactions ($F(2,2208) = 1.07$, $p = 0.34$).

Second, we reanalyzed the PPVT data, including only the children with both PPVT and PLS-4 data to allow us to examine whether differences in findings between the English PPVT and the Spanish PLS-4 Auditory Comprehension were due to the inclusion of more children in the PPVT analyses. These analyses included 720 fewer children ($n = 1100$) with 1–7 repeated assessments. Results were also consistent with those reported with the whole sample.

Follow-up analyses also indicated that the DLL children showed substantial gains over time in their PPVT score ($F(1,760)=61.7$, $p<0.001$), but did not show different mean levels ($F(2,760)=0.56$) or rates of change over time ($F(2,760)=0.05$) across the three language exposure groups.

7. Discussion

Using secondary analysis of data collected over multiple years, the current study provided an opportunity to examine the role of classroom language use and child language outcomes with a large number of children across multiple sites and multiple years. This study examined DLL children's growth over time relative to national norms on three measures of language learning (English Receptive Language, Spanish Auditory Comprehension and Spanish Expressive Communication) and compared language skills in English and Spanish for DLL children enrolled in Educare Head Start for children whose teachers spoke English with little or no Spanish, English with some Spanish, and both English and Spanish. Irrespective of classroom group, DLL children showed gains comparable to test norms on all three measures, with gains at rates greater than expected in English and in Spanish Auditory Comprehension (AC). Test norms are adjusted for child age. So, for example, if a child scored 87 one year and 87 a second year, he/she would be considered consistent relative to same-age peers; however, if scores were 87 in year 1 and 92 in year 2 and 97 in year 3, scores would demonstrate the child was gaining relative to same-age peers in the norming sample. In our sample, children's growth in English was large compared to test norms (about 5 points per year they were enrolled in Educare) and occurred across all three classroom groups. There was also greater growth than expected from test norms in Spanish Auditory Comprehension (AC) on average with the children in classrooms with both English and Spanish scoring higher than children in the other classroom language groups. Spanish Expressive Communication scores were comparable to expectations over time, and there were no differences in Spanish Expressive Communication (EC) by classroom language group. The findings suggest children were gaining (English and AC) or holding steady (EC) compared to peers in norming studies. Classroom language group as we defined it in this study did not make a difference for English learning, but did for Spanish, at least for Spanish Auditory Comprehension.

7.1. English language learning

The finding that children in all three of our classroom groups improved in English is consistent with that of Yoshikawa et al. (2013) but is inconsistent with that of several studies that have found the use of home language was associated with gains in English in preschool (Bernhard et al., 2006; Farver et al., 2009; Méndez et al., 2015), grade 1 (Collier & Thomas, 2004; Tong et al., 2011) and grade 2 (Rodríguez, Irby, Brown, Laura-Alecio, & Galloway, 2005). However, in contrast to Vitiello, Downer, and Williford (2011), who reported gains in English inverse with the amount of home language used in the classroom, children in all three of our classroom groups improved considerably in English and across classroom groups at fairly comparable rates. Said another way, children in this study, regardless of L1 language instruction exposure, learned English (L2), the primary language of instruction, at similar rates and all three groups exceeded expectations in English language growth (averaging about 5 points per year more than would be expected by norms), thereby and importantly demonstrating substantial progress in reducing the achievement gap, a gap well documented among low-income DLL children (Espinosa, 2013b). These findings also support the idea that increased Spanish instruc-

tion does not harm or delay English language development (Barnett et al., 2007; Durán et al., 2010).

It has been argued that developing language skills in preschool is dependent on both the quantity (amount) and quality of language input that children receive in early childhood settings, with children experiencing differential growth in home and second languages depending on input (Hammer et al., 2014; Hoff, 2013). The classroom quality measure we used assesses a number of aspects of language and teacher interactions with children and our classrooms were consistently of good quality. Thus, it is possible that English input and interactions children received from teachers in all three instructional contexts in this study were sufficient to support children's English language skills and may have attenuated differences in lexical amount (quantity).

7.2. Spanish Auditory Comprehension

The second major finding of the current study was that DLL children did have better Spanish Auditory Comprehension outcomes when teachers used more extensive Spanish in the classroom, consistent with a number of other studies (Barnett et al., 2007; Collier & Thomas, 2004; Farver et al., 2009; Solari et al., 2011). This finding is not surprising given studies have found improvements in home language with classroom home language instruction. This also supports existing research that shows associations between the amount of input a child receives in a given language and competency in that language (Hammer, Komaroff et al., 2012; Hoff, 2006; Hoff et al., 2012). On average, children in this study improved relative to norms on Spanish Auditory Comprehension, but children in the classrooms where there were more Spanish exceeded other classroom groups in Spanish Auditory Comprehension.

7.3. Spanish Expressive Communication

More puzzling was that we did not find an advantage for Spanish Expressive Communication as a function of classroom group; nor did we find greater growth than expected. One possibility is that the findings may be linked to the extent to which teachers use large and small group instruction compared to engaging in individual conversations in Spanish with children (Hammer, Blair, López, Leong, & Bodrova, 2012). Anecdotally, program staff in the current study suggested that complex language—the deepest language using advanced constructivist cognitions—may not have been occurring in Spanish, even in the more intensive English-Spanish classrooms. It could be that although teachers were often using both English and Spanish for instructional purposes of the classroom, the children themselves tended to express more in English, rather than in the home language, which would be consistent with the tendency of minority language speakers to shift to the mainstream language (Cummins & Swain, 1986; Gort & Pontier, 2013; Guiberson, Barnett, Jancosek, & Yoshinaga-Itano, 2006). Bohman, Bedore, Peña, Mendez-Perez, and Gillam (2010) argue that children's language use (i.e., speaking a language expressively) makes the learner process the language in a way that simply hearing the language receptively does not. It is important to point out that children were making expected progress in Spanish Expressive Communication, but were not exceeding norms or growing in Spanish Expressive Communication more in classrooms where more Spanish was spoken by teachers (our third classroom type). Development of bilingualism is complex. Indeed, longitudinal studies examining the oral language and literacy development of Hispanic DLLs suggest that children continue to develop their home language upon preschool entry, even when exposed to English-only instruction (Hammer, Lawrence, & Miccio, 2008; López, 2012; Tabors, Pérez, & López, 2003), a claim also supported by cross-language transfer of L1 to L2 (Cummins, 1979). Further, because the oral language devel-

opment of Spanish and English in DLL children is interactive and dynamic, complex relations of mutual supports may be apparent across language systems and provoke variability in the acquisition of either language (Brisk & Harrington, 2010).

7.4. The role of contextual variables

The broader context of DLL language development must certainly contribute to the complex relations of mutual supports and input across language systems. In this study, we provided rich descriptive information about classroom and family contextual variables (Tables 1 and 2), such as use of Spanish in the classroom, concentration and number of DLL students in classrooms and classroom quality, site location, and about family contextual variables such as race/ethnicity, home language, child's first language, parental immigrant status, parental food insecurity (as a proxy for deep poverty), maternal depression, child health status, disability status and sex, together with maternal education and age, and child age of entry into Educare. With these variables, we sought to determine and demonstrate the comparability of our three classroom groups and to entertain or eliminate alternative explanations for findings associated with our classroom groups, given our research questions. But, any of these factors could be hypothesized to contribute to the complex relations of mutual supports across language systems.

In our case, family and individual child contextual variables varied less than did classroom or site contextual variables, across the three language group types. DLL children and their families were similar in language background, and on most other variables, with the exception of parent report of child health and maternal education, the latter of which was controlled in the final analysis. However, classroom contextual variables provided more variability. First, the observers reported a high proportion of time in Spanish for all purposes (social, procedural and instructional) in classrooms with higher percentage of time spent speaking Spanish and the reverse. At first blush, this seems obvious, but it did provide a validation of the amount of Spanish in use (or not in use) in the classrooms and demonstrated that when Spanish was in use, it was for a variety of functions. Second, the proportion and number of DLL students increased as amount of Spanish in the classroom increased. We think this finding could be regarded as a limitation and a gateway to further study and so we address it further in the Limitations/Implications for Future Research section below.

Third, there were differences by region in the use of substantial Spanish in the classroom, led by classrooms in the Southeast in this sample. Fourth, classroom quality, while high for all classroom types, was lower in English-only classrooms. The differences by site and classroom quality were controlled in the final analyses. However, we return to site differences in our discussion of Generalizability and again when we turn to Limitations/Implications for Future Research.

7.5. Generalizability

An important question is to which ECE programs does this current study generalize? Which programs might be able to apply the lessons learned? Findings from the current study come from DLLs enrolled in high-quality enhanced Head Start programs within the Educare network; thus, the findings generalize to high-quality programs serving low-income children, Head Start and pre-k programs. Previously, high-quality programs have not been typical in the United States (Cost, Quality, and Child Outcomes Study Team, 1995). However, recently, there has been considerable progress throughout the country in identifying and improving quality of ECE programs, whether through increasing accreditation (McDonald, 2009), state early learning standards and programs (National

Center on Early Childhood Quality Assurance, 2016), new Head Start Program Performance Standards (Administration for Children and Families, 2016), and quality rating and improvement systems (Quality Rating and Improvement Systems Learning Network, 2017). For example, Aikens, Bush, Gleason, Malone, and Tarullo (2016) document significant Head Start quality improvements from 2006 to 2014 in all ECERS-R factors and in at least one CLASS domain. The level of quality of the Educare programs, with ECERS-R scores averaging 5.7 overall is consistent with Tiers 3–5 of Quality Rating and Improvement System programs in many states and with measurement in the Head Start FACES study showing 54% of Head Start programs scored in this range (for Teaching and Interactions factor) and 32% (for Provisions for Learning factor) (Aikens et al., 2016). At the same time as quality improvements in recent years, early childhood programs have also seen an increase in attendance, particularly among young Latino preschoolers, the majority of whom are DLLs: nationally about 60% of low-income Latino preschoolers attend early childhood programs, and in some areas, participation rates reach over 80% (López, Grindal, Zanoni, & George, 2017). Thus, the early childhood field is moving towards a higher preponderance of good quality centers serving linguistically diverse children, providing more opportunities for studies such as ours, conducted in good-quality programs and minimizing confounds due to quality, with a nationally geographically diverse sample of young DLLs, and with increasing applicability of effects of ECE for young diverse learners.

Next, the relatively large sample size with geographic representation of DLL classrooms from five regions of the U.S. adds to the generalizability of the study. The study had representation from multiple regions in each classroom language group and children and families differed little within the language classroom groups; study findings can be generalized to all parts of the country serving DLL children. However, some regions of the U.S. do seem to be leading the way in providing substantial Spanish instruction, a topic we return to under Limitations/Implications for Future Research.

7.6. Limitations/implications for future research

Several limitations should be noted in regard to this study. In a number of cases, we think the limitation applies to or is due to the state of knowledge within the field of DLL classroom studies. Despite limitations, we believe our study contributes to the knowledge base, as we note. Each limitation also makes apparent opportunities for further study in this field.

A major limitation is that data are correlational in nature. While there have been some experimental studies to examine L1 and L2 effects (Bernhard et al., 2006; Durán et al., 2010; Farver et al., 2009; Matera & Gerber, 2008; Méndez et al., 2015) as we have noted, the many variables potentially requiring study in terms of both inputs and outcomes has thus far restricted definitive conclusions among studies. For example, Buysse et al. (2014) noted experimental variability in design, definition of DLLs, assessment of DLLs (one language or two, receptive or expressive), measurement of language instruction in classroom-specific activities vs. DLL use throughout the day, and sample size. Thus, studies with large samples such as the current one are helpful in identifying operative elements that represent everyday early childhood environments today (e.g., what languages do teachers speak in the classroom). However, we acknowledge the need for more causal studies and identify how experimental research in the future could more systematically be used to overcome limitations of our specific study as well as to expand the knowledge base in this critical area.

Our broad categorization of dual language input/instruction was a limitation. We identified three general categories of dual language use and one category included a range from 26% to 90% dual-language. We have argued that our three categories are edu-

cationally meaningful, including whether L1 language by teachers was little to none, limited or more extensive. From the point of view of a DLL child, hearing his/her home language a quarter of the time (or more) constitutes a very different experience from rarely hearing the home language in the classroom. However, questions remain about the precise language goals, language use of caregivers throughout the school year, and variation across multiple early childhood sites in our study, and these questions suggest directions for future research, including experimental studies. While ECE programs within our study shared general goals to advance English language learning while encouraging Spanish language growth, the *specific* goals for DLL early education may have varied considerably (Garcia, 2011). Moreover, precisely *how* instructional models and strategies are implemented to support DLL goals remains unclear. Our study did employ a supplemental analysis to determine if purpose (e.g., social, procedural, or instructional) of Spanish use in the classrooms added information, but associations among purpose and type were highly correlated, and although this analysis validated our categorization, this additional measure did not add to the precision of our analyses and therefore was not included in the final statistical model. Regarding reliability of DLL instruction measurement, our measure of language use was checked by teachers and with the use-of-Spanish-in-the-classroom question, as noted, but not by formal means. While we were confident that Spanish-use statistics in Table 1 verified our categories, lack of a formal reliability process may be regarded as a limitation of the study. In the future, reliability work can be aided by more consistent definitions of DLL instructional models as we describe later in this section (Administration for Children and Families, 2017).

Next, we recommend experimental and other work to examine issues that we identified as limits of the current study, for example, whether there are critical cut points in regard to quantity of L1 that children need to hear to demonstrate growth (in accordance with national norms) in both languages over time; if using L1 and L2 is more important in some areas of instruction than in others; or if intentionally using L1 in conversation impacts expressive L1 growth. We also recommend more study in areas that, while beyond the scope of the current study, potentially have bearing on the outcomes of interest here, including if L1 or L2 classroom use matters more if children are sequential vs. simultaneous learners of English, and how teacher attitudes in regard to bilingualism may affect outcomes in either L1 or L2, to name some possibilities. We also suggest more research exploring the fascinating questions regarding the choices children make about whether and how to express themselves in Spanish or English in the context of different instructional models.

We note variation within our own sample that raises further questions for future study. Within the English–Spanish classrooms in the current study there were two programs that practiced explicit dual-language programs (half-day English and half-day Spanish instruction for all children), while other English–Spanish programs were less explicit. Future studies could document implementation for precise purpose, time and use of instructional Spanish, quality of teacher–child interactions conducted in both languages, family members' beliefs about the purpose of home language use in the school, children's perceptions of what language should be spoken at school, children's language choices when given the opportunity and for which types of purposes, and explicit reports of children's higher-order language use in both languages. Our questions may not exhaust the possibilities, but the current study does demonstrate that there is a great deal more to be learned about how instruction is offered. Using the classroom language models (CLMs; i.e., dual immersion; home language with English support; English with home language support, and use of English-only) proposed in the Administration on Children and Families' recent policy statement on supporting DLL in early childhood pro-

grams would be a useful beginning point for conceptualizing DLL instruction in future research studies (Administration for Children and Families, 2017).

Additionally, there is a need to understand the role classroom quality plays for young DLL children's outcomes in both languages (Peisner-Feinberg et al., 2014). Altogether, we think the study reported here may reveal more general challenges in understanding how DLL instruction is measured and represented and offers the field a challenge to think about how to obtain a more discriminating way of characterizing and studying DLL models of instruction, while taking into account both general aspects of ECE quality, and specific aspects of quality that may be particularly critical for young DLLs to benefit from classroom experiences. For example, we suggested that the good quality ratings on the ERS may have meant that children had English interactions that were “sufficiently rich” in English lexical diversity and lexical amount to support English-language growth but future research could test this assumption.

The current study was limited by the nature of the language assessments available. It is important to consider reasonable measurement burden for young children in the context of an ECE program, a burden that is increased when children's language development in two areas of two languages is ultimately of interest. The Educare Learning Network selected an in-depth measure of auditory and expressive Spanish and a measure of receptive language in English for all children (PPVT-4), the measures used for the current study. The PLS-5 was not available when our study began, but also we used data collected with PLS-4 because the PLS-5 does not provide a report of scores standardized for Spanish, only for comprehensive Spanish and English together. This is a drawback in the measurement of DLL children's progress in the home language. We acknowledge that the combination English-Spanish score provided by the PLS-5 makes a contribution by providing a comprehensive score of the child's English AND Spanish, albeit a different one from the standardized scores sought for Spanish only for the current study. Moreover, the opportunity to study DLLs in languages other than Spanish (e.g., Tagalog, Haitian Creole) is even more hampered by a lack of linguistically and culturally sensitive measurement tools available in languages other than English (Snow & Van Hemel, 2008). While we argue that the field is constrained by a lack of measures for assessing dual-language learners in L1 and L2 in everyday program contexts, the current study analyzed data from a large sample of DLL's in both auditory and expressive Spanish, as well in English Receptive Language, using reliable measures with good predictive validity. While there are criticisms of using normed assessments for charting classroom progress of DLL children (Barrueco et al., 2012), the current study documented gains of DLL children compared to norms by using standardized measures, while informing the focus on reducing the achievement gap across the Educare network. Nonetheless, more work is needed on valid, low-burden, equivalent measures of children's language progress in L1 and L2 that can inform classroom practice, the research literature and policy in regard to English, Spanish and other language use found in programs serving low-income preschool-age children, consistent with a key recommendation of Administration for Children and Families policy statement on supporting DLL language growth (Administration for Children and Families, 2017).

As we have noted, descriptive findings (see Table 1) seemed to suggest that DLL child concentration could be somewhat associated with the type of DLL instruction as we defined it in this study, with higher concentrations of DLL children in classrooms with more Spanish language use. It is possible that the higher concentrations of Spanish-speaking peers contributed to the relatively greater Spanish Auditory Comprehension growth we saw in this study. However, we would have projected that if peer Spanish language use was driving Spanish language growth, it would have been in both Expressive Communication and Audi-

tory Comprehension, not just Auditory Comprehension. Relatively greater growth in Auditory Comprehension seems to imply that children are expanding Spanish vocabulary by listening (e.g., to the teacher). Nonetheless, the descriptive findings about greater concentrations of DLL students in Spanish-use classrooms do suggest further research questions about peer effects and as to whether higher concentrations of DLL students drive greater Spanish use in the classroom, or, conversely, if children/families are attracted to classrooms with higher concentration of Spanish use/Spanish speakers. These questions were not within the purview of our study, and we suspect that they are underpinned by geographic differences in populations (and the need for still broader contextual studies). Geographic differences are underscored by our descriptive finding of more substantial Spanish use in the Southeast classrooms, possibly because of more established DLL populations and availability of and/or more longstanding employment of Spanish-speaking teaching staff, which may in turn attract the enrollment of more DLL children. Altogether, questions about context call for future study.

7.7. Implications for practice

Taken together with other research, findings from this study have implications for practice. First, Spanish-speaking children seem to learn English in good quality classrooms irrespective of classroom type—whether mostly English, English and some Spanish, or both English and Spanish in combination. The current study verifies alternative teacher language approaches for English learning that have been found in previous studies (e.g. Barnett et al., 2007; Farver et al., 2009; Yoshikawa et al., 2013) as determined by direct observations of classroom language use. Our finding is important as it highlights that DLL children learn English at equal (and advanced) rates regardless of L2 classroom exposure, when in high-quality classrooms, and that Spanish use in the classroom at varying levels does not impede English acquisition.

Some programs incorporate Spanish in the classroom as a way to help communicate with the child, or to facilitate English language development. However, many programs have the deeper value for children to not only maintain home language, but also to grow in its use and understanding, i.e. to become truly bilingual. Given the large body of research that demonstrates an array of cognitive advantages of bilingualism, including in executive functioning (e.g. Barac, Bialystok, Castro, & Sanchez, 2014; Bialystok & Majumder, 1998; Carlson & Meltzoff, 2008; White & Greenfield, 2017), the importance of fostering dual-language development in DLL populations is being increasingly recognized. Taken together with our findings on the positive associations with Spanish use in the classroom and DLL children's Spanish language development (at least receptively), one implication of findings from this study is the need to increase techniques and teacher training for Spanish use in classrooms, with a particular focus on encouraging children to express themselves in their home language.

On this note, the current study demonstrates that children in high-quality classrooms did maintain Spanish in both Auditory Comprehension and Expressive Communication. But only relatively greater progress in Auditory Comprehension was associated with more Spanish used in the classroom. Thus, the study challenges the field to demonstrate strategies that lead to greater growth in expressive Spanish, as well as receptive, for young DLL children. That is, another implication is that training for teachers who are learning Spanish (or whatever the home language of the DLL population of interest may be) and preparing to work with young DLLs needs to be more fine grained so that teachers learn not just how to speak the children's home language, but also how to promote child conversation and creative and conceptual expression in L1, and learn to move past the tendency reported by others for lim-

ited classroom conversations in L1 (Cummins & Swain, 1986; Gort & Pontier, 2013; Guiberson et al., 2006).

This study also demonstrates that the ECE programs seeking home-language-speaking teachers are on the right track for promoting DLL children's outcomes. Presumably, programs serving children with more than two languages could provide similar benefits if teachers speaking in other languages within a classroom deliver consistent high-quality interactions in a way that benefits the language environment of the classroom, but that remains to be tested. The current study did not provide evidence of any advantage to a classroom of having someone who speaks "some Spanish," as children grew comparably in both English and Spanish Expressive Communication in all three classroom types. Rather, the study does suggest that for children to demonstrate positive outcomes beyond the norm in the home language, at least when it comes to Auditory Communication, they may need more intensive Spanish language in the classroom as we saw in our most balanced English and Spanish classroom language type.

7.8. Contributions of the current study

This study was conducted with a relatively large sample of Spanish-speaking DLL children attending a high-quality, expanded Head Start program model aiming for DLL students' growth in both English and Spanish. Interpretable findings from a relatively large sample are important given a field of small studies with varying methodologies demonstrating mixed findings, and the relatively large sample from many regions of the U.S. contributes to geographical generalizability, which proves unique among existing DLL studies. This study demonstrates that quality dual-language preschool was beneficial for DLL children's Spanish (L1) and English (L2), and was even more beneficial for L1 Auditory Comprehension language skills with substantial L1 inputs in the classroom, altogether going beyond studies that (1) only examine outcomes in one language (e.g., Buysse et al., 2014; Restrepo et al., 2010), (2) those criticized for limited generalizability due to small sample sizes (Buysse et al., 2014), or (3) those with L1 and L2 inputs confounded with classroom quality.

These findings are also important because they address concerns about the language and academic achievement of low-income Spanish/English DLL children, a group that traditionally performs below monolingual peers on language and reading assessments once they reach school (Hoff, 2013). The study adds to the case that DLL children's language skills grow considerably in quality preschool programs. Put another way, the study contributes to the case that limited or no exposure to quality pre-kindergarten ECE programming places DLL children at a disadvantage with regards to first (L1) and second (L2) language development, jeopardizing their more global preparedness for elementary school. The current study clearly points to patterns of change and calls for replication with improvements in the measurement of DLL children's language development across their full language repertoire in future national data sets examining early childhood education participation.

Author note

The Educare Learning Network Investigative Team consists of the local evaluators at each Educare school who supplied data. For the data in this paper, they were: Jane Atwater, University of Kansas; Lori Ellingford, Arizona State University; Sheridan Green, Clayton Early Learning, Denver; Daryl Greenfield, University of Miami; Miriam Hirschstein, University of Washington; Diane Horm, University of Oklahoma-Tulsa; Brenda Jones-Harden, University of Maryland; Brian Mangus, Anselm Center, Edmund, Oklahoma; Mark Nagasawa, the Erikson Institute; Nancy Perry, Arizona State University; Lisa St. Clair, University of Nebraska

Medical Center; Susan Spieker, University of Washington; Mary Sweet-Darter, Anselm Center, Edmund Oklahoma; Helen Raikes, University of Nebraska-Lincoln, and Dale Walker, University of Kansas.

Acknowledgments

We gratefully acknowledge the funding support for this work provided by the Buffett Early Childhood Fund and the George Kaiser Family Foundation. The content of this publication does not necessarily reflect the views or policies of the Buffett Early Childhood Fund or the George Kaiser Family Foundation. We also thank Karen Taylor and John Cashwell for valuable assistance with the data for this paper. Finally, we thank the programs for providing the data, including the Educare staff, families, and children.

References

- Administration for Children and Families. (2008). *Dual language learning: What does it take? Head Start dual language report*. Washington, DC: U.S. Department of Health and Human Services. Retrieved from: http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/teaching/eecd/Individualization/Learning%20in%20Two%20Languages/DLANA_final_2009%5B1%5D.pdf
- Administration for Children and Families. (2010). *Head Start impact study. Final report*. Washington DC: Department of Health and Human Services. Retrieved from: <http://www.acf.hhs.gov/opre/resource/head-start-impact-study-final-report>
- Administration for Children and Families. (2013). *Report to Congress on dual language learners in Head Start and Early Head Start programs*. Washington, DC: Department of Health and Human Services. Retrieved from: <http://www.acf.hhs.gov/sites/default/files/opre/report.to.congress.pdf>
- Administration for Children and Families. (2016). *Head Start Program Performance Standards (HSPPS) final rule ACF-PH-16-04*. Washington, DC: Department of Health and Human Services. Retrieved from: <http://eclkc.ohs.acf.hhs.gov/policy/pi/acf-pi-hs-16-04>
- Administration for Children and Families. (2017). *Policy statement on supporting the development of children who are dual language learners in early childhood programs*. Washington, DC: Department of Health and Human Services. No: ODAS; ECD-ACF-PS-2017-2. Retrieved from: https://www.acf.hhs.gov/sites/default/files/ecdf/dll_policy_statement_final.pdf
- Aikens, N., Bush, C., Gleason, P., Malone, L., & Tarullo, L. (2016). *Tracking quality in Head Start classrooms: FACES 2006 to FACES 2014 research brief (OPRE Report #2016-82)*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families Office of Planning, Research and Evaluation. Mathematica Policy Research.
- August, D., & Shanahan, T. (2006). *Executive summary: Developing literacy in second-language learners: A report of the National Literacy Panel on Language-Minority Children and Youth*. Mahway, NJ: Lawrence Erlbaum Associates.
- Barac, R., Bialystok, E., Castro, D. C., & Sanchez, M. (2014). The cognitive development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, 29, 699–714.
- Barnett, W. S., Yarosz, D. J., Thomas, J., Jung, K. H., & Blanco, D. (2007). Two-way and monolingual English immersion in preschool education: An experimental comparison. *Early Childhood Research Quarterly*, 22, 277–293.
- Barrueco, S., López, M., Ong, C., & Lozano, P. (2012). *Assessing Spanish-English bilingual preschoolers: A guide to best measures and approaches*. Baltimore, MD: Paul H. Brookes.
- Beretvas, S. N. (2011). Cross-classified and multiple membership models. In *Handbook of advanced multilevel analysis*. pp. 313–334. New York: NY: Routledge.
- Bernhard, J. K., Cummins, J., Campoy, I., Ada, A., Winsler, A., & Bleiker, C. (2006). Identity texts and literacy development among preschool English language learners: Enhancing learning opportunities for children at risk of learning disabilities. *Teachers College Record*, 108(11), 2380–2405.
- Bialystok, E., & Majumder, S. (1998). The relationship between bilingualism and the development of cognitive processes in problem solving. *Applied Linguistics*, 19, 69–85.
- Bohman, T. M., Bedore, L. M., Peña, E. D., Mendez-Perez, A., & Gillam, R. B. (2010). What you hear and what you say: Language performance in Spanish-English bilinguals. *International Journal of Bilingual Education and Bilingualism*, 13, 325–344.
- Brisk, M. E., & Harrington, M. M. (2010). *Literacy and bilingualism: A handbook for all teachers*. Mahwah, NJ: Erlbaum Associates.
- Burchinal, M., Roberts, J., Nabors, L., & Bryant, D. (1996). Quality of center child care and infant cognitive and language development. *Child Development*, 67, 606–620.
- Burchinal, M., Peisner-Feinberg, E., Bryant, D. M., & Clifford, R. (2000). Children's social and cognitive development and child-care quality: Testing for differential associations related to poverty, gender, or ethnicity. *Applied Developmental Science*, 4(3), 149.
- Burchinal, M., Field, S., López, M. L., Howes, C., & Pianta, R. (2012). Instruction in Spanish in pre-kindergarten classrooms and child outcomes for English language learners. *Early Childhood Research Quarterly*, 27(2), 188–197.
- Buyse, V., Peisner-Feinberg, E., Pérez, M., Hammer, C. S., & Knowles, M. (2014). Effects of early education programs and practices on the development and learning of dual language learners: A review of the literature. *Early Childhood Research Quarterly*, 29, 765–785.
- Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, 11, 282–298.
- Castro, D. C., Espinosa, L. M., & Pérez, M. M. (2011). Defining and measuring quality early childhood practices that promote dual language learners' development and learning. In M. Zaslow, I. Martinez-Beck, K. Tout, & T. Halle (Eds.), *Quality measurement in early childhood settings* (pp. 257–280). Baltimore MD: Paul H. Brookes.
- Chang, F., Crawford, G., Early, D., Bryant, D., Howes, C., Burchinal, M., et al. (2007). Spanish-speaking children's social and language development in pre-kindergarten classrooms. *Early Education and Development*, 18(2), 243–269.
- Collier, V. P., & Thomas, W. P. (2004). The astounding effectiveness of dual language education for all. *NABE Journal of Research and Practice*, 2(1), 1–20.
- Collins, B. A. (2014). Dual language development of Latino children: Effect of instructional program type and the home and school language environment. *Early Childhood Research Quarterly*, 29(3), 389–397.
- Cost, Quality, and Child Outcomes Study Team. (1995). *Cost, Quality, and child outcomes in child care centers, public report* (2nd ed.). Denver, CO: Economics Department University of Colorado at Denver.
- Cummins, J., & Swain, M. (1986). *Bilingualism in education*. London: Longman.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49, 222–251.
- Currie, J., & Thomas, D. (1996). *Does Head Start help hispanic children?* Cambridge, MA: Mimeo, National Bureau of Economic Research, September.
- Dickinson, D. K., McCabe, A., Clark-Chiarelli, N., & Wolf, A. (2004). Cross-language transfer of phonological awareness in low-income Spanish and English bilingual preschool children. *Applied Psycholinguistics*, 25, 323–347.
- Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4 manual*. Bloomington, MN: NCS Pearson, Inc.
- Dunn, L. M., Padilla, E., Lugo, D., & Dunn, D. M. (1986). *Test de Vocabulario en Imágenes Peabody*. Circle Pines, MN: American Guidance Service.
- Durán, L. K., Roseth, C. J., & Hoffman, P. (2010). An experimental study comparing English-only and transitional bilingual education on Spanish-speaking preschoolers' early literacy development. *Early Childhood Research Quarterly*, 25(2), 207–217.
- Educare Learning Network [ELN]. (2015). *What is Educare?* Retrieved from: <http://www.educarenetwork.org/about/what-is-educare.php>
- Espinosa, L. M., & Magruder, E. S. (2015). Practical and proven strategies for teaching young dual language learners. In L. M. Espinosa (Ed.), *Getting it right for young children from diverse backgrounds: Applying research to improve practice with a focus on dual language learners* (2nd ed., pp. 76–113). Upper Saddle River, NJ: Pearson.
- Espinosa, L. M. (2013). *PreK-3rd: Challenging common myths about dual language learners: An update to the Seminal 2008 Report* (10). New York, NY: Foundation for Child Development. Retrieved from: <http://fcd-us.org/sites/default/files/Challenging%20Common%20Myths%20Update.pdf>
- Espinosa, L. M. (2018). *Early education for dual language learners: Promoting school readiness and early school success*. Migration Policy Institute. Retrieved from: <http://fcd-us.org/sites/default/files/DualLanguageLearners.pdf>
- Farver, J. A. M., Lonigan, C. J., & Eppe, S. (2009). Effective early literacy skill development for young Spanish-speaking English language learners: An experimental study of two methods. *Child Development*, 80(3), 703–719.
- Federal Interagency Forum on Child and Family Statistics. (2015). *America's children: Key national indicators of well-being, 2015*. Washington, DC: U.S. Government Printing Office.
- García, E. (2011). Classroom experiences and learning outcomes for dual language learning children: An early childhood agenda for educators and policy. In C. Howes, J. Downer, & R. Pianta (Eds.), *Dual language learners in the early childhood classroom* (pp. 1–16). Baltimore MD: Paul H. Brookes.
- Goldenberg, C., Hicks, J., & Lit, I. (2013). Dual language learners: Effective instruction in early childhood. *American Educator*, 37(2), 26–29.
- Goldstein, H. (1994). Multilevel cross-classified models. *Sociological Methods & Research*, 22(3), 364–375.
- Gormley, W. T., Jr., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Developmental Psychology*, 41(6), 872–884.
- Gormley, W. T. (2008). The effects of Oklahoma's pre-k program on Hispanic children. *Social Science Quarterly*, 89, 916–936.
- Gort, M., & Pontier, R. W. (2013). Exploring Bilingual pedagogies in dual language preschool classrooms. *Language and Education*, 27(3), 223–245.
- Guiberson, M., Barrett, K. C., Jancosek, E. G., & Yoshinaga-Itano, C. (2006). Language maintenance and loss in young preschool age children of Mexican immigrants: Longitudinal study. *Communication Disorders Quarterly*, 28, 4–14.
- Guss, S. S., Norris, D. J., Horm, D. M., Monroe, L. A., & Wolfe, V. (2013). Lessons learned about data utilization from classroom observations. *Early Education & Development*, 24(1), 4–18.
- Halle, T. G., Whittaker, J. V., Zepeda, M., Rothenberg, L., Anderson, R., Daneri, P., et al. (2014). The social-emotional development of dual language learners: Looking back at existing research and moving forward with purpose. *Early Childhood Research Quarterly*, 29, 734–749.

- Hammer, C. S., Lawrence, F. R., & Miccio, A. W. (2008). The effect of summer vacation on bilingual preschoolers' language development. *Clinical Linguistics & Phonetics*, 22(9), 686–702.
- Hammer, C. S., Hoff, E., Uchikoshi, Y., Gillanders, C., Castro, D. C., & Sandilos, L. E. (2014). The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, 29, 715–733.
- Hammer, C. S., Blair, C., López, L., Leong, D., & Bodrova, E. (2012, March). Tools of the Mind: Promoting the school readiness of ELLs. In D. C. Farran (Chair) (Ed.), *Enhancing executive function and achievement in prekindergarten classrooms: The effectiveness of Tools of the Mind*. Symposium conducted at the meeting of the Society for Research on Educational Effectiveness, Washington DC.
- Hammer, C. S., Komaroff, E., Rodriguez, B. L., López, L. M., Scarpino, S. E., & Goldstein, B. (2012). Predicting Spanish–English bilingual children's language abilities. *Journal of Speech, Language, and Hearing Research*, 55(5), 1251–1264.
- Harms, T., Clifford, R. M., & Cryer, D. (2005). *Early childhood environment rating scale—Revised*. New York, NY: Teachers College Press.
- Harms, T., Cryer, D., & Clifford, R. M. (2006). *Infant/toddler environment rating scale—Revised edition*. New York, NY: Teachers College Press.
- Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, 39(1), 1–27.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26(1), 55–88.
- Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing the achievement gaps. *Developmental Psychology*, 49(1), 4–14.
- López, L. M., & Greenfield, D. B. (2004). The cross-language transfer of phonological skills of Hispanic Head Start children. *Bilingual Research Journal*, 28(1), 1–18.
- López, M., Grindal, T., Zanoni, W., & George, R. (2017). *Hispanic children's participation in early care and education: A look at utilization patterns of Chicago's Publicly Funded Programs (Publication #2017-20)*. Washington, D.C: National Research Center on Hispanic Children & Families.
- López, L. M. (2012). Assessing the phonological skills of bilingual children from preschool through kindergarten: Developmental progression and cross-language transfer. *Journal of Research in Childhood Education*, 26(4), 371–391.
- Méndez, L. L., Crais, E. R., Castro, D. C., & Kainz, K. (2015). A culturally and linguistically responsive vocabulary approach for young Latino dual language learners. *Journal of Speech Language and Hearing Research*, 58, 93–106.
- Matera, C., & Gerber, M. M. (2008). Effects of a literacy curriculum that supports writing development of Spanish-speaking English learners in Head Start. *NHSA Dialog*, 11(1), 25–43.
- Matthews, H. (2011). *Meeting the early learning challenge: Supporting English language learners*. Washington, DC: CLASP.
- McDonald, D. (2009). *Elevating the field: Using NAEYC early childhood program accreditation to support and reach higher quality in early childhood programs*. Washington, DC: NAEYC.
- National Association for the Education of Young Children. (1995). *Responding to linguistic and cultural diversity recommendations for effective early childhood education: A position statement*. Washington, DC: National Association for the Education of Young Children.
- National Association for the Education of Young Children. (2009). *Responding to linguistic and cultural diversity recommendations for effective early childhood education: A position statement*. Washington, DC: National Association for the Education of Young Children.
- National Center on Early Childhood Quality Assurance. (2016). *Early learning standards and guidelines (Publication No. 075)*. Washington, DC: Department of Health and Human Services. Retrieved from https://childcareta.acf.hhs.gov/sites/default/files/public/state_elgs_web_final_2.pdf
- National Task Force on Early Childhood Education for Hispanics. (2007). *Para nuestros niños: Expanding and improving early education for Hispanics..* Retrieved online from <http://fcd-us.org/resources/para-nuestros-ninos-expanding-and-improving-early-education-hispanics>
- Paradis, J., Nicoladis, E., Crago, M., & Genesee, F. (2011). Bilingual children's acquisition of the past tense: A usage-based approach. *Journal of Child Language*, 38(3), 554–578.
- Pearson, B. Z., Fernandez, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, 18(1), 41–58.
- Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L., et al. (1999). *The children of the cost, quality and outcomes study go to school: Executive summary*. Chapel Hill, NC: Frank Porter Graham Child Development Center at the University of North Carolina.
- Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L., et al. (2001). The relation of preschool child-care quality to children's cognitive and social development trajectories through second grade. *Child Development*, 72(5), 1534–1553.
- Peisner-Feinberg, E., Buysse, V., Fuligni, A., Burchinal, M., Espinosa, L., Halle, T., et al. (2014). Using early care and education quality measures with dual language learners: A review of the research. *Early Childhood Research Quarterly*, 29(4), 786–803.
- Phillipsen, L., Burchinal, M., Howes, C., & Cryer, D. (1998). The prediction of process quality from structural features of child care. *Early Childhood Research Quarterly*, 12, 281–303.
- Quality rating and improvement systems learning network. (n.d.). About QRIS [Web information]. Retrieved from <http://qrisnetwork.org/>.
- Restrepo, M. A., Castilla, A. P., Schwanenflugel, P. J., Neuhauser-Pritchett, S., Hamilton, C. E., & Arboleda, A. (2010). Effects of a supplemental Spanish oral language program on sentence length, complexity, and grammaticality in Spanish-speaking children attending English-only preschools. *Language, Speech, and Hearing Services in Schools*, 41(1), 3–13.
- Rodriguez, J., Irby, B. J., Brown, G., Lara-Alecio, R., & Galloway, M. (2005). An analysis of second grade reading achievement related to pre-kindergarten Montessori and transitional bilingual education. *Review of Research and Practice*, 3, 45–65.
- Rumberger, R. W., & Tran, L. (2010). State language policies, school language practices, and the ELL achievement gap. In P. Gándara, & M. Hopkins (Eds.), *Forbidden language: English learners and restrictive language policies* (pp. 86–101). New York, NY: Teachers College.
- Shatz, M., & Wilkinson, L. C. (Eds.). (2010). *The education of English language learners: Research to practice*. New York, NY: The Guilford Press.
- Snow, C. E., & Van Hemel, S. B. (Eds.). (2008). *Early childhood assessment: Why, what and how*. Washington, DC: The National Academies Press.
- Solari, E. J., Landry, S. H., Zucker, T. A., & Crawford, A. (2011). The importance of sensitive measurement tools for understanding what instructional practices promote school readiness for dual language learners. In C. Howes, J. T. Downer, & R. C. Pianta (Eds.), *Dual language learners in the early childhood classroom* (pp. 45–68). Baltimore, MD: Paul H. Brookes.
- Stein, A., Freel, K., Hanson, A. T., Pacchiano, D., & Eiland-Williford, B. (2013). The Educare Chicago research-program partnership and follow-up study: Using data on program graduates to enhance quality improvement efforts. *Early Education & Development*, 24(1), 19–41.
- Tabors, P. O., Pérez, M. M., & López, L. M. (2003). Dual language abilities of Spanish-English bilingual four-year olds: Initial finding from the early childhood study of language and literacy development of Spanish-speaking children. *NABE Journal of Research and Practice*, 1, 70–91.
- Tong, F., Lara-Alecio, R., Irby, B. J., & Mathes, P. G. (2011). The effects of an instructional intervention on dual language development among first-grade Hispanic English-learning boys and girls: A two-year longitudinal study. *The Journal of Educational Research*, 104, 87–99.
- U.S. Department of Health and Human Services. (2012). *Advisory Committee on Head Start research and evaluation final report*. Washington, DC: Author.
- Vitiello, V. E., Downer, J. T., & Williford, A. P. (2011). Preschool classroom experiences of dual language learners: Summary of finding from publicly funded programs in 11 states. In C. Howes, J. Downer, & R. Pianta (Eds.), *Dual language learners in the early childhood classroom* (pp. 69–92). Baltimore, MD: Paul H. Brookes.
- White, L. J., & Greenfield, D. B. (2017). Executive functioning in Spanish- and English-speaking Head Start preschoolers. [Special Issue Article]. *Developmental Science*, 20(1), 1–14.
- Whitebook, M., Howes, C., & Phillips, D. (1989). *Who cares? Child care teachers and the quality of care in America. National child care staffing study*. Oakland, CA: Child Care Employee Project.
- Winsler, A., Burchinal, M. R., Tien, H. C., Peisner-Feinberg, E., Espinosa, L., Castro, D. C., et al. (2014). Early development among dual language learners: The roles of language use at home, maternal immigration, country of origin, and socio-demographic variables. *Early Childhood Research Quarterly*, 29, 750–764.
- Winsler, A., Kim, Y. K., & Richard, E. R. (2014). Socio-emotional skills, behavior problems, and Spanish competence predict the acquisition of English among English language learners in poverty. *Developmental Psychology*, 50(9), 2242–2254.
- Yazejian, N., Bryant, D., & Kennel, P. (2013). Implementation and replication of the Educare model of early childhood education. In T. G. Halle, A. J. Metz, & I. Martinez-Beck (Eds.), *Applying implementation science in early childhood programs and systems* (pp. 209–225). Baltimore, MD: Paul H. Brookes.
- Yazejian, N., Bryant, D., Freel, K., Burchinal, M., & the Educare Learning Network (ELN) Investigative Team. (2015). High-quality early education: Age of entry and time in care differences in student outcomes for English-only and dual language learners. *Early Childhood Research Quarterly*, 32, 23–39.
- Yazejian, N., Bryant, D., Hans, S., Horm, D., St. Clair, L., File, N., et al. (2017). Child and parenting outcomes after one year of Educare. *Child Development*, <http://dx.doi.org/10.1111/cdev.12688>
- Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W. T., et al. (2013). *Investing in our future: The evidence base on preschool education*. Society for Research in Child Development: Foundation for Child Development.
- Zimmerman, I., Steiner, V., & Pond, R. (2002). *Preschool language scale* (4th ed.). San Antonio, TX: The Psychological Corporation.