MIND, BRAIN, AND EDUCATION

Bilingual Baby: Foreign Language Intervention in Madrid's Infant Education Centers

Naja Ferjan Ramirez[1] and Patricia Kuhl[1]

ABSTRACT— The first years of life represent a unique window of opportunity for foreign language learning. However, key questions are: How much and what kind of foreign language exposure is needed to ignite learning? We conducted a foreign language (English) intervention in four public Infant Education Centers in Madrid, Spain. Intervention children ($N = 126$, ages 7–33.5 months) experienced 18 weeks of daily, hour-long, group English sessions with native-speaking tutors, using a brain-based method of infant language development. Intervention infants were compared to a matched Current Practice Comparison (CPC) group of peers in the same schools. Intervention children outperformed the CPC group, showing rapid gains on measures of English word comprehension and English speech production. Schools' neighborhood wealth was *not* a significant factor in learning. Follow-up analyses show that the language gains were fully retained 18 weeks post-intervention. Children's foreign language skills advance quickly in school using this research-to-practice curriculum.

In today's globalized world, knowing two languages presents important communicative, social, and economic advantages (Callahan & Gándara, 2014), in addition to the well-known cognitive benefits associated with bilingualism, such as advantages in executive functioning (Bialystok, 2011; Carlson & Meltzoff, 2008), metalinguistic awareness (Barac & Bialystok, 2012), and more robust cognitive abilities with increased age (Craik, Bialystok, & Freedman, 2010). Studies show that the young brain is adept at acquiring two languages—brain responses at 11 months of age are equivalent for monolingual English-learning infants and Spanish-English bilingual infants in response to English syllables, whereas only Spanish-English bilinguals show strong brain responses to Spanish syllables (Ferjan Ramirez, Ramirez, Clarke, Taulu, & Kuhl, 2017). Infants are born with the ability to discriminate phonetic units across all languages (Kuhl et al., 2006; Werker & Tees, 1984), but perception narrows at about 12 months of age when first words appear, as infants begin to focus on the sounds contained in the language(s) they hear. Infants growing up in bilingual families achieve language milestones, such as babbling ("bababa" and "dadada") and first words at the same age as their monolingual peers (Hoff et al., 2012; Petitto et al., 2001).

Age of Acquisition

Research on second language learning outcomes (Flege, MacKay, & Meador, 1999; Johnson & Newport, 1989; Mackay & Flege, 2004) and language in international adoptees (Jacobs, Miller, & Tirella, 2010; Mcacham, 2006) show that exposure to a second language during the first 5 years of life produces the best possible outcomes for full acquisition of the second language. These findings are supported by brain research showing that the processing of early acquired second languages is more efficient compared to later acquired second languages (Jasinska & Petitto, 2013; Wartenburger et al., 2003; Weber-Fox & Neville, 1996, 2001). Using voxel-based morphometry (VBM), Mechelli et al. (2004) have shown that exposure to two languages before the age of 5 years results in the highest levels of brain tissue density in areas related to language, memory, and attention. A recent magnetic resonance imaging

[1]Institute for Learning & Brain Sciences, University of Washington

Address correspondence to Naja Ferjan Ramirez, Institute for Learning & Brain Sciences, University of Washington, Box 357988, Seattle, WA 98195-7988, United States; naja@u.washington.edu

(MRI) study shows no differences in brain structure between monolinguals and bilinguals who started acquiring both languages before the age of 3, whereas bilinguals who started acquiring the second language after age 3 had modified brain structure (Klein, Mok, Chen, & Watkins, 2014).

The benefits of early exposure to a second language align well with the concept of "neural commitment" (Kuhl, 2004), especially when timed to occur during the "critical period" for language acquisition (Lenneberg, 1967). Early in development, learners "commit" neural networks to patterns in language input, and this initial coding of native-like patterns may interfere with second language learning at later ages (Flege, 1995; Iverson et al., 2003; Zhang et al., 2009). In infancy, on the other hand, brain networks are not fully developed, interference is weak, and infants can simultaneously set up neural networks for two distinct languages. In agreement with this model, a recent magnetoencephalography (MEG) study indicates that the infant brain is indeed capable of simultaneously "committing" to two languages (Ferjan Ramirez et al., 2017). Thus, young children have the neural capacity to learn two languages from birth. However, this capacity begins to fade over time (see also Berken, Chai, Chen, Gracco, & Klein, 2016).

What Do Young Children Need to Learn Two Languages?

In addition to early exposure, research findings show that positive second language outcomes are dependent, to a large degree, on the quantity (Garcia-Sierra, Ramirez-Esparza, & Kuhl, 2016) and quality (Ramirez-Esparza, Garcia-Sierra, & Kuhl, 2014, 2016) of language input that children receive. These studies show that the prevalence of "parentese" speech in one-on-one social interactions at home is strongly predictive of future language gains and that the strength of brain responses to the critical phonetic units of a language depends on the quantity of input in that language (Garcia-Sierra et al., 2016).

Children learn language from exposure, in ways that are unique to humans, combining pattern detection abilities (Saffran, Aslin, & Newport, 1996) with uniquely human social skills (see Kuhl, 2007 for discussion). Laboratory studies indicate that foreign language learning in infancy can be surprisingly rapid: infants can detect words in foreign languages when both transitional probabilities and stress-related cues are available, in brief exposure sessions (Hay, Pelucchi, Graf-Estes, & Saffran, 2011). Moreover, infants exposed to natural foreign language at 9 months of age in play sessions by a live tutor learn to discriminate foreign language sounds at levels equivalent to infants exposed to that language from birth, and do so in 12 sessions over 5 weeks—a total of 6 hours of foreign language experience. However, no learning occurs if the same material on the same schedule is presented via video, showing that a social setting provides essential enhancements for foreign language learning (Kuhl, Tsao, & Liu, 2003). Short-term foreign language exposure through social interactions in infancy modulates brain responses to foreign language sounds (Conboy & Kuhl, 2011) that are predictive of language skills up to 3 years of age (Kuhl et al., 2008), further supporting the idea that language acquisition in infancy may be "gated" by the social brain (Kuhl, 2007). These findings have important practical implications. However, it is currently unknown how the plethora of findings translates to real-life environments, such as infant education centers.

Second Language Programs for Young Children

Several research studies have evaluated programs for children who learn the majority language as a second language in preschool or elementary school. Second language skills benefit greatly from high-quality language learning environments (Buysse, Peisner-Feinberg, Paez, Hammer, & Knowles, 2014); however it is important to note that most of these studies focus on children 3 years and older and do not consider infants or toddlers. Along similar lines, public school foreign language learning immersion programs typically focus on older children (see e.g., Hermanto, Moreno, & Bialystok, 2012; Muñoz, 2006), and do not capitalize on the abilities of very young children to rapidly and efficiently learn two languages. While researchers and educators agree that high-quality learning environments in the first 3 years of life are critical because they influence later functioning (see National Academy of Sciences, Engineering, and Medicine [NASEM], 2017, chap. 5), to our knowledge, there are no theoretically guided, research-based, scientifically tested methods for infant and toddler second language teaching in early educational (school) settings. Existing private bilingual classes for infants and toddlers are expensive and not widely available, and use methods that have not been rigorously tested.

Bilingual Language Learning in Madrid's Public Infant Education Centers

Many communities around the world aspire to create bilingual programs for very young children. Spain has recently introduced an infant bilingual program in the Community of Madrid's Public Infant Education Centers *(escuelas infantiles)*, serving children 0–3 years of age. The program consists of approximately 2 hours of weekly instruction led by a Spanish-English bilingual teacher who introduces English word labels, sings English nursery rhymes, reads English books, and exposes children to English phrases.

Current Study

One important question is how much, and what kind of foreign language exposure is necessary to ignite foreign language learning in infancy. The aim of the current study was to test a language learning intervention, based on a theoretical model and research in infant language development (Kuhl et al., 2008), against Madrid's current Infant Bilingual Program, over a period of 18 weeks in four public infant education centers in Madrid. Across the four participating centers, we randomly assigned children ($N = 250$; age range: 7–33.5 months) to the "Intervention" or "Current Practice Comparison (CPC)" group. Intervention and CPC group children were matched within each of the four centers. Each day during the 18-week period, Intervention children left their regular classroom, and went to a different room to participate in 1-hour English sessions (Intervention). CPC children experienced 18 weeks of English exposure as usual (see Methods for details on both kinds of exposure). Spanish and English comprehension and English production outcomes were assessed in Intervention and CPC children.

Three specific hypotheses were tested: (a) Intervention children were hypothesized to show greater English *vocabulary comprehension* growth than CPC children. Spanish vocabulary comprehension was hypothesized to grow at the same rate in both groups; (b) for the Intervention group, significant gains were predicted in the frequency of English *vocalizations produced* per hour; and (c) at the end of the 18-week Intervention period, a significantly higher frequency of English vocalizations was predicted in the Intervention compared to CPC group. We also measured language outcomes in relation to age and neighborhood income, and assessed children's retention of English knowledge after an 18-week delay following the Intervention.

METHODS

Participants

The participants were 126 children (62 boys) in the Intervention group and 124 children (66 boys) in the CPC group. Each of the four schools had an Intervention group and a CPC group consisting of 30–36 children. The mean age at the start of the study was 22 months (range 7–33.5 months). Prior to random assignment to the Intervention and CPC groups, children were divided into four age groups: 7–14 months, 14–20.5 months, 20.5–27 months, and 27–33.5 months. Intervention and CPC children were matched on age at the group level, and within each of the four age groups.

The schools were public Infant Education Centers in the Community of Madrid, and served families who lived in the school's neighborhood. Two schools served predominantly low-income neighborhoods (Moratalaz and Carabanchel), and two served mid-income neighborhoods (Arganzuela and Fuencarral-El Pardo), as measured by the average gross domestic product (GDP) per capita (Ayuntamiento de Madrid, 2008). To assess the role of neighborhood GDP per capita (wealth) on English language learning, children were divided into two groups according to their school's neighborhood wealth (low and mid).

Intervention English Play Sessions

The Intervention sessions were highly social and led by "tutors," who were native English speakers and undergraduate students or recent graduates of the University of Washington. None of the tutors had advanced degrees, and none were licensed teachers or professional researchers. Prior to the Intervention, all tutors participated in a 2-week training at the Institute for Learning & Brain Sciences at the University of Washington, during which they received instructions and practiced the method. Sessions were conducted in small groups (four tutors playing with 12 children over 1 hour; children worked with the same four tutors throughout the Intervention), allowing for frequent one-on-one joint engagement and consistent practice of routines. Tutors' language production in the classroom was monitored via bimonthly LENA recordings, which were reviewed by a trained researcher immediately after collection, to assure the fidelity of the Intervention implementation (see Results section and Table S1 in Appendix S1, Supporting Information).

There were six key theory- and research-grounded principles of the Intervention method:

1 Tutors addressed children often to achieve a *high quantity* of English input. The amount of language that children hear is known to affect language acquisition and processing (Hart & Risley, 1995; Weisleder & Fernald, 2013), and studies have shown that infants' brain responses are modulated by the quantity of language that they hear (Garcia-Sierra et al., 2016).
2 Tutors addressed children using "*parentese*" or infant-directed speech, which is characterized as having higher pitch, slower tempo, and exaggerated intonation contours when compared to adult-directed speech. Research with monolingual and bilingual infants shows that the prevalence of parentese at home, particularly in one-on-one interactions, is associated with advanced future language skills (Ramirez-Esparza et al., 2014, 2016), and event-related brain responses are enhanced in infants listening to parentese as opposed to standard speech (Zhang et al., 2011).
3 The learning context was *highly social* with games and activities that prompt face-to-face interaction

(Figure 1) (Kuhl et al., 2003; Roseberry, Hirsh-Pasek, Parish-Morris, & Golinkoff, 2009), and tutors were trained to provide prompt, contingent responses (Bornstein, Tamis-LeMonda, Hahn, & Haynes, 2008). Research shows that infants learn foreign language sounds from live social tutors, but not from videos presenting the same material (Kuhl et al., 2003), and infants' social behaviors during social language exposure correlate significantly with brain measures that indicate second language phonetic learning (Conboy, Brooks, Meltzoff, & Kuhl, 2015).

4 Children were encouraged to "*talk*" (even if it involved just "babbling") and interact. Early vocalizations in back-and-forth exchanges predict future language and those who "babble" earlier advance more rapidly in language learning (Ferguson, Menn, & Stoel-Gammon, 1992). Recent MEG studies show that infants at 7 and 11 months of age activate motor brain areas when listening to speech (Kuhl, Ramirez, Bosseler, Lin, & Imada, 2014), and findings on speech perception suggest that infants may rely on sensorimotor connections when listening to speech (Bruderer, Danielson, Kandhadai, & Werker, 2015). The Intervention method and curriculum encouraged children to vocalize and interact right from the beginning.

5 Children heard English from *multiple native speakers*. The proportion of language input provided by native speakers explains the variation in 2-year-old bilingual's language skills (Place & Hoff, 2011), and recent research indicates that parentese in late-bilingual mothers is more variable and more inconsistent across phonetic contexts, compared to parentese of native-speaking mothers (Fish, Garcia-Sierra, Ramirez-Esparza, & Kuhl, 2017). Research also shows that exposure to multiple speakers improves learning (Rost & McMurray, 2009; Seidl, Onishi, & Cristia, 2014).

6 Instruction was *play-based*, with weekly themes serving as the guide for daily activities. Research shows that young children learn vocabulary better through adult-scaffolded play compared to direct instruction (Han, Moore, Vukelich, & Buell, 2010; Weisberg, Hirsh-Pasek, & Golinkoff, 2013). Children learned through meaningful and engaging contexts, and language content was delivered in a way that ensured distributed exposure, further fostering learning and retention (Childers & Tomasello, 2002).

Madrid Infant Bilingual Program (Current Practice Comparison, CPC)

The Community of Madrid Infant Bilingual Program focuses on activities that are part of the existing everyday (Spanish) program. The additional experiences in English consist of approximately 2 hours of weekly instruction led by a Spanish-English bilingual teacher who introduces simple vocabulary and phrases through typical nursery school activities such as book reading, nursery rhymes, and singing. All teachers in this Madrid program hold the Cambridge B1 or B2 Certificate. The schools have English books, and other English materials, such as posters and wall decor with English words, which are utilized by the teachers during the English instruction period.

Fig. 1. English Intervention session at one of the participating Infant Education Centers in Madrid. The sessions were highly interactive and children were encouraged to participate.

Parents of the Intervention and CPC children were instructed not to purposely expose their children to additional English tutoring experiences outside of the school for the duration of the program, and were asked to verify that they followed this guideline.

Language Measures

The outline of study design is shown in Figure 2. At the start of the study, CPC and Intervention children's baseline Spanish level was assessed with the European Spanish MacArthur-Bates Communicative Developmental Inventory (CDI; López Ornat et al., 2005). At both the start and at the end of the 18-week Intervention period, CPC and Intervention children who were at least 18 months of age participated in a Spanish and English vocabulary comprehension assessment using the Computerized Comprehension Task (CCT; Friend & Keplinger, 2003), an assessment of early word comprehension administered on a touch screen.

In Intervention and CPC children, English language production was assessed with LENA technology (LENA Research Foundation, 2015). Each child wore a LENA vest, with a digital language processor (DLP): a small, light-weight recorder placed into the vest's pocket, designed to optimally record the child's voice and the language they hear. Intervention children were recorded every 2 weeks, on

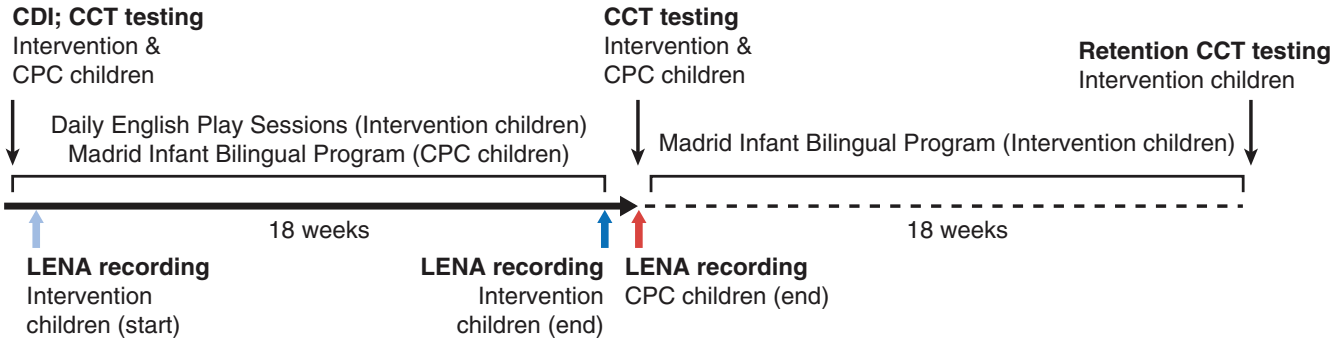

Fig. 2. Outline of study design.

the third day of the week (Wednesday); LENA recordings from the initial (week 1) and final (week 18) weeks are used in analyses reported here. CPC children were recorded with LENA at the end of the 18-week period, and their recordings were obtained in the same environment as the Intervention children's. That is, CPC children were also recorded during an English play session, which used the same format as that used during the Intervention. They were recorded in the same room, interacted with the same tutors who worked with the Intervention children, came to the recording classroom for two days prior to the day they were recorded, and "practiced" putting on the LENA vests to ensure that they were completely comfortable playing in the setting before they were recorded.

LENA produces an automatic, machine-based count of child and nearby adult vocalizations. Because no normative or reliability data exist for classroom LENA recordings or for recordings shorter than 8 hours, the automatic LENA counts were used exclusively to oversee the fidelity of method implementation. To assess children's growth of English, manual coding was performed, and all statistical analyses were conducted on manually coded data. For each child, English vocalizations were identified and divided into five categories: (a) repeated English word (child repeats an English word within 20 s of hearing it from the tutor); (b) repeated English phrase (child repeats an English phrase within 20 s of hearing it from the tutor); (c) spontaneous English word (child spontaneously produces an English word, without hearing it from the tutor for at least 20 s before production); (d) spontaneous English phrase (child spontaneously produces an English phrase, without hearing it from the tutor for at least 20 s before production); and (e) mixed phrase (phrase containing a mix of Spanish and English words [e.g., Eso es "pink"]).

In a follow-up study conducted 18 weeks after the completion of the Intervention, Intervention children were retested on Spanish and English vocabulary comprehension with CCT to measure retention after returning to their normal classrooms (see description under Madrid Infant Bilingual Program) with no additional Intervention sessions. Parents were again instructed not to purposely expose their children to additional English outside of the school during this post-Intervention 18-week period.

Additional information on methods appears in Appendix S1, Supporting Information.

RESULTS

Baseline Measures of Spanish Vocabulary (Communicative Developmental Inventory, CDI)

At the start of the 18-week period, the Intervention and CPC children were statistically equivalent on their knowledge of Spanish as measured by the CDI, with the mean score being the 41st percentile. Statistical comparison can be found in the in Results section in Appendix S1, Supporting Information.

Start- and End-of-Study Measures of English and Spanish Comprehension (Computerized Comprehension Task, CCT)

Mean English and Spanish comprehension scores and standard deviations for the Intervention and CPC children are shown in Table 1.

For English CCT, a repeated measures analysis of variance (ANOVA) with Time (start/end of Intervention) and Group (Intervention/CPC) as independent variables showed a significant main effect of Time, $F(1, 102) = 93.00$, $p < .0001$ (Greenhouse-Geisser), partial $\eta^2 = 0.48$, and an interaction of Time × Group, $F(1, 102) = 9.64$, $p = .002$ (Greenhouse-Geisser), partial $\eta^2 = 0.09$. Follow-up t-tests revealed that the Intervention and the CPC groups' English scores were equivalent at the start of the 18-week period, $t(102) = 0.19$, $p = .8$, $d = 0.04$, confidence interval (CI) (−2.3, 2.8). At the end of the 18-week period, the Intervention group's scores were significantly higher than the CPC group's, $t(54) = -9.91$, $p < .001$, $d = -1.34$, CI (−10.4, −6.9), and reached levels typical of English-learning 20-month-old

Table 1
Mean English and Spanish Computerized Comprehension Task (CCT) Scores and Standard Deviations at the Start and End of the 18-Week Intervention Period

English CCT				*Spanish CCT*			
Intervention		*CPC*		*Intervention*		*CPC*	
Start	*End*	*Start*	*End*	*Start*	*End*	*Start*	*End*
14.7 (6.8)	23.4 (4.5)	14.5 (6.2)	18.8 (5.0)	29.4 (8.0)	34.4 (5.4)	30.1 (6.6)	33.6 (4.8)

Note: CCT = Computerized Comprehension Task; CPC = Current Practice Comparison.

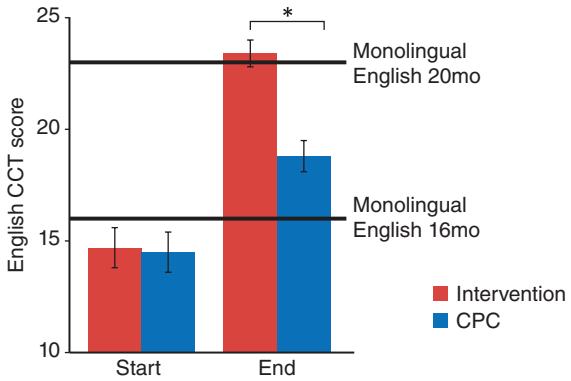

Fig. 3. Mean English Computerized Comprehension Task (CCT) scores at the start and at the end of the 18-week period, for Intervention (red; $n = 55$) and Current Practice Comparison (CPC; blue; $n = 49$) participants 18 months and older. * Indicates significance at $p < .001$. Comparison data for English monolingual children from Friend and Keplinger (2008).

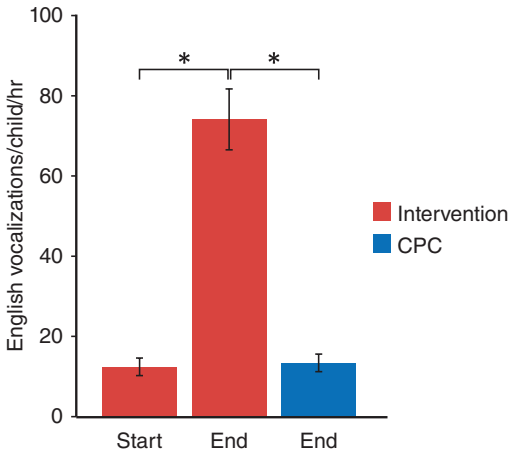

Fig. 4. Mean number of English vocalizations per child per hour for Intervention group (red; $n = 126$) at the beginning and end of the 18-week period, and for Current Practice Comparison (CPC) group (blue; $n = 124$) at the end of the 18-week period. * Indicates significance at $p < .001$.

infants in the United States (Friend & Keplinger, 2008; see Figure 3).

For Spanish CCT, there was also a main effect of Time, $F(1, 102) = 64.69$, $p < .0001$ (Greenhouse-Geisser), partial $\eta^2 = 0.38$, but the interaction of Time × Group was not significant, $F(1, 102) = 2.20$, $p = .14$ (Greenhouse-Geisser), partial $\eta^2 = 0.021$. The Spanish CCT scores in the two groups were matched at the beginning and at the end of the Intervention, and, as predicted, increased in both groups equally over the course of the 18-week period.

Measures of English Speech Production (LENA)

At the end of the 18-week period, the Intervention group produced an average of 74.1 English vocalizations per child per hour, a significant gain from the beginning of the Intervention when the average number was just over 12 English vocalizations per hour, $t(125) = -14.87$, $p < .001$, $d = -1.33$, CI (−76.1, −47.2) (Figure 4). At the end of the Intervention, the CPC group produced just over 13 English vocalizations per child per hour, a number equivalent to the average produced by the Intervention group at the beginning of the 18-week period, $t(248) = -0.95$, $p = .34$, $d = -0.12$, CI (−7.12, 5.13).

The same pattern of results was found across all five categories of English vocalizations: English words, English phrases, repeated English vocalizations, spontaneous English vocalizations, and mixed phrases (Figure S1a–e, Supporting Information). For each of the five categories, *t*-tests showed that the Intervention group produced significantly more English vocalizations per child per hour at the end of the Intervention compared to the beginning of the Intervention ($ps < .05$), and the CPC group produced significantly fewer English vocalizations than the Intervention group at the end of the Intervention ($ps < .05$; for details see Figure S1a–e, Supporting Information). Multivariate analysis of variance (MANOVA) statistics are shown in the in Results section in Appendix S1, Supporting Information.

Age effects were assessed on the total number of English vocalizations produced per child per hour. For Intervention children, a repeated measures ANOVA showed a significant interaction of Age (four levels) with Time (start/end of Intervention), Wilks's $\lambda = 0.91$, $F(3, 122) = 3.7$, $p = .013$, partial $\eta^2 = .08$. A Bonferroni post hoc test revealed that all four age groups were different from each other at $p < .05$. In a comparison of Intervention and CPC groups (Group),

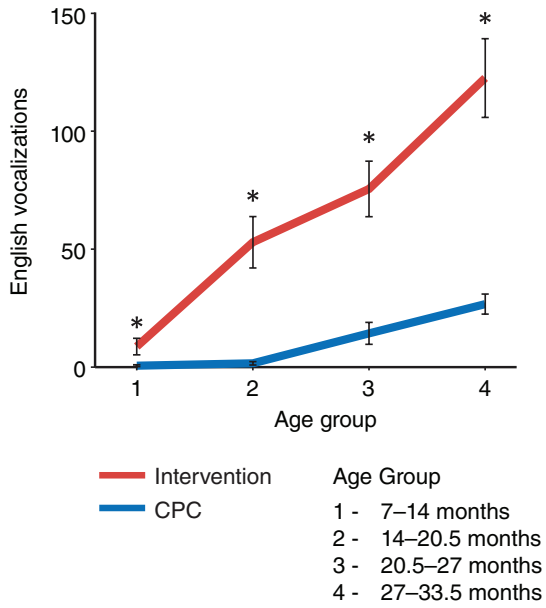

Fig. 5. Mean number of English vocalizations per child per hour at the end of the 18-week period, for the Intervention (red) and Current Practice Comparison (CPC) (blue) children, by age group. All * indicate that the average per child per hour is significantly higher in the Intervention compared to the CPC group ($p < .05$).

a univariate ANOVA showed a main effect of Age, $F(3, 242) = 67.2$, $p < .001$, partial $\eta^2 = .4$, a main effect of Group, $F(1, 242) = 169.0$, $p < .001$, partial $\eta^2 = .4$, and an interaction between Age and Group, $F(3, 242) = 5.3$, $p = .001$, partial $\eta^2 = .06$. Bonferroni post hoc tests revealed that the CPC children performed significantly below the Intervention children at the end of the Intervention at all four ages ($p < .05$; see Figure 5).

Significant gains were also observed in English language complexity, as measured by an increase in the mean length of English phrases (Figure S1f, Supporting Information.)[1] In the Intervention group, English phrases increased from 2.2 words in length at the beginning of the Intervention to 2.8 words in length by the end of the Intervention, $t(794) = 5.22$, $p < .001$, $d = 0.50$, CI (−0.97, −0.23). The mean English phrase length in the CPC group at the end of the Intervention was 2.2, equivalent to that of the Intervention group at the beginning of the Intervention, $t(219) = 0.0$, $p = 1.0$, $d = 0$, CI (0, 0).

Effects of Neighborhood Wealth on English Comprehension and Production

Two separate two-way ANOVAs were conducted to assess the role of neighborhood wealth on learning, one for the English CCT scores at the end of the Intervention (comprehension) and one for the English vocalizations produced at the end of the Intervention (production). The two effects tested were Wealth (Low/Mid) and Group (Intervention/CPC). The ANOVA for CCT scores yielded a significant main effect of Group, $F(1, 100) = 22.67$, $p < .001$, partial $\eta^2 = .19$. The main effect of Wealth was not significant, $F(1, 100) = 0.25$, $p = .62$, partial $\eta^2 = .003$, nor was the interaction between Group and Wealth, $F(1, 100) = 1.79$, $p = .18$, partial $\eta^2 = .018$. The analyses of English vocalizations revealed very similar results: There was a significant main effect of Group, $F(1, 246) = 95.76$, $p < .001$, partial $\eta^2 = .28$, but not Wealth, $F(1, 246) = 0.003$, $p = .96$, partial $\eta^2 = 0$, and the interaction between Group and Wealth was not significant, $F(1, 246) = 1.64$, $p = .20$, partial $\eta^2 = .007$. This indicates that on measures of English comprehension and production, children in low-income neighborhood schools learned at the same rate as children in mid-income schools.

Retention Test: 18 Weeks Post-Intervention

We were interested in the degree to which children retained the English vocabulary they learned, and this was measured by readministering the CCT (comprehension) to Intervention children 18 weeks after the completion of the study. We also retested the Spanish CCT. The average scores in the follow-up testing were 23.25 (standard deviation [*SD*] = 4.5) for the English CCT and 36.60 (*SD* = 4.6) for the Spanish CCT. Two separate repeated measures ANOVAs were conducted, one for each language, with Time as the 3-level independent variable (start of Intervention = Time 1, end of Intervention = Time 2, 18 weeks post-Intervention = Time 3). For English and for Spanish, there was a significant Time effect (English: Wilks's $\lambda = 0.37$, $F(2, 50) = 42.9$, $p < .001$, partial $\eta^2 = .63$; Spanish: Wilks's $\lambda = 0.40$, $F(2, 50) = 36.9$, $p < .001$, partial $\eta^2 = .60$). Bonferroni post hoc tests revealed that for English, Time 1 was significantly different from Time 2 ($p < .001$) and Time 3 ($p < .001$), but there was no difference between Time 2 and Time 3 ($p > .05$). For Spanish, Time 1 was significantly different from Time 2 ($p < .001$) and Time 3 ($p < .001$), and Time 2 was also significantly different from Time 3 ($p < .001$). This shows that the effects of English learning were retained 18 weeks after the completion of the Intervention, and that Spanish continued to increase during this time (Figure 6).

DISCUSSION

The current study demonstrates that 18 weeks of foreign language exposure in daily, social, 1-hour long play sessions with native-speaking tutors using a research-based method leads to highly significant gains in foreign language comprehension and production in children between 7 and 33.5 months of age. The learning effects generalize across five categories of English vocalizations, and across the entire age

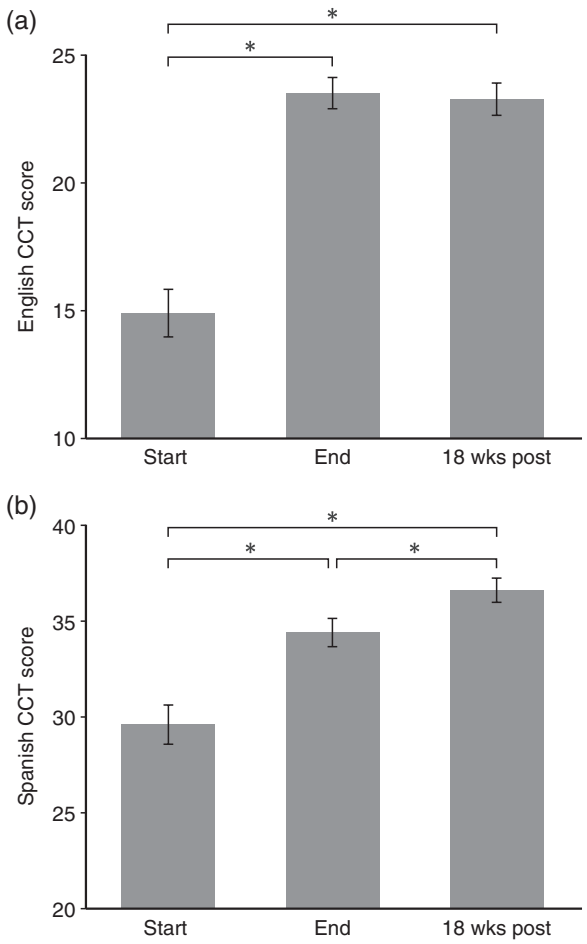

Fig. 6. Mean English (a) and Spanish (b) Computerized Comprehension Task (CCT) scores for the Intervention group at the beginning of the 18-week period, at the end of the 18-week period, and 18-weeks post-Intervention ($n = 52$). * Indicates significance at ($p < .001$).

range tested. While older children produced a greater number of English vocalizations compared to younger children, even children who were between 7 and 14 months old at the beginning of the Intervention showed significant English production gains. Retention measures show that the gains in English learning remain unchanged for at least 18 weeks after social exposure ends. Importantly, children in the Intervention group simultaneously made significant progress in Spanish, advancing in Spanish at the same rate as the children in the CPC group.

Parents in Madrid, in the United States, and around the world have a common concern when it comes to bilingual education—many do not speak foreign languages, and cannot afford expensive private programs designed to teach them. Public schools typically do not offer foreign language instruction until elementary school, therefore missing the brain's critical window for native-like language acquisition. The language Intervention we tested here was conducted in four public Infant Education Centers serving low- to mid-income families and showed significant comprehension and production gains regardless of neighborhood wealth. These findings are interesting considering the wealth of literature showing the negative effects of lower socioeconomic status (SES) on language learning and processing (Fernald, Marchman, & Weisleder, 2013; Ramey & Ramey, 2004). Our results support the notion that all infants, regardless of socioeconomic background, can readily acquire a second language in a social environment that engages children with high quantity and quality language input (McCabe et al., 2013; Werker & Byers-Heinlein, 2008).

Importantly, our findings also underscore the fact that the nature and quality of foreign language programs for infants and toddlers play an important role in learning. The Intervention tested here was based on six learning principles, each of which is supported by theory and research. From a scientific perspective, we cannot yet specify the minimal conditions that are necessary and sufficient for producing the observed gains. For example, the English speech production gains in the Intervention group may have been partly due to increasing the children's general level of engagement. It is also possible and likely that learning is aided by increasing the amount of language exposure above Madrid's standard bilingual teaching program, lowering the child to teacher ratio, or employing highly motivated language tutors (Burchinal & Cryer, 2004; Yoshikawa et al., 2013).

Further laboratory and intervention studies will be needed to test specific subsets of variables and their effects on foreign language learning. Brain studies using infant- and toddler-friendly techniques, such as MEG brain recordings, will play a critical role in improving our understanding of the neural basis of variation in second language acquisition outcomes. For example, examining MEG responses during social and nonsocial language experiences, a study now underway will allow us to test whether the brain is activated differentially during these two conditions. MEG studies comparing infants and toddlers with preschoolers can shed light on the differential involvement of brain networks in simultaneous versus sequential second language learning in early childhood. Follow-up intervention studies spanning longer periods of time (an entire school year or longer) are necessary to assess children's syntactic growth. Due to the relatively short duration of the Intervention (18 weeks), the current study focused primarily on the early stages of language learning, and did not go beyond a rudimentary assessment of growth in syntactic complexity at the group level. Further studies are needed to assess children's grammatical growth, and to determine whether research-based interventions in early education settings can in fact create truly bilingual minds.

Follow-up studies will also be needed to evaluate the necessary conditions under which children retain second

language knowledge. In the present study, we showed that infants retained their knowledge of English for 18 weeks post-Intervention. During the Intervention and post-Intervention periods, we were careful to ensure that children did *not* receive additional English exposure outside of the school. Future studies can explore avenues to provide additional support for learning and retention. For example, family engagement has been linked to multiple positive child outcomes (see NASEM, 2017, chap. 5), and studies should consider ways in which parents can support learning at home.

Taken together, the current study shows that a rigorous and well-described multifactor intervention rapidly serves to advance second language learning in infants and young children. Our findings underscore an important point about the human ability to acquire more than one language: Children have the capacity to begin second language learning in early childhood and they make rapid gains from social language exposure when the method of language exposure incorporates key evidence-based features that have been demonstrated to be effective in laboratory research. Here we demonstrate the effectiveness of the method in a school-based program lasting only 1 hour/day. Translational science that takes basic research into the classroom to create well-designed experimental interventions—in the present case an intervention for infants and toddlers on early bilingual language learning—has an impact that has the potential to improve second language learning worldwide.

Acknowledgments—The authors thank Denise Padden and Rechele Brooks of the University of Washington Institute for Learning & Brain Sciences (UW I-LABS), and Margaret Friend, Alicia Muñoz Sánchez, and Manuel Rubio Sánchez for valuable assistance throughout data collection, manuscript preparation, and review. The research was supported by the Madrid Regional Ministry of Education, Youth, and Sports, and the UW I-LABS Ready Mind Project.

NOTE

1 Phrase length was calculated by taking a mean of all phrases produced by all children, and is thus not comparable to standardized measures of grammatical development (such as mean length of utterance [MLU]).

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix S1. Supplementary materials and results with references list.

Table S1. Adult words per hour, child vocalizations per hour, and adult–child conversational exchanges per hour in Intervention and Current Practice Comparison (CPC) children.

Figure S1. Mean number of English vocalizations per child per hour, by category, for Intervention Group (red; $n = 126$) at the beginning and end of the 18-week period, and for Current Practice Comparison (CPC) group (blue; $n = 124$) at the end of the 18-week period. Repeated English vocalizations (a), spontaneous English vocalizations (b), English words (c), English phrases (d), and mixed phrases (e). Average phrase length across all English phrases produced by all participating children (f). All * indicate that the mean per child per hour is significantly greater in the Intervention group compared to the CPC control group ($p < .05$).

REFERENCES

Ayuntamiento de Madrid, Área de Gobierno de Economía y Empleo. (2008). *Numero 16 del Barómetro de Economía de la Ciudad de Madrid*. Madrid, Spain: Author.

Barac, R., & Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child Development*, *83*, 413–422.

Berken, J. A., Chai, X., Chen, J. K., Gracco, V. L., & Klein, D. (2016). Effects of early and late bilingualism on resting-state functional connectivity. *Journal of Neuroscience*, *36*, 1165–1172.

Bialystok, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology*, *65*, 229–235.

Bornstein, M. H., Tamis-LeMonda, C. S., Hahn, C.-S., & Haynes, M. (2008). Maternal responsiveness to young children at three ages: Longitudinal analysis of a multidimensional, modular, and specific parenting construct. *Developmental Psychology*, *44*, 867–874.

Bruderer, A. G., Danielson, D. K., Kandhadai, P., & Werker, J. F. (2015). Sensorimotor influences on speech perception in infancy. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 13531–13536.

Burchinal, M. R., & Cryer, D. (2004). Diversity, child care quality, and developmental outcomes. *Early Childhood Research Quarterly*, *18*, 401–426.

Buysse, V., Peisner-Feinberg, E., Paez, M., Hammer, C., & 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.

Callahan, R., & Gándara, P. (2014). *The bilingual advantage: Language, literacy, and the labor market*. Bristol, England: Multilingual Matters.

Carlson, S., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, *11*, 282–298.

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, *38*, 967–978.

Conboy, B. T., Brooks, R., Meltzoff, A. N., & Kuhl, P. K. (2015). Social interaction in infants' learning of second-language phonetics: An exploration of brain–behavior relations. *Developmental Neuropsychology*, *40*, 216–229.

Conboy, B. T., & Kuhl, P. K. (2011). Impact of second-language experience in infancy: Brain measures of first- and second-language speech perception. *Developmental Science*, *14*, 242–248.

Craik, F., Bialystok, E., & Freedman, M. (2010). Delaying the onset of Alzheimer disease: Bilingualism as a form of cognitive reserve. *Neurology*, *75*, 1726–1729.

Ferguson, C. A., Menn, L., & Stoel-Gammon, C. (Eds.) (1992). *Phonological development: Models, research, implications*. Timonium, MD: York Press.

Ferjan Ramirez, N., Ramirez, R. R., Clarke, M., Taulu, S., & Kuhl, P. K. (2017). Speech discrimination in 11-month-old bilingual and monolingual infants: A magnetoencephalography study. *Developmental Science*, *20*, e12427.

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skills are evident at 18 months. *Developmental Science*, *16*, 234–48.

Fish, M. S., Garcia-Sierra, A., Ramirez-Esparza, N., & Kuhl, P. K. (2017). Nonnative language input in infancy: A comparative study of monolingual and bilingual caregiver speech. *Journal of Phonetics*, *63*, 19–34.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 233–277). Timonium, MD: York Press.

Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America*, *106*, 2973–2987.

Friend, M., & Keplinger, M. (2003). An infant-based assessment of early lexicon acquisition. *Behavior Research Methods, Instruments, and Computers*, *35*, 302–309.

Friend, M., & Keplinger, M. (2008). Reliability and validity of the computerized comprehension task (CCT): Data from American English and Mexican Spanish infants. *Journal of Child Language*, *35*, 77–98.

Garcia-Sierra, A., Ramirez-Esparza, N., & Kuhl, P. K. (2016). Relationships between quantity of language input and brain responses in bilingual and monolingual infants. *International Journal of Psychophysics*, *110*, 1–17.

Han, M., Moore, N., Vukelich, C., & Buell, M. (2010). Does play make a difference? How play intervention affects the vocabulary learning of at-risk preschoolers. *American Journal of Play*, *3*, 82–104.

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

Hay, J. F., Pelucchi, B., Graf-Estes, K., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, *63*, 93–106.

Hermanto, N., Moreno, S., & Bialystok, E. (2012). Linguistic and metalinguistic outcomes of intense immersion education: How bilingual? *International Journal of Bilingual Education and Bilingualism*, *15*, 131–145.

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*, 1–27.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, *87*, B47–B57.

Jacobs, E., Miller, L., & Tirella, L. (2010). Development and behavioral performance of internationally adopted preschoolers: A pilot study. *Child Psychiatry and Human Development*, *41*, 15–29.

Jasinska, K. K., & Petitto, L. A. (2013). How age of bilingual exposure can change the neural systems for language in the developing brain: A functional near infrared spectroscopy investigation of syntactic processing in monolingual and bilingual children. *Developmental Cognitive Neuroscience*, *6*, 87–101.

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

Klein, D., Mok, K., Chen, J. K., & Watkins, K. E. (2014). Age of language learning shapes brain structure: A cortical thickness study of bilingual and monolingual individuals. *Brain and Language*, *131*, 20–24.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*, 831–843.

Kuhl, P. K. (2007). Is speech learning gated by the social brain? *Developmental Science*, *10*, 110–120.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B*, *363*, 979–1000.

Kuhl, P. K., Ramirez, R. R., Bosseler, A., Lin, J.-F. L., & Imada, T. (2014). Infants' brain responses to speech suggest analysis by synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 11238–11245.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*, F13–F21.

Kuhl, P. K., Tsao, F. -M., & Liu, H. -M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 9096–9101.

LENA Research Foundation. (2015). *Software version V3-4.0*. Retrieved from http://www.lenafoundation.org

Lenneberg, E. H. (1967). *Biological foundations of language*. New York, NY: John Wiley & Sons.

López Ornat, S., Gallego, S., Gallo, P., Karousou, A., Mariscal, S., & Martínez, M. (2005). *MacArthur: Inventario de desarrollo comunicativo. Manual y Cuadernillos*. Madrid, Spain: TEA Ediciones.

MacKay, I. R. A., & Flege, J. E. (2004). Effects of the age of second language learning on the duration of first and second language sentences: The role of suppression. *Applied Psycholinguistics*, *25*, 373–396.

McCabe, A., Tamis-LeMonda, C. S., Bornstein, M. H., Cates, C. B., Golinkoff, R., Hirsh-Pasek, K., … Guerra, A. W. (2013). Multilingual children: Beyond myths and towards best practices. *Social Policy Report, Society for Research in Child Development*, *27*, 1–37.

Meacham, A. (2006). Language learning and the internationally adopted child. *Early Childhood Education Journal*, *34*, 73–79.

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Neurolinguistics: Structural plasticity in the bilingual brain. *Nature*, *431*, 757.

Muñoz, C. (2006). Age and the rate of foreign language learning. *Applied Linguistics*, *28*, 627–630.

National Academy of Sciences, Engineering, and Medicine (NASEM). (2017). *Promoting the educational success of children and youth learning English: Promising futures* (pp. e24677). Washington, DC: National Academies Press.

Petitto, L. A., Katerelos, M., Levy, B., Gauna, K., Tetreault, K., & Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: Implications for the mechanisms underlying early bilingual language acquisition. *Journal of Child Language*, *28*, 453–96.

Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence 2-year-olds' bilingual proficiency. *Child Development*, *82*, 1834–1849.

Ramey, C. T., & Ramey, S. L. (2004). Early learning and school readiness: Can early intervention make a difference? *Merrill-Palmer Quarterly*, *50*, 471–491.

Ramirez-Esparza, N., Garcia-Sierra, A., & Kuhl, P. K. (2014). Look who's talking: Speech style and social context in language input are linked to concurrent and future speech development. *Developmental Science*, *17*, 880–891.

Ramirez-Esparza, N., Garcia-Sierra, A., & Kuhl, P. K. (2016). The impact of early social interactions on later language development in Spanish-English bilingual infants. Advance online publication. *Child Development*, e12648.

Roseberry, S., Hirsh-Pasek, K., Parish-Morris, J., & Golinkoff, R. M. (2009). Live action: Can young children learn verbs from video? *Child Development*, *80*, 1360–1375.

Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, *12*, 339–49.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Seidl, A., Onishi, K. H., & Cristia, A. (2014). Talker variation aids young infants' phonotactic learning. *Language Learning and Development*, *10*, 297–307.

Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., & Perani, D. (2003). Early setting of the grammatical processing in the bilingual brain. *Neuron*, *37*, 159–170.

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, *8*, 231–256.

Weber-Fox, C. M., & Neville, H. J. (2001). Sensitive periods differentiate processing of open and closed-class words: An ERP study of bilinguals. *Journal of Speech, Language, and Hearing Research*, *44*, 1338–1353.

Weisberg, D. S., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Guided play: Where curricular goals meet a playful pedagogy. *Mind, Brain, and Education*, *7*, 104–112.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*, 2143–2152.

Werker, J. F., & Byers-Heinlein, K. (2008). Bilingualism: First steps in perception and comprehension. *Trends in Cognitive Science*, *12*, 144–151.

Werker, J. F., & Tees, R. C. (1984). Cross language speech perception: Evidence for perceptual reorganization. *Infant Behavior and Development*, *7*, 49–63.

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinoza, L. M., Gormley, W. T., … Zaslow, M. J. (2013). *Investing in our future: The evidence base on preschool education*. New York, NY: Foundation for Child Development.

Zhang, Y., Koerner, T., Miller, S., Grice-Patil, Z., Svec, A., Akbari, D., … Carney, E. (2011). Neural coding of formant-exaggerated speech in the infant brain. *Developmental Science*, *14*, 566–581.

Zhang, Y., Kuhl, P. K., Imada, T., Iverson, P., Pruitt, J., Stevens, E. B., … Nemoto, I. (2009). Neural signatures of phonetic learning in adulthood: A magnetoencephalography study. *NeuroImage*, *46*, 226–240.