HHS Public Access

Author manuscript

Dev Sci. Author manuscript; available in PMC 2018 January 01.

Published in final edited form as:

Dev Sci. 2017 January ; 20(1): . doi:10.1111/desc.12425.

Caregiver talk to young Spanish-English bilinguals: Comparing direct observation and parent-report measures of dual-language exposure

Virginia A. Marchman[1], **Lucía Z. Martínez**[1], **Nereyda Hurtado**[1], **Theres Grüter**[2], and **Anne Fernald**[1]

[1]Stanford University

[2]University of Hawai'i at Manoa

Abstract

In research on language development by bilingual children, the early language environment is commonly characterized in terms of the relative amount of exposure a child gets to each language based on parent report. Little is known about how absolute measures of child-directed speech in two languages relate to language growth. In this study of 3-year-old Spanish-English bilinguals (*n* = 18), traditional parent-report estimates of exposure were compared to measures of the number of Spanish and English words children heard during naturalistic audio recordings. While the two estimates were moderately correlated, observed numbers of child-directed words were more consistently predictive of children's processing speed and standardized test performance, even when controlling for reported proportion of exposure. These findings highlight the importance of caregiver engagement in bilingual children's language outcomes in both of the languages they are learning.

The linguistic richness of a child's early learning experience is critical for language acquisition and cognitive growth, whether that child is learning one language or two. For monolingual English-speaking children, it is well-known that the quantity and quality of talk children hear from caregivers at home predicts later vocabulary and IQ (Hart & Risley, 1995; Hoff, 2003; Huttenlocher, Haight, Bryk, Seltzer & Lyons, 1991). For monolingual Spanish-speaking children as well, hearing more talk from caregivers predicts both children's vocabulary development and their skill in processing spoken language in real time (Hurtado, Marchman & Fernald, 2008; Weisleder & Fernald, 2013). For bilingual children learning two languages, the link between early language experience and language learning is also a critical issue (Song, Tamis-LeMonda, Yoshikawa, Kahana-Kalman & Wu, 2011). The linguistic environment of children hearing two languages has typically been characterized in terms of estimates of the proportion of a child's exposure to each language based on parent report, rather than by observing the absolute amounts of speech directed to the child in each language (Pearson, Fernández, Lewedeg & Oller, 1997; Hoff et al., 2012). In this study of 3-year-old Spanish-English bilinguals, we compare the predictive validity of reported and

Address for correspondence: Virginia A. Marchman, Ph.D., Department of Psychology, Jordan Hall, Building 420, Stanford University, Stanford, CA 94305, marchman@stanford.edu.

observed measures of language input in each language in accounting for children's language proficiency in Spanish and English.

From reported to observed measures of language exposure in bilinguals

Bilingual children's language input is necessarily distributed over two languages. One widely used measure of a child's relative experience with each language is the proportion of waking hours during which the child is reported to be exposed to one language versus the other, derived from detailed parent interviews (Marchman & Martínez-Sussmann, 2002; Marchman et al., 2010; Gutiérrez-Clellen & Kreiter, 2003) or from diary records (Place & Hoff, 2011). Using such measures, Pearson et al. (1997) found that the proportion of reported exposure to Spanish versus English predicted the number of words that bilingual infants and toddlers could produce in one language compared to the other, a result that has been replicated in other bilingual populations as well (e.g., de Houwer, 2007; Gathercole & Thomas, 2009; Patterson, 1998; Hoff et al., 2012). Some studies suggest that relative exposure also predicts grammar (Elin Thordardottir, 2011; Elin Thordardottir, Rothenberg, Rivard & Naves, 2006; Gutiérrez-Clellen & Kreiter, 2003; Hoff et al., 2012; Marchman, Martínez-Sussmann & Dale, 2004), although others argue that syntactic development is less influenced by relative exposure than is vocabulary (e.g., Paradis & Genesee, 1996; Unsworth, 2013). Hurtado, Grüter, Marchman and Fernald (2013) found that children's relative exposure to Spanish vs. English was also correlated with their relative efficiency in processing Spanish vs. English words, further evidence that children's early proficiency in lexical comprehension is tied to their experience with a particular language.

These studies using proportion of exposure to two languages have addressed important questions about bilingual development: For example, is the balance between two languages heard by the child reflected in the balance between those two languages in the child's emerging vocabulary (e.g., Pearson, et al., 1997)? However, it is important to note that this question is quite different from the central question that has motivated research exploring effects of early language experience in monolingual children (e.g., Hart & Risley, 1995), which ask: Do those children who hear more child-directed speech overall from caregivers make more rapid gains in vocabulary learning and other language skills than do children whose parents are less verbally engaged?

This latter question, framed in terms of quantity of child-directed speech, is just as relevant to bilingual children as it is to monolingual children. Some children are balanced in their exposure, in that they hear 50% Spanish and 50% English according to their parents' reports. Previous findings suggest that although these children might differ in the absolute size of their vocabularies, they would be more or less balanced between Spanish and English (Pearson et al., 1997). However, such proportional measures tell only half the story, since there are large individual differences across families in how much talk is addressed to young children over the course of a day. For example, in a study examining speech to toddlers in monolingual Spanish speaking families during a typical 10-hour day, Weisleder and Fernald (2013) found that total amount of child-directed speech ranged from 12,000 words/day to 700 words/day. Comparing bilingual to monolingual families, De Houwer (2014) showed that Dutch-speaking mothers also varied tremendously in the amount of talk directed to their

child during home-based video-recorded sessions. Critically, based on 13 different indices of input frequency, the extent of the variation among families in quantity of caregiver talk was comparable for monolingual and bilingual families.

Thus, for bilingual children, both the overall balance of exposure to each language and the absolute amount of speech provided by caregivers will be sources of variation in the total number of Spanish and English words available in the environment (Grüter, Hurtado, Marchman & Fernald, 2014). However, bilingual children who have the same overall balance of exposure may vary in terms of the amounts of child-directed speech that they hear in each of their languages. This variation in quantity of speech could have a powerful influence on lexical learning, resulting in large differences in vocabulary size in the two languages resulting from factors beyond exposure balance, similar to effects seen in monolinguals (Hart and Risley, 1995; Weisleder & Fernald, 2013).

Very little is known about how amount of exposure to child-directed speech in each language is related to bilingual language growth. A major reason for the lack of previous research on this question is that it is much more difficult for researchers to obtain a representative sample of the language environment of a bilingual child than a monolingual child (Place & Hoff, 2011). In their landmark study, Hart and Risley (1995) showed that 1-hour recordings of primary caregivers interacting with their young children at home yielded stable measures of children's language environment from month to month. In contrast, comparable home recordings in bilingual families are much less likely to be representative. If a mother speaks Spanish to her child several hours each day, the grandparents speak Spanish to the child on weekends, and the child hears English from siblings in the afternoon and from the father in the evening, a 1-hour recording session cannot possibly capture this diversity. Thus researchers often rely on parent-report interviews or diaries to estimate the proportions of time bilingual children are engaged with speakers of each language, a reasonable compromise given that direct observations of sufficiently large samples of children's bilingual learning environments are much more demanding in terms of time, transcription and coding.

Given these logistical challenges, few studies have examined how individual differences in the quantity of caregiver talk matter for outcomes in young bilinguals. De Houwer (2011) first explored this question indirectly in a reanalysis of 3-hour home recordings of five child-caregiver pairs speaking English and Inuktitut (Allen, 2007). By extrapolating absolute frequencies from the proportions of utterances in each language, De Houwer reported that children's productions in each language were more closely aligned with the absolute counts of caregiver's utterances than with the relative proportion of utterances reported in each language. In this study we conduct a direct test of this hypothesis. We first examine links between reported language proportion and child outcomes. We then evaluate the associations between the observed measures of the number of words bilingual children heard in each language based on naturalistic recordings of the language directed to children during interactions with others. To the extent that variation in the amount of words children hear adds important information about bilingual children's learning environments, observed measures of talk in each language should predict children's proficiency in each language, even when controlling for estimated exposure based on parent report.

New approaches to exploring the language environment in bilingual families

To assess the amounts of child-directed speech each child was exposed to in Spanish and English, we took advantage of the LENA™ recording and speech analysis system (Ford, Baer, Xu, Yapanel & Gray, 2009). This powerful new technology makes use of a small digital language processor (DLP) to make recordings up to 16 hours in duration. The DLP is worn by the child in the front pocket of specially-designed clothing, so recordings are made from the child's perspective. Since no outside observer is present, the LENA™ system yields more extensive recordings of children's daily interactions that are also more representative and less intrusive than is possible using other methods. Analysis of these recordings relies on specialized speech-recognition algorithms to estimate all clear adult speech spoken in proximity to the target child. Although the accuracy of adult word counts (AWC) based on the LENA™ automated measures was first established for English (Xu, Yapanel & Gray, 2009; Oller et al., 2010; Oetting, Hartfield & Pruitt, 2009), this technology has been used effectively with other languages as well. In a study of mother-child interaction in Latino families, Weisleder & Fernald (2013) validated the use of LENA™ with Spanish child-directed speech, showing a high correlation between the automated AWC estimates per hour and the numbers of child-directed words/hour derived from word-for-word transcripts of the same speech samples.

It is important to note that the LENA™ technology is not yet able to provide automated measures that distinguish between Spanish and English. Thus a native Spanish-speaking coder listened to the recordings to identify whether each 5-min segment of child-directed speech was in Spanish, English, or both. By combining these codes with the LENA™ automated measures, we were able to estimate the absolute number of Spanish and English child-directed words that children actually hear over many hours of natural interactions with others.

Assessing bilingual children's language proficiency in two languages

We focused on emerging bilingual children who were 3 years old, an age at which we were able to use school-relevant assessments of language outcomes, including vocabulary and grammatical skills, in both Spanish and English. All children were administered the Spanish and English versions of the Clinical Evaluation of Language Fundamentals, Preschool (CELF-P; Wiig, Secord, & Semel, 2009; Semel, Wiig, & Secord, 2004), a standardized test of receptive and expressive language knowledge.

All children were also tested in both languages in the Looking-while-Listening (LWL) task, to assess their speed in interpreting familiar words in Spanish and English. The LWL task taps into young children's developing efficiency in spoken language processing (Fernald, Zangl, Portillo & Marchman, 2008). Children look at two pictures on a screen while listening to speech naming one of the objects (e.g., *Where's the birdy?* or *¿Dónde está el pájaro?*). Children's eye movements as they shift to the named target picture reveal the speed with which they can interpret familiar object names in real time. Research with monolingual English-speaking children has shown that infants become faster and more accurate in

identifying referents of familiar words presented in continuous speech over the second year of life (Fernald, Pinto, Swingley, Weinberg & McRoberts, 1998). Moreover, greater efficiency in early language processing is associated with more accelerated vocabulary growth and better long-term language and cognitive outcomes (Fernald, Perfors & Marchman, 2006; Marchman & Fernald, 2008). Studies with Spanish-speaking children from Latino families in the U.S. report similar findings (Hurtado, Marchman & Fernald, 2007). The predictive validity of reaction time (RT) in the LWL task has been further demonstrated with English-speaking late talkers, for whom efficiency in language processing at 18 months was a significant predictor of risk status at 30 months (Fernald & Marchman, 2012). Such findings indicate that early efficiency in real-time processing plays a critical role in supporting language development for monolingual children from a variety of backgrounds.

In the first study to extend this paradigm to bilingual children, Marchman et al. (2010) investigated links between vocabulary size and processing efficiency in 30-month-old Spanish-English simultaneous bilinguals. The efficiency with which children processed spoken language in Spanish was related to their vocabulary size in Spanish, and the efficiency with which they processed words in English was related to their vocabulary size in English, even after controlling for children's relative exposure to Spanish versus English. These results indicated that efficiency of real-time comprehension of Spanish or English words is tied to a child's particular level of skill in a particular language, a striking parallel to what has been observed previously with monolingual children. Thus the ability to process spoken words efficiently was closely linked with a particular body of vocabulary knowledge, regardless of whether the children were learning one language or two. However, the Marchman et al. (2010) study did not specifically explore the relation between language exposure and children's language learning outcomes.

Research questions

The participants were Latino preschoolers whose home language was Spanish but who were hearing different amounts of English from various sources including parents and other family members. Audio-recordings of daily interactions enabled us to obtain estimates of the absolute frequency with which these emergent bilingual children heard words in Spanish and English directed to them, frequencies that typically vary across the different speakers in the children's lives. In language background interviews, parents reported how much Spanish and English their child heard in a typical week, estimates that were converted to a proportion of waking hours in which children were in contact with speakers of each language. To assess children's language skills in both languages, we administered parallel Spanish and English versions of a comprehensive standardized test of language knowledge. We also measured their language processing efficiency in Spanish and English using the LWL procedure.

These diverse measures of children's language environment and their language skills in both Spanish and English enabled us to address two main questions: First, how does variability in the observed number of Spanish and English words directed to children, based on the naturalistic recordings, align with variability in estimated language exposure, based on parent report? And second, to what extent do reported and observed measures of the

language experiences of these emergent bilingual children in Spanish and English predict children's outcomes? Does individual variation in quantity of caregiver talk predict children's language proficiency in each of their languages, even after controlling for reported relative exposure?

Method

Participants

Participants were 18 typically-developing Spanish-English bilingual children (13 female, 5 male), ranging in age from 36-40 months (*M* = 37.8 mos, *SD* = 1.4) who were participating in a study of language development in emergent bilingual children. All children were born in the United States, while 100% of the parents were born in Mexico and were native speakers of Spanish. Although Spanish was the predominant home language in all families, these children were also regularly exposed to English, from parents or other caregivers, a situation that is typical of many emergent bilingual children in the U.S. Mothers reported less than a high school education on average (*M* = 10.1 years, range = 5 to 16), with 61% (11 of 19) completing 9 years of schooling or less. Only 4 (22%) reported attending at least some college. Socioeconomic status (SES) was estimated with an updated version of the Hollingshead Four Factor Index of Social Status (Hollingshead, 1975), a composite scale based on both parents' education and occupation (possible scores = 8 to 66). The mean HI score was 23.8 (*SD* = 11.2, range = 13 to 53) in this predominately lower-SES sample. Since there was some range in age at test and HI scores in this sample, these factors are included as control variables in all analyses in order to explore relations between reported and observed exposure and child outcomes after these factors were taken into account.

Procedure and measures

This study was conducted in a community-based laboratory located about 10 miles away from a major university campus. All research staff were bilingual/bicultural and communicated with families in either Spanish or English, depending on the preferences of the family. Families visited the laboratory twice, to assess children's skill in Spanish and English on separate visits one week apart. The language used during the first and second visit was counterbalanced.

Reported measures of Spanish and English exposure—At the first testing session, an experimenter conducted a comprehensive language background environment interview with one or both parents, typically the mother (Marchman & Martínez-Sussmann, 2002). The parent was first asked to describe their child's typical weekday and weekend, including wake-up, night-time, and nap times. Next, the parent listed those people with whom their child came into regular contact, when that contact occurred, and the proportion of Spanish vs. English they used with the child. For each person in the child's life, the total number of hours that person spent with the child, and then the number of hours using Spanish and English, was computed. Total person hours was the total number of contact hours summed across all people who regularly interacted with the child (e.g., mother, father, grandmother). *Reported Spanish Proportion* was then defined as the proportion of total person hours in Spanish out of total person hours in Spanish + English. Since the Spanish to English

proportion scores were not normally distributed, nonparametric correlations (Spearman's rho) are reported in all analyses.

Measures of Spanish and English exposure based on word counts—Audio recordings of each child's interactions with caregivers and others who engaged with the child were made using the LENA™ recording and speech analysis system. Parents were asked to record a typical weekday following their regular routine, turning on the DLP recorder at the beginning of the day and preferably leaving it on for a minimum of 8 hours. Because the LENA™ analysis software is less accurate in noisy environments (Xu et al., 2009), families were encouraged to record on a day when there were no special activities planned (e.g., parties or trips to an amusement park). Families typically recorded when they were at home; however, they frequently recorded in other locations, including outside or in other peoples' homes. The DLP could be turned off whenever the family chose to do so, or it automatically turned off after 16 hours of recording. Raw recording lengths ranged from 5.3 to 13.1 (M = 9.4) hours, with some variability in recording length due to the fact that some families turned off the recorder at different points in the day (e.g., while the child was sleeping), while other families did not. To ensure that all estimates were based only on recordings when the child was awake and when the automated analyses were likely to be accurate, trained listeners “cleaned” all of the recordings, deleting segments when the child was sleeping, the child was not wearing the recording device, or the recording quality was poor. The lengths of the cleaned recordings included in the analysis ranged from 5.3 to 12.6 hours (M = 8.6 hours). Some recordings were conducted over multiple days (M = 2.2 days). An additional 14 families were invited to participate but were excluded because they chose not to take the recorder (n = 5), they returned it with less than 4 hours of recording (n = 3), their recording had technical problems (n = 2), or they did not complete testing at the second visit (n = 4).

Automated measures of the recordings were derived using the LENA™ analysis software, including counts of all adult words spoken “near and clear” to the child (AWC) in each 5-minute recording segment. Note that the LENA™ system identifies words spoken by “adults” based on the acoustic properties of the speech signal, and includes all male and female speakers near to the target child who have a non-child-like voice register, including older children. The LENA™ system does not compute word-count estimates for those portions of each 5-minute segment that include speech produced by the target child, by other small children, by television or other electronic media, by adults who are not in close proximity to the target child, or in “overlapping” speech when two individuals speak at the same time.

Because the LENA™ software identifies words spoken near the child, but which are not necessarily spoken to the child, trained bilingual listeners coded each 5-min segment of speech to determine whether the AWC counts derived from the LENA™ automated analysis were primarily child-directed speech (CDS) spoken by an adult or older child. And because the LENA™ software does not differentiate Spanish from English word tokens, these bilingual coders also assessed what proportion of the words in each segment (in 5% increments) was in each language, ranging from all Spanish/no English (100/0%) to no Spanish/all English (0/100%). Reliability was assessed by double coding 20% of the

sessions, yielding an initial concordance within 5-percentage points of 94%. Those 5-min segments identified as being primarily child-directed were used to derive estimates of the CDS in Spanish and English for each child. For those 5-min segments of CDS in which words heard were 100% Spanish or 100% English, the AWCs were added to the Spanish or English CDS totals as appropriate. For segments containing both Spanish and English words, the AWCs were divided and added to the Spanish vs. English totals based on the allocation percentage. To control for differences in recording length, *Spanish CDS/hr* and *English CDS/hr* were computed as the sum of the child-directed word counts identified as Spanish or English, divided by the length of the recording after excluding naps and non-valid segments, and converted to a rate per hour. We also calculated *Observed Spanish Proportion* (Spanish CDS/hr out of Spanish + English CDS/hr) based on word counts to facilitate comparison with the proportion of Spanish exposure based on parent report. Note that total CDS/hr was uncorrelated with recording length (r_s = -.12, p = .66), indicating that longer recordings were not necessarily capturing denser samples of CDS than shorter recordings.

Children's language outcomes in Spanish and English

Language knowledge—Children were administered both the Spanish- and English-language versions of the Clinical Evaluation of Language Fundamentals, Preschool (CELF-P; Wiig, et al., 2009; Semel, et al., 2004) at the appropriate Spanish- and English-language testing sessions. The CELF-P is one of the few standardized language assessments designed to provide parallel comprehensive evaluations in English- and Spanish-speaking 3- to 6-year-olds from a range of ability levels. Standard scores for the Core Language sub-scales were derived.

Language processing efficiency—Children's language processing efficiency in Spanish and English was assessed using the looking-while listening (LWL) procedure, conducted in two separate but parallel testing sessions. When the parent and child were comfortable, they were directed to a testing booth where the child sat on his/her parent's lap facing a 54″ LCD television screen. On each trial, children saw two pictures of familiar objects for 2 seconds, and then a pre-recorded voice directed the child's attention to one of the pictures in Spanish or English (e.g., *¿Dónde está el carro?* or *Where's the car?*). A videocamera recorded children's gaze patterns which were later coded. Each testing session lasted about 7 minutes.

The target words were 10 nouns typically familiar to both Spanish- and English-learning children in this age range [apple, *la manzana*; balloon, *el globo*; banana, *el plátano*; bird, *el pájaro*; car, *el carro*; cookie, *la galleta*; cow, *la vaca*; frog, *la rana*; horse, *el caballo*; juice, *el jugo*; shoe, *el zapato*; spoon, *la cuchara*] (Dale & Fenson, 1996; Jackson-Maldonado et al., 2003). Visual stimuli were identical in the Spanish- and English-language stimulus sets, consisting of digitized photographs of familiar objects corresponding to each target word, matched for visual salience. Parents completed additional questionnaires indicating whether their child understood and said the stimulus words in Spanish and in English. All children were reported to understand all of the target words in both languages, although some children were reported not to produce some of the words in one or both languages.

In both Spanish and English, children were tested on 24 trials consisting of familiar sentences with the target nouns in sentence-final position. All target nouns (Spanish M = 735 ms, range = 670 - 800 ms; English M = 720, 650 - 875) were presented in yoked pairs matched in duration and syllable length. Three filler trials were interspersed (e.g., *¿Te gustan las fotos? ¡Aquí vienen más!; Do you like the pictures? Here are some more!*). A female native Spanish-English bilingual speaker recorded several tokens of each sentence in each language. Final tokens were chosen based on cross-token comparability in duration of the carrier phrase and target word.

Video-recordings of children's eye movements were coded offline using custom software. On each 33-ms frame, coders indicated whether the child was fixating the left or right picture, or was away from both pictures. Custom software determined the duration of fixations and the time of initiation of shifts between images. Trials on which the child was inattentive or not looking at the pictures during the stimulus sentence were prescreened out of all analyses. Reaction time (RT) was the mean latency to initiate a shift from the distracter to the target picture, calculated on each trial on which the child was looking at the distracter picture at the onset of the critical word and shifted within 300-1800 ms from critical word onset. Trials with RTs shorter than 300 ms or longer than 1800 ms were excluded since they were not likely to be in response to the stimuli (Fernald et al., 2008). Estimates of RT were based on M = 11.2 trials (range = 5 to 16) for Spanish and M = 7.1 trials (range = 2 – 11) for English.

Results

We first examine descriptive statistics of reported language exposure to Spanish and English. We next report the observed numbers of child-directed Spanish and English words that these emergent bilingual 3-year-olds heard during interactions with others, based on the LENA™ recordings. We then compare these two estimates of exposure, asking whether observed measures of exposure to each language align with traditional estimates based on parent report. Finally, we ask to what extent the reported and observed measures of children's language experience in Spanish and English predict processing speed and language knowledge in Spanish and English.

Reported vs. observed measures of bilingual children's exposure to Spanish and English

As shown in Table 1, parents reported that their children were exposed to more Spanish than English - about 80% Spanish (20% English) on average - although there was variability across families. Table 1 also reports the observed numbers of child-directed words (*CDS/hr*) identified as Spanish and English during the LENA™ recordings. Both estimates consistently showed that these children were exposed to more Spanish than English, however, the observed Spanish proportion was significantly higher than the reported proportion, Z = 3.7, p < .001. Nevertheless, a nonparametric test of association (Spearman's rho, r_s) revealed that these two measures were moderately, albeit only marginally, correlated, $r_s(16)$ = .46, p < .06. This suggests that proportion scores based on absolute word counts and those based on parent report were somewhat consistent in their estimates of the rank order of children's relative exposure to Spanish and English.

To further explore the relations between parents' reports of exposure and the observed numbers of words children heard, Figures 1a and 1b plot *CDS/hr* and relative reported exposure to Spanish and English, respectively. These figures reveal that there was substantial variability among families in the number of words directed toward the children in Spanish and English, variability that was not always reflected in parents' estimates of the proportion of exposure to each language. The shaded area in Figure 1a demarcates the 10 families in which parents reported that their child heard Spanish more than 80% of the time. Within this group of predominantly Spanish-speaking families, one child heard around 130 child-directed words in Spanish per hour, while another heard seven times more than that amount.

This relation is illustrated in Figure 1b for English. Here, we see less striking variability, but nevertheless, some variation among families in the numbers of English words these children were observed to hear. The majority of parents in this sample reported that their child heard less than 20% English overall, and while most of these children did indeed hear relatively few words in English, there was one notable exception. Further, in the 8 families who were reported to hear more than 20% English, one child heard as little as 20 words/hr in English, whereas another heard nearly five times that amount. In sum, reported language exposure failed to capture some of the sources of variability in children's actual language experience, as estimated by the LENA recordings.

How do different measures of exposure to Spanish and English predict bilingual children's language outcomes?

Table 2 presents descriptives of the child measures of language processing efficiency (RT) and scores on the CELF-P, a standardized test of language knowledge in Spanish and English. Children were slightly, but not reliably, faster to comprehend familiar words in Spanish than English. In contrast, language test scores were higher in Spanish than English ($Z = 3.6$, $p < .0001$). There was substantial variation in both measures of child outcomes.

Our main question is: To what extent do measures of reported and observed exposure predict individual variation in these child outcomes? To explore this issue, the first column of Table 3 presents nonparametric partial correlations (r_S) between reported Spanish/English proportion and children's processing speed (RT), controlling for SES and age at test. Associations between reported exposure proportion and RT were weak in both languages. However, reported exposure was moderately associated with scores on the CELF-P, accounting for 15 and 22% of the variance remaining after age and SES in Spanish and English, respectively.

The two right-hand columns of Table 3 report the associations between observed CDS/hr and lexical processing speed (RT) and CELF-P. The first number in the pair indicates the nonparametric correlation, partialing out SES and child age at test; the second number reflects the association further partialing out reported Spanish/English proportion. Looking first at RT, the link between observed CDS/hr in Spanish and children's RT in Spanish was strong, accounting for nearly 49% of the variance remaining after the covariates. Those children who heard more Spanish words were faster to interpret familiar Spanish words in real time, demonstrating more efficient language processing skills than those children who heard fewer words in Spanish. This relation is also illustrated in Figure 2a, which shows the

time course of shifting to the named target picture by children who heard relatively more Spanish CDS/hr ($n = 9$) vs. those who heard less Spanish CDS/hr ($n = 9$). Note that those children in the higher Spanish CDS/hr group were faster to initiate a shift to the target picture, and therefore, displayed increased looking to the target picture sooner in the sentence than those children who heard less Spanish CDS/hr. These effects are also reflected in mean RT differences between the two groups, with children who heard more Spanish CDS/hr being faster ($M = 689$, $SD = 130.7$) as a group, than children who heard less Spanish CDS/hr ($M = 836$, $SD = 138$), $t(16) = 2.3$, $p < .04$, $d = 1.2$.

Associations between observed child-directed words and RT were somewhat weaker in English than in Spanish, as indicated by the partial nonparametic correlations. However, when children were grouped based on higher ($n = 9$) or lower ($n = 9$) English CDS/hr scores, those bilingual children who heard relatively more English CDS tended to shift to the target picture more quickly than children who heard less English CDS, as shown in Figure 2b. This effect is also seen in a comparison of mean RTs scores, resulting in faster mean RTs for those children in the higher English CDS/hr ($M = 730$ ms, $SD = 128$) compared to the lower English CDS/hr group ($M = 902$, $SD = 244$), a result that was marginally significant, $t(16) = 1.9$, $p < .08$, $d = .88$. Note that cross-language associations were much weaker than the within-language ones: Spanish CDS/hr was not reliably linked to RT in English, and English CDS/hr was not reliably linked to RT in Spanish.

A similar pattern is seen in predicting to the CELF-P. As shown in Table 3, variability in the Spanish CDS/hr moderately predicted children's scores on the CELF-P in Spanish, and English CDS/hr moderately predicted scores on the English test ($r_s = .41$-$.53$), controlling for age at test and SES. While not all r_s reached standard criterion levels for significance due to our small sample size, the amount of variance accounted for, after controlling for SES and age of test, ranged from 17 to 28%. Moreover, these values remained in the moderate range even after further controlling for reported Spanish/English proportion, as shown by the second number in each pair. Note also that these values are substantially higher than the cross-language associations. Finally, those children who heard more Spanish CDS/hr had higher standard scores on the Spanish CELF-P ($M = 106.9$, $SD = 14.3$) than did children who heard less Spanish CDS/hr ($M = 94.3$, $SD = 7.7$), $t(16) = 2.3$, $p < .04$, $d = 1.09$. Analogously, those children who heard more English CDS/hr had higher scores on the English CELF-P ($M = 74.7$, $SD = 13.4$) than did children who heard less CDS/hr ($M = 65.1$, $SD = 8.9$), a large effect size although only marginally significant in this small sample, $t(16)=1.8$, $p = .10$, $d = .84$. Thus, there is evidence that individual differences in the number of Spanish and English words that bilingual children hear during interactions with others support their language learning, as measured by both in a test of real-time processing efficiency and on a standardized test of language skills.

Discussion

Three main findings emerged from this research: First, there was substantial variability among bilingual Latino families in the numbers of words spoken to 3-year-old children in Spanish and English, consistent with earlier findings on variability in the amount of speech addressed to children among monolingual Spanish-speaking Latino families (Weisleder &

Fernald, 2013). Second, measures of language exposure in bilingual families based on parents' estimates of the proportion of Spanish vs. English exposure were moderately aligned with children's actual language experience in these two languages based on naturalistic recordings. And third, the quantity of talk to 3-year-olds in Spanish and English was associated with children's outcomes in each language, including processing speed and language knowledge, even when controlling for reported Spanish/English exposure proportion.

The merits of reported vs. observed measures of bilingual language exposure

One main goal of the current study was to compare two different methods for measuring bilingual learning environments, contrasting parent-report estimates of Spanish vs. English exposure with observed measures of child-directed talk in both languages based on naturalistic recordings. As in monolingual children, the recordings picked up on important variability across children in the quantity of language they heard. We also found that traditional parent-report estimates of children's proportion of exposure to Spanish were sometimes inconsistent with estimates of exposure based on absolute word counts. For example, parents in six of the 18 families in this study estimated that their children heard ca. 50% to 75% Spanish each day, although the proportions of exposure to Spanish based on absolute word counts in naturalistic recordings were much higher, ranging from ca. 86% to 98%. However, such discrepancies are understandable given that parent-report estimates are likely to take into account children's exposure to English in activities in more places outside the home than were captured on our recordings. Moreover, when parents' are interviewed about their child's language exposure, they are typically asked about interactions with different caregivers across a whole week, while the LENA™ recordings in this study were made on just one or two days. Thus the samples of speech we used to calculate the CDS/hr may not have sampled the full range of interactions a child may have with different speakers of Spanish and English over a longer period of time. Despite these inconsistencies, the two different estimates of proportion of language exposure - those based on parent-report interviews and those based on observed word counts - were moderately correlated, suggesting that these techniques provide somewhat comparable information about children's relative Spanish-to-English exposure levels. We should note that the only true test of a comparison of reported vs. observed estimates of exposure would be to obtain naturalistic recordings in the identical set of circumstances on which the estimates were derived based on the parental interview.

A second main goal of this study was to compare the power of these two estimates of exposure to predict language outcomes in emerging bilingual children. The results indicated some links between reported exposure and children's scores on a standardized test of language knowledge. However, observed child-directed words were linked to both RT and standardized test performance, even after controlling for reported exposure. The results for Spanish were particularly striking: those children who heard more Spanish words were more efficient in interpreting Spanish words in real time and scored higher on a test of Spanish knowledge. In addition, those children who heard more words in English were marginally faster to recognize English words during real-time comprehension and also tended to score higher on a test of English language proficiency. Such links between variation in bilingual

children's exposure and scores on a standardized test of language knowledge are consistent with previous findings showing relations between language experience and children's performance on standardized tests (Umbel, Pearson, Fernandez, & Oller, 1992). In this study, we further explored children's speed of processing during a test of real-time language comprehension. The links between exposure and RT found here are important given that differences among children in early language processing skills have been shown to predict school-relevant abilities several years later (Marchman & Fernald, 2008). It is not immediately clear why the associations to RT were substantially stronger in Spanish than English, however, it is possible that limitation in the range of English CDS/hr may have been a contributor.

Limitations of the present study

While providing the first direct comparison of reported vs. observed measures of bilingual children's language exposure in relation to early language outcomes, this study also had several limitations. One limitation is that our sample size was quite small, with only modest power to detect significant relations. Moreover, the children were tested at only one age, 3 years, and we focused only on language outcomes. Future studies should continue to explore these relations in larger samples of children from a broader age range using additional measures of language and other cognitive abilities.

A second limitation is that our sample was drawn from a sub-population of Latino families in which children heard primarily Spanish with limited exposure to English. While there was some range in amount of English exposure, the mainly Spanish-dominant character of our sample may limit the generalizability of these results to other bilingual populations in which children are receiving more balanced input from birth on. Nevertheless, this particular group of Spanish-dominant children is of considerable social and economic importance. Latino children now represent 24% of the U.S. population 18 years of age and younger, with almost half of these children living in poverty (Garcia, 2015). Among those in the lowest SES group, the majority are growing up in families where caregivers speak primarily Spanish at home, with minimal knowledge of English, like many of the families in our sample. These children are likely to be classified as English-Language Learners (ELL) when they enter kindergarten, and prognoses for their future academic achievement are low. According to a recent analysis of data from a large, representative sample of children in the U.S., Spanish-speaking ELL children are at greater risk for deficiencies in reading and math than any other minority group (Garcia, 2015). Although our results may not extend to other populations of bilingual children, they are relevant to a growing population of vulnerable Latino children in Spanish-dominant families who have relatively little access to English before they enter school.

Other limitations of this study have to do with the inherent challenges of estimating bilingual children's exposure to each of their languages. Estimates of bilingual exposure based on parent report were shown here to have some validity for predicting children's outcomes, however, they were only moderately associated with the direct measures based on the home recordings. Thus, it is possible that parents may have difficulty reporting on children's language exposure or they may be biased to under- or over-report particular sources of one

or the other language. The “day in the life” method used here was designed to ease the burden of these judgments compared to global assessments, however, it may nevertheless be the case that such techniques may be difficult for some parents.

Our hypothesis was that direct observations, such as those based on our home recordings, have the potential to identify important variation in children's experience in Spanish and English that is not captured by the estimates based on parent report. However, the fact remains that it is extremely difficult to capture a representative sample of a bilingual child's language experience. Bilingual environments are notoriously heterogeneous, with many more influential sources of variation than in monolingual environments. For example, Place and Hoff (2011) showed that the English language skills of Spanish/English bilinguals from a higher-SES Latino population were influenced by numerous factors, including the number of different speakers from whom the child heard English, and the extent to which the child's exposure to English was from native speakers. Our estimates of the numbers of words a child heard in Spanish and English based on the naturalistic recordings accounted for significant variance in predicting children's outcomes, but these estimates must have been influenced by the idiosyncrasies of the particular day that parents chose to record, and many other factors. In future research, it will be important to make more extensive recordings over longer time periods, taking full advantage of new recording and analysis technology that provides unique access to the daily interactions of bilingual children and their caregivers. Moreover, learning outcomes are supported not just by more talk, but by more high quality talk (Hart & Risley, 1995). Exploring the specific features of talk that support language learning, both in terms of general language knowledge and processing skill, are an ongoing effort in our and others' research programs (e.g., Rodríguez et al., 2009; Byers-Heinlein, 2012).

Conclusions

The results of this research highlight an important parallel between the language development of children from bilingual and monolingual learning environments, showing the critical role of verbal engagement between caregivers and children in shaping children's early language outcomes in both of the languages they were learning. The finding that links between variation in quantity of caregiver talk and processing speed reinforces the view that early engagement with caregivers not only provides young children with information that supports the development of language knowledge, but also tunes up information-processing skills fundamental for later language and cognitive growth (Hurtado et al., 2008; Weisleder & Fernald, 2013). Moreover, those children who are experiencing less supportive learning environments early in development are less likely to gain strength in those critical language learning skills and thus may be at increased risk for poorer language and cognitive outcomes than children from more supportive early learning environments (Fernald et al., 2012).

We hope that these results will encourage researchers to explore new methods of sampling bilingual children's home learning environments. With further refinement of the techniques for capturing variation in bilingual children's early language experiences, future work will lead to greater understanding of the intricate relations between language exposure and the development of language processing skill and other dimensions of language proficiency.

These relations are of critical importance in early language learning, regardless of whether a child is learning one language or two.

Acknowledgments

We are grateful to the children and parents who participated in this research. Special thanks to Adriana Weisleder, Araceli Arroyo, Lucía Rodríguez Mata, Ricardo Hoffman Bion, Katherine Adams, Kyle MacDonald and the staff of the Language Learning Lab at Stanford University. This work was supported by grants from the National Institutes of Health to Anne Fernald (DC 008838), and an NIH Research Supplement for Underrepresented Minorities to Nereyda Hurtado and Lucía Martínez.

References

Allen S. The future of Inuktitut in the face of majority languages: Bilingualism or language shift? Applied Psycholinguistics. 2007; 28:515–536.

Byers-Heinlein K. Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. Bilingualism: Language and Cognition. 2012; doi: 10.1017/S1366728912000120

Dale PS, Fenson L. Lexical development norms for young children. Behavior Research Methods, Instruments & Computers. 1996; 28(1):125–127.

De Houwer A. Parental language input patterns and children's bilingual use. Applied Psycholinguistics. 2007; 28:411–424.

De Houwer A. Language input environments and language development in bilingual acquisition. Applied Linguistics Review. 2011:221–239.

De Houwer, A. The absolute frequency of maternal input to bilingual and monolingual children: A first comparison. In: Grüter, T.; Paradis, J., editors. Input and Experience in Bilingual Development. Amsterdam: John Benjamins; 2014. p. 37-58.

Thordardottir, Elin; Rothenberg, A.; Rivard, M-E.; Naves, R. Bilingual assessment: Can overall proficiency be estimated from separate measurement of two languages? Journal of Multilingual Communication Disorders. 2006; 4:1–21.

Thordardottir, Elin. The relationship between bilingual exposure and vocabulary development. International Journal of Bilingualism. 2011; 15:426–445.

Fernald, A. When infants hear two languages: Interpreting research on early speech perception by bilingual children. In: McCardle, P.; Hoff, E., editors. Childhood Bilingualism: Research on Infancy through School Age. Clevedon: Multilingual Matters; 2006. p. 19-29.

Fernald A, Marchman VA. Individual differences in lexical processing at 18 months predict vocabulary growth in typically-developing and late-talking toddlers. Child Development. 2012; 83:203–222. [PubMed: 22172209]

Fernald A, Perfors A, Marchman VA. Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the second year. Developmental Psychology. 2006; 42:98–116. [PubMed: 16420121]

Fernald A, Pinto JP, Swingley D, Weinberg A, McRoberts GW. Rapid gains in speed of verbal processing by infants in the second year. Psychological Science. 1998; 9:72–75.

Fernald, A.; Zangl, R.; Portillo, AL.; Marchman, VA. Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In: Sekerina, I.; Fernández, EM.; Clahsen, H., editors. Developmental Psycholinguistics: On-Line Methods in Children's Language Processing. Amsterdam: John Benjamins; 2008. p. 97-135.

Ford, M.; Baer, CT.; Xu, D.; Yapanel, U.; Gray, S. The LENA Language Environment Analysis System: Audio Specifications of the DLP-0121. 2009. p. 1-8.http://www.lenafoundation.org/TechReport.aspx/Audio_Specifications/LTR-03-2

Garcia E. Inequalities at the starting gate: Cognitive and non-cognitive skills gaps between 2010–2011 Kindergarten classmates. Economic Policy Institute Report – June 17, 2015. 2015

Gathercole VCM, Thomas EM. Bilingual first-language development: Dominant language takeover, threatened minority language take-up. Bilingualism - Language and Cognition. 2009; 12:213–237.

Grüter, T.; Hurtado, N.; Marchman, VA.; Fernald, A. Language exposure and online processing efficiency in bilingual development: Relative versus absolute measures. In: Grüter, T.; Paradis, J., editors. Input and Experience in Bilingual Development. Amsterdam: John Benjamins; 2014. p. 15-36.

Gutiérrez-Clellen VF, Kreiter J. Understanding child bilingual acquisition using parent and teacher reports. Applied Psycholinguistics. 2003; 24:267–288.

Hart, B.; Risley, TR. Meaningful Differences in the Everyday Experience of Young American Children. Baltimore, MD: Brookes Publishing Co.; 1995.

Hoff E. The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. Child Development. 2003; 74:1368–1878. [PubMed: 14552403]

Hoff E. How social contexts support and shape language development. Developmental Review. 2006; 26:55–88.

Hoff E, Core C, Place S, Rumiche R, Señor M, Parra M. Dual language exposure and early bilingual development. Journal of Child Language. 2012; 39:1–27. [PubMed: 21418730]

Hoff-Ginsberg E. How should frequency in input be measured? First Language. 1992; 12:233–244.

Hollingshead, AB. Four factor index of social status. Yale University; New Haven, CT: 1975. Unpublished manuscript

Hurtado N, Marchman VA, Fernald A. Spoken word recognition by Latino children learning Spanish as their first language. Journal of Child Language. 2007; 33:27–249.

Hurtado N, Marchman VA, Fernald A. Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. Developmental Science. 2008; 11:F31–F39. [PubMed: 19046145]

Hurtado N, Grüter T, Marchman VA, Fernald A. Relative language exposure, processing efficiency and vocabulary in Spanish–English bilingual toddlers. Bilingualism: Language and Cognition. 2013; Advance online publication. doi: 10.1017/S136672891300014X

Huttenlocher J, Haight W, Bryk A, Seltzer M, Lyons T. Vocabulary growth: Relation to language input and gender. Developmental Psychology. 1991; 27:236–248.

Jackson-Maldonado, D.; Thal, DJ.; Marchman, VA.; Newton, T.; Fenson, L.; Conboy, B. MacArthur Inventarios del Desarrollo de Habilidades Comunicativas: User's Guide and Technical Manual. Baltimore, MD: Brookes Publishing Co.; 2003.

Marchman VA, Fernald A. Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. Developmental Science. 2008; 11:F9–F16. [PubMed: 18466367]

Marchman VA, Fernald A, Hurtado N. How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish–English bilinguals. Journal of Child Language. 2010; 37:817–840. [PubMed: 19726000]

Marchman VA, Martínez-Sussmann C. Concurrent validity of caregiver/parent report measures of language for children who are learning both English and Spanish. Journal of Speech, Language & Hearing Research. 2002; 45:993–997.

Marchman VA, Martínez-Sussmann C, Dale PS. The language-specific nature of grammatical development: Evidence from bilingual language learners. Developmental Science. 2004; 7:212–224. [PubMed: 15320381]

Oetting JB, Hartfield L, Pruitt S. Exploring LENA as a tool for researchers and clinicians. The ASHA Leader. 2009; 14:20–22.

Oller DK, Niyogi P, Gray S, Richards JA, Gilkerson J, Xu D, Yapanel U, et al. Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107:13354–13359. [PubMed: 20643944]

Patterson J. Expressive vocabulary development and word combinations of Spanish–English bilingual toddlers. American Journal of Speech–Language Pathology. 1998; 7:46–56.

Paradis J, Genesee F. Syntactic acquisition in bilingual children: Autonomous or interdependent? Studies in Second Language Acquisition. 1996; 18:1–25.

Pearson BZ, Fernández SC, Lewedeg V, Oller DK. The relation of input factors to lexical learning by bilingual infants. Applied Psycholinguistics. 1997; 18:41–58.

Place S, Hoff E. Properties of dual language exposure that influence 2-year-olds' bilingual proficiency. Child Development. 2011; 82:1834–1849. [PubMed: 22004372]

Rodríguez ET, Tamis-LeMonda CS, Spellmann ME, Pan BA, Raikes H, Lugo-Gil J, Luze G. The formative role of home literacy experiences across the first three years of life in children from low-income families. Journal of Applied Developmental Psychology. 2009; 30:677–694.

Semel, E.; Wiig, EH.; Secord, WA. Clinical Evaluation of Language Fundamentals, Preschool (CELF-P). Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company; 2004.

Song L, Tamis-LeMonda CS, Yoshikawa H, Kahana-Kalman R, Wu I. Language experiences and vocabulary development in Dominican and Mexican infants across the first 2 years. Developmental Psychology. 2011; 48(4):1106–1123. DOI: 10.1037/a0026401 [PubMed: 22142187]

Unsworth S. Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. Bilingualism: Language and Cognition. 2013; 16:86–110.

Umbel VM, Pearson BZ, Fernandez MC, Oller DK. Measuring bilingual children's receptive vocabularies. Child Development. 1992; 63:1012–1020. [PubMed: 1505238]

Weisleder A, Fernald A. Talking to children matters: Early language experience strengthens processing and builds vocabulary. Psychological Science. 2013; doi: 10.1177/0956797613488145

Wiig EH, Secord WA, Semel E. Clinical Evaluation of Language Fundamentals. Preschool (Spanish version). 2009

Xu, D.; Yapanel, U.; Gray, S. Reliability of the LENA™ language environment analysis system in young children's natural home environment. 2009. Retrieved from http://www.lenafoundation.org/TechReport.aspx/Reliability/LTR-05-2

Research Highlights

- Traditional estimates of Spanish-English exposure based on parent report are compared to measures of the number of Spanish and English words that bilingual children heard during naturalistic audio recordings of their interactions with others.
- Observed estimates revealed considerable variability in the number of words that children heard in each of their languages.
- Number of child-directed words was a more consistent predictor of children's language outcomes than estimates based on parent reports, even when controlling for parent-report estimates of relative exposure.
- Variation in quantity of child-directed words is a significant predictor of bilingual children's language outcomes in both of the languages they are learning.

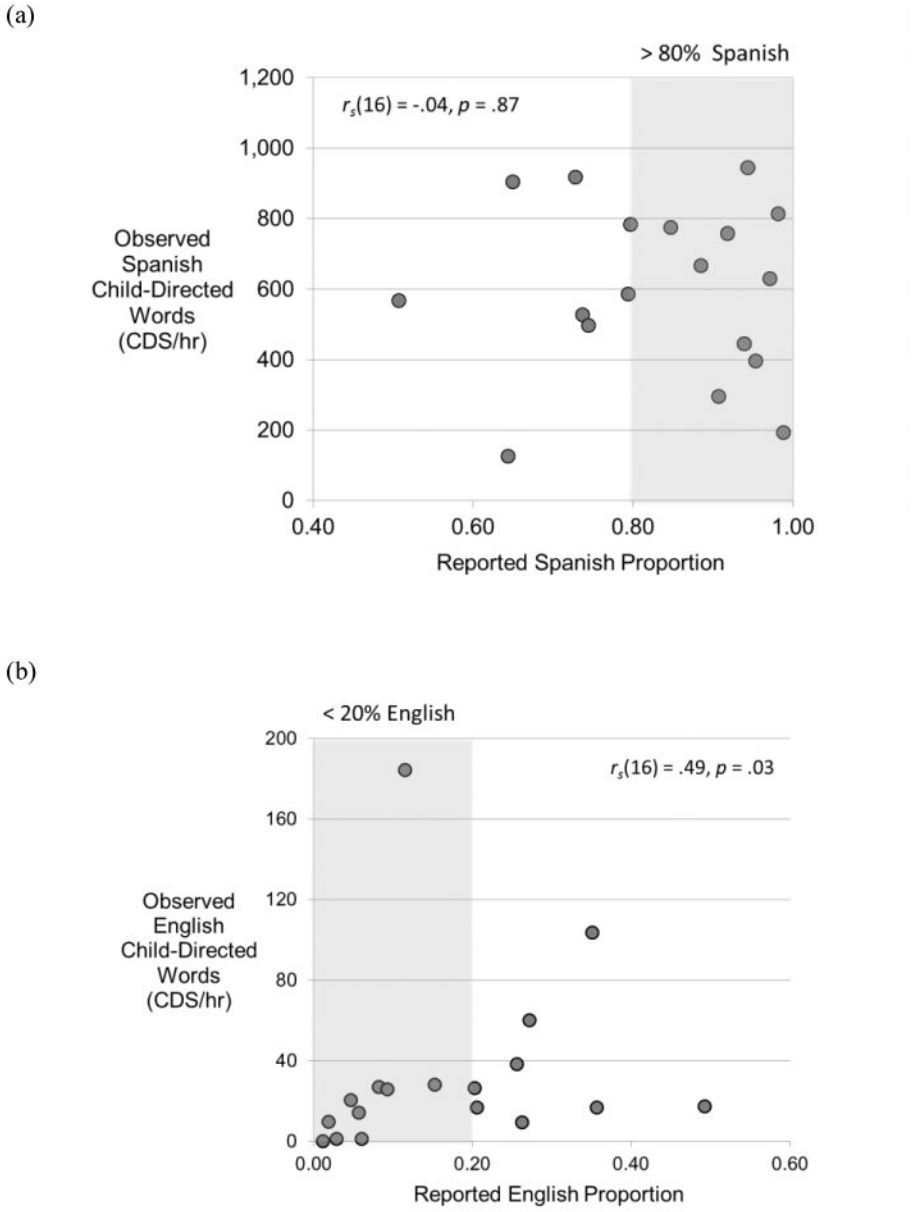

Figure 1.

a and b. Observed measures of child-directed speech (Mean CDS/hr) in relation to proportion of children's reported exposure in (a) Spanish and (b) English.

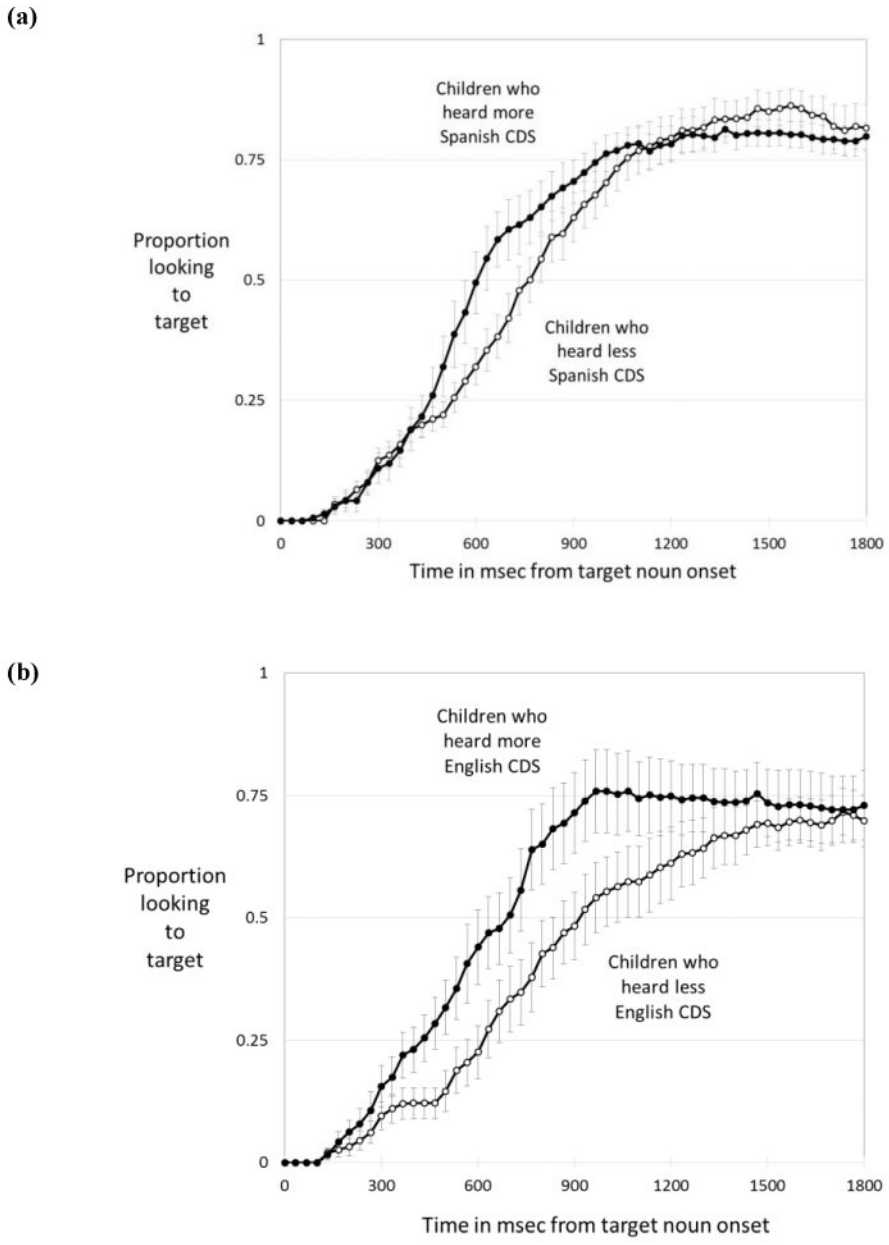

Figure 2.

a and b. Time course of spoken language processing to shift from distracter to target picture in the LWL task in children hearing more vs. less child-directed speech (CDS/hr) in (a) Spanish and (b) English.

Table 1

Descriptives of language exposure measures (*n* = 18)

Measure of Exposure	*M (SD)*	range
Reported		
Proportion Spanish[a]	.83 (.14)	.51-.98
Observed		
Spanish CDS/hr[b]	602.6 (244.9)	127.5 – 946.5
English CDS/hr[b]	33.6 (44.9)	0.1 – 185.5
Proportion CDS/hr[c]	.95 (.05)	.78– 1.0

[a]Proportion of person hours reported for Spanish out of all Spanish and English person hours based on a comprehensive Language Background Interview.

[b]Mean (*SD*) number of child-directed words (CDS/hr) identified as Spanish or English, normalized for length of recording.

[c]Mean proportion of child-directed words (CDS) that were identified as Spanish out of all Spanish and English child-directed words produced during the recording, normalized for length of recording.

Table 2

Descriptives of child measures of speed of processing on the LWL task and scores on a test of expressive/receptive language in Spanish and English (*n* = 18)

	Spanish	English
Looking-while-listening (LWL)		
RT[a]	763 (151)	813 (202)
CELF-P		
Core Language[b]	100.6 (12.9)	70.0 (12.0)

[a]Mean (*SD*) latency to shift from the distracter to target picture on all distracter-initial trials within 300-1800 ms from critical word onset.

[b]Standard score on the Clinical Evaluation of Language Fundamentals-Preschool (CELF-P) Core Language composite in Spanish (Wiig et al., 2009) and English (Semel et al., 2004).

Table 3

Nonparametric partial correlations (Spearman's Rho, r_S) between children's outcomes (Language Processing Speed and Language Knowledge) and Reported Spanish/English Proportion and Observed Spanish and English words in each language (adjusted for recording length), controlling for SES and age at test. For the Observed measures, the second number in each pair is the nonparametric partial correlation controlling for SES, age and Reported Spanish/English proportion ($n = 18$).

	Reported	Observed	
	Spanish/English Proportion	**Spanish CDS/hr**	**English CDS/hr**
Lexical Processing Speed (RT)			
Spanish	-.01	**-.71****/**-.72****	-.12/-.13
English	.03	-.12/-.12	**-.47**$^{\#}$/**-.53***
Language Knowledge (CELF-P)			
Spanish	.39	.41/**.48**$^{\#}$	-.15/.05
English	**-.46**$^{\#}$	.10/.07	**.53***/.40

Note:

$^{\#}p = .07$;

$^{*}p < .05$;

$^{**}p < .01$