Published in final edited form as:

Child Dev. 2018 September ; 89(5): 1674–1690. doi:10.1111/cdev.12818.

Caregiver Talk and Medical Risk as Predictors of Language Outcomes in Full Term and Preterm Toddlers

Katherine A. Adams^a, Virginia A. Marchman^a, Elizabeth C. Loi^b, Melanie D. Ashland^a, Anne Fernald^a, and Heidi M. Feldman^b

^aDepartment of Psychology, Stanford University

^bDepartment of Pediatrics, Stanford University School of Medicine

Abstract

This study examined associations between caregiver talk and language skills in full term (FT) and preterm (PT) children (n=97). All-day recordings of caregiver-child interactions revealed striking similarities in amount of caregiver talk heard by FT and PT children. Children who heard more caregiver talk at 16 months demonstrated better knowledge- and processing-based language skills at 18 months. The unique contributions of caregiver talk were tempered by medical risk in PT children, especially for processing speed. However, there was no evidence that birth status or medical risk moderated the effects of caregiver talk. These findings highlight the role of caregiver talk in shaping language outcomes in FT and PT children and offer insights into links between neurodevelopmental risk and caregiver-child engagement.

Identifying the causes of variation in children's early language development is a critical public health concern because early oral language skills form the foundation for later cognitive and language outcomes (Marchman & Fernald, 2008). One powerful correlate of language outcomes is the amount of talk directed to children during interactions with caregivers. Amount of caregiver talk is an index of a child's language nutrition (Zauche, Thul, Mahoney, & Stapel-Wax, 2016), shown to predict individual variation in vocabulary knowledge and early processing skills that support vocabulary learning in children born full term (FT) (e.g., Hoff, 2003; Hurtado, Marchman, & Fernald, 2008; Weisleder & Fernald, 2013). Children born preterm (PT) are at particular risk for poor developmental outcomes, yet also demonstrate considerable variability in their language skills (Marchman, Adams, Loi, Fernald, & Feldman, 2015; van Noort-van der Spek, Franken, & Weisglas-Kuperus, 2012). Such variation has been attributed in part to the neurobiological risks associated with PT birth (Allen, 2008), as well as to variation in features of children's learning environments (e.g., Landry, Smith, Miller-Loncar, & Swank, 1997; Smith, Swank, Denson, Landry, Baldwin, & Wildin, 1996). Using measures of number of adult words heard during naturalistic day-long home recordings in FT and PT children, the goals of this study were to determine if PT children are exposed to different amounts of caregiver talk compared to their FT peers, to examine relations between the quantity of caregiver talk and both knowledge-

and processing-based language outcomes in FT children, and to explore whether variation in caregiver talk predicted PT language outcomes in a similar way as in the FT group. Finally, we evaluated whether medical risk is a moderator of relations between quantity of caregiver talk and language skills of PT children.

Caregiver talk to FT and PT children

In a seminal observational study, Hart and Risley (1995) explored the quantity of caregiver talk across a diverse sample of families. They found that children from families low in socioeconomic status (SES) heard, on average, thousands fewer words each day than their higher-SES peers, and this gap contributed to meaningful differences in long-term language outcomes. In studies using home observation and lab-based paradigms, children who heard more talk from their caregivers showed greater gains in both vocabulary (Huttenlocher, Haight, Bryk, Seltzer & Lyons, 1991; Rowe, 2012) and language processing efficiency, compared to children who heard less talk from caregivers (Hurtado et al., 2008; Weisleder & Fernald, 2013). One study with children ages 2–48 months assessed over a 6-month time period found that each 1000 additional words spoken to a child was associated with a 0.44 point increase in average scores on the Preschool Language Scale, Fourth Edition (PLS-4), a measure of receptive and expressive vocabulary (Zimmerman, Gilkerson, Richards, Christakis, Xu, Gray, & Yapanel, 2009). On the basis of these and other findings, intervention efforts targeting quantity of caregiver talk as a means to improve language skills in children growing up in poverty are now increasing (e.g., Suskind, Leffel, Graf, Hernandez, Gunderson, Sapolich, & Levine, 2015).

Additional research is required to determine if quantity of caregiver talk is an important factor in language development for other populations at risk for language delay, such as children born PT. Children born PT are generally slower than their full term peers to reach early language milestones (Landry, Smith, & Swank, 2002), and language delays can persist into the school years, even in the absence of general cognitive delays (Lee, Yeatman, Luna, & Feldman, 2011). Studies suggest that there is great variability in parents' responsiveness during interactions with PT children (Landry et al., 1997; Smith et al., 1996). Parents of children born PT may need to modulate their interactions to the appropriate level of stimulation to avoid over- or underwhelming the PT child's information processing or arousal regulation capacities (e.g., Eckerman, Hsu, Molitor, Leung, & Goldstein, 1999). This narrower "activation band" for optimal parent-child interaction (Field, 1982) may increase the likelihood of low-quality interactions because PT children may show greater negative social responses to parental bids, and parents may engage in more controlling and over-stimulating behaviors compared to dyads with FT children (Forcada-Guex, Pierrehumbert, Borghini, Moessinger, & Muller-Nix, 2006). PT children and their caregivers may be less successful as social partners (De Schuymer, De Groote, Beyers, Striano, & Roeyers, 2011), possibly resulting in a reduction in the amount of caregiver talk heard by PT compared to FT children. Other studies find no differences between FT and PT groups in measures of maternal sensitivity (for a review, see Bilgin & Wolke, 2015) or the amount of caregiver talk heard during play (Salerni, Suttora, & D'Odorico, 2007). Given these inconsistent findings, additional research is needed to explore group differences in the language environment of FT and PT children.

In this study, we collected day-long naturalistic interactions using the LENA™ audio-recording and analysis system (Xu, Yapanel, & Gray, 2009) which provides estimates of the adult word count (AWC) heard by children over the course of a day. Most previous research on caregiver talk with PT children has relied on short recordings that sample caregivers' speech during activities associated with peak language use, such as playing with toys or book reading. While such samples provide useful snapshots of many features of caregiver-child interactions, these data do not reflect the natural ebb and flow of caregiver-child engagement throughout a full day. Moreover, these samples often involve goal-directed tasks that may necessitate higher levels of directive or over-stimulating behavior which could alter the natural dynamics of caregiver-child interactions (Field, 1982). LENA-recorded samples of caregiver talk are collected over a much longer time period (10–16 hours), typically in the home environment, and thus may have fewer observer artifacts (Zauche et al., 2016) than samples collected in shorter, laboratory-based samples. Using full-day naturalistic recordings may be more sensitive than laboratory observations in capturing meaningful group differences in caregiver-child language interactions.

Previous studies using LENA™ with PT children suggest that higher amounts of caregiver talk support language function in this at-risk population. For example, PT children who heard more words from caregivers during Neonatal Intensive Care Unit (NICU) visits had higher rates of child vocalizations (Caskey, Stephens, Tucker, & Vohr, 2011) and larger vocabularies when the children were 18 months (Caskey, Stephens, Tucker, & Vohr, 2014). However, estimates of amount of caregiver talk in the NICU may not necessarily reflect the amount of talk that children experience during daily interactions with caregivers outside of the NICU environment. Here, we directly compare caregiver talk captured in the naturalistic language environments of FT and PT children when they were 16 months old.

Predictors of language delay in PT children

Preterm birth is associated with increased risk for delay in knowledge-based domains of language function, including vocabulary and syntactic structures (Adams-Chapman, Bann, Carter, Stoll, & NICHD Neonatal Research Network, 2015; Foster-Cohen, Friesen, Champion, & Woodward, 2010; Sansavini, Guarini, Savini, Broccoli, Justice, Alessandrini, & Faldella, 2011). Children born PT are also at high risk for delays in tasks that involve real-time processing of linguistic stimuli or verbal working memory (Jansson-Verkasalo, Valkama, Vainionpää, Pääkkö, Ilkko, & Lehtihalmes, 2004; Lee et al., 2011). Some studies suggest that PT children show greater difficulty in complex language processing tasks than in knowledge-based assessments (Bosch, 2011). Tasks requiring efficient language processing may reveal differences that are not captured as effectively in untimed tasks (Lee et al., 2011; Venker & Kover, 2015). In young children, individual differences in language processing speed predict later language outcomes in FT (Fernald & Marchman, 2012) and PT populations (Marchman et al., 2015). In this study, we use both knowledge- and processing-based measures of language proficiency.

Higher risk of language delay in PT children has been attributed in part to perinatal factors, such as gestational age, birth weight, and medical complications. Foster-Cohen and colleagues found that younger gestational age was associated with smaller parent-reported

vocabulary (Foster-Cohen, Edgin, Champion, & Woodward, 2007). The Nursery Neurobiologic Risk Score (Brazy, Eckerman, Oehler, Goldstein, O'Rand, 1991), capturing conditions such as intraventricular hemorrhage, periventricular leukomalacia, and hypoglycemia, is a robust predictor of cognitive and language skills at 24 months (Wickremasinghe, Hartman, Voigt, Katusic, Weaver, Colby, & Barbaresi, 2012). Similarly, neonatal illnesses associated with PT birth, such as bronchopulmonary dysplasia and necrotizing enterocolitis, increase risk for neurodevelopmental impairments related to language delay (Behrman & Butler, 2007; Landry, Chapieski, Fletcher, & Denson, 1988; Rees, Pierro, & Eaton, 2007; Singer, Siegel, Lewis, Hawkins, Yamashita, & Baley, 2001). Other studies suggest that variability in language skills may be attributable to a combination of neurobiological, perinatal, and environmental factors that together shape language development (Cusson, 2003; Foster-Cohen et al., 2010), such as length of hospital stay, birth weight, APGAR scores, infant irritability and features of parenting quality.

Here we examined associations between variability in the amount of talk heard from caregivers and young children's knowledge- and processing-based language outcomes, with particular emphasis on how medical factors might contribute to and interact with this relation in children born PT. One possibility is that there may be a weaker relation between caregiver talk and child outcomes in PT compared to FT infants (Field, 1982). If increased sensitivity to stimulation in PT children, particularly those of high medical risk, leads to more variability in responsiveness to caregiver engagement, then increases in quantity of caregiver talk might not be linked to the same level of improvements in PT as FT children. A second possibility is that optimal parent-child interactions might play an especially large role in supporting PT language development, following a model of differential susceptibility (Gueron-Sela, Atzaba-Poria, Meiri, & Marks, 2015; Madigan, Wade, Plamondon, Browne, & Jenkins, 2015). That is, stronger associations between caregiver talk and language outcomes would be predicted for higher-risk PT compared to lower-risk PT or FT children (Landry et al., 1997; Smith et al., 1996). A third possibility is that caregiver talk might be an equally strong predictor of outcomes in FT and PT children, with parallel effects across groups regardless of medical risk status.

Research Questions

This study examined relations between children's learning environment, measured by AWC, the number of words spoken by caregivers during day-long naturalistic recordings when their toddlers were 16 months old, and language outcomes at 18 months in FT and PT children. We applied stringent criteria for exploring these relations, controlling not only for family SES, child sex, and the child's own number of vocalizations, but also the child's level of performance at an earlier time point. We assessed children's outcomes in two ways. First, we evaluated language knowledge using a well-known standardized measure of expressive and receptive language, the Bayley Scales of Infant and Toddler Development, 3rd Edition (BSID-III, Bayley, 2006). We also assessed language skills using an information-processing task that captures children's accuracy and speed in comprehending familiar words in real time. Our four main hypotheses were:

1. Group differences will be observed between FT and PT children in the number of adult words (AWC) heard during naturalistic recordings over the course of a typical day.
2. For children born FT, variation in AWC will account for variance in children's language knowledge and processing outcomes, after controlling for SES, gender and the child's earlier language skill.
3. For children born PT, variation in AWC will be associated with outcomes in ways that are different to those seen in the FT group.
4. Medical risk will account for additional variance in outcomes in PT children, beyond caregiver talk, and will moderate the associations between AWC and child outcomes.

Method

Participants

Participants were 97 children (48 males, 49 females) born either full term (FT, $n=56$, 27 males, 29 females), with a gestational age ≥ 37 weeks, or preterm (PT, $n=41$, 21 males, 20 females), with a gestational age ≤ 32 weeks and birth weight (BW) <1800 grams. Most families with PT children were recruited between the summer of 2011 and fall of 2014 from the NICU, intermediate care nursery, and the High-Risk Infant Follow-up Clinic of the local children's hospital. Others were recruited during that time period through a research registry, local parent groups, an early intervention service provider, or birth records. Children born FT were recruited between the summer of 2012 and winter of 2015 through birth records, staff or sibling referral, and parent registration on the lab's website. Exclusionary criteria for all children were known developmental disorders and medical conditions that would prevent child participants from engaging in our research assessments, such as a history of meningitis, a seizure disorder requiring medication, a ventriculoperitoneal shunt, a genetic disorder, and visual or hearing impairments. All children were primarily English speaking and heard a non-English language $\geq 25\%$ of the time, as reported by parents. The protocol was approved by a university IRB and parents gave signed consent at each visit.

All participants were drawn from a larger longitudinal study in which children born FT were individually matched (on the basis of SES, age, and sex) to a child born PT. This one-to-one matching procedure allowed for more reliable testing of the effect of birth status on children's outcomes. Children in both the FT and PT groups completed visits at approximately 16 months (FT chronological age: $M = 16.3$, range = 15.1 to 17.4 months; PT corrected age: $M = 16.1$, range = 15.2 to 17.6) and 18 months (FT chronological age: $M = 18.8$, range = 17.7 to 20.4 months; PT corrected age: $M = 18.6$, range = 17.7 to 20.3 months). The sample reported here includes unequal numbers of FT and PT children because some children were missing data or otherwise did not meet inclusion criteria for the current analyses. Table 1 shows the sociodemographic background of both groups of children. Participants were from mid- to high-SES backgrounds, with no difference across groups, $t(95) = 0.02$, $p = .99$, $d = .01$, although there was substantial variability in SES backgrounds in both groups.

Birth History and Medical Risk

Information about birth history is also presented in Table 1. Similar to Marchman et al. (2015), we catalogued the factors of extreme preterm birth (gestational age < 28 weeks), extended length of hospital stay (> 51 days), and 11 medical conditions (small for gestational age, respiratory distress syndrome, chronic lung disease/bronchopulmonary dysplasia, patent ductus arteriosus, necrotizing enterocolitis, intraventricular hemorrhage, periventricular leukomalacia, retinopathy of prematurity, seizures, hearing loss and hyperbilirubinemia/jaundice). For each child, the presence or absence of each factor was coded by research assistants in consultation with the last author using information in the daily progress notes, NICU discharge summaries, and summary letters from outpatient follow-up visits, and then summed to yield a medical risk score (max = 13, *Mdn* = 3.0). In this relatively healthy group of PT children, none had a hearing or visual impairment that precluded participation in the study. Based on either an MRI or a head ultrasound conducted during hospitalization, two children were identified as having white matter injury associated with periventricular leukomalacia. Seven children were identified as having an intraventricular hemorrhage, six of which were mild grades. For some analyses, the PT children were categorized as lower ($n = 22$) vs. higher ($n = 19$) medical risk based on median medical risk score.

Naturalistic day-long recordings

At 16 months, audio recordings were made using the LENA™ recording and speech analysis system. A digital language recorder worn by the child in the chest pocket of specialized clothing contains a directional microphone that enables unobtrusive recording of all adult-like speech that is produced “near and clear” to the target child during daily activities. Parents were asked to record on a typical weekday, following their regular routine. The recorder could be turned off whenever the caregiver chose to do so (e.g., when visitors requested not to be recorded), and all families were given the option to delete the entire recording. Families generally recorded when they were at home; however, they frequently also recorded in other locations, including outside or in other peoples’ homes, which reflected their normal routines. In the case of twins, each target child wore separate recorders on different days and thus, each recording captured the speech occurring during the interactions of that particular child with others. Raw recording lengths ranged from 8.4 to 16.0 hours ($M = 14.7$ hours), with variability due in part to the fact that some families turned off the recorder at different points in the day, while other families did not. Trained listeners “cleaned” all of the recordings, deleting segments when the child was sleeping or not wearing the recorder, or when the recording quality was poor. The lengths of the cleaned recordings ranged from 5.4 to 13.4 hours ($M = 9.5$ hours) over a single day, with no group differences (PT $M = 9.6$ hours, FT $M = 9.5$ hours), $t(95) = 0.10$, $p = .92$, $d = .02$. An additional 38 families were eligible for participation but not included in the final sample because they chose not to take the recorder (FT: $n = 3$; PT: $n = 9$), they returned it with less than 3.5 hours of recording (FT: $n = 2$; PT: $n = 2$), or they did not have complete data at both testing points within the target age window (FT: $n = 6$; PT: $n = 16$). These additional families came from similar SES backgrounds as the families in the final sample ($M = 56.7$, $SD = 9.9$), $t(133) = 1.4$, $p = .18$. While the PT children who did not meet eligibility criteria were slightly less healthy than the PT children in the final sample, the group difference on the

index of medical risk did not achieve statistical significance ($M = 4.3$, $SD = 2.1$), $t(65) = 1.9$, $p = .07$.

The home recordings were analyzed using the LENA™ analysis software which automatically derives the following measures for each 5 minute segment: (1) Adult word count (AWC): the number of words spoken “near and clear” to the child, excluding overlapping adult and child speech, TV, and music; (2) Child vocalizations (CVC): the number of speech-like sounds produced by the target child, including words, babbling, and pre-speech communicative sounds or “protophones” such as squeals, growls, or raspberries, but excluding crying, whining, and vegetative sounds (e.g., breathing, burping). This CVC measure captures the frequency with which children are vocalizing and thereby may influence the frequency of caregiver vocalizations. This measure is not intended to capture the complexity or developmental sophistication of children’s vocalizations. The accuracy of the AWC estimates for English-language recordings has been established in previous studies as highly reliable (Oller et al., 2010; Xu et al., 2009). Accuracy for CVC estimates are reported to be somewhat weaker (75%) (Oller et al., 2010; Xu et al., 2009), but are nevertheless within an acceptable range. To control for differences in recording length, AWC and CVC counts for each 5-minute segment were summed to create a total count over each child’s cleaned recording, and then divided by the number of hours in the cleaned recording length. This procedure produced the final measures of AWC/hr and CVC/hr.

Parent report and laboratory measures of child outcomes

Vocabulary size—Expressive vocabulary was assessed at 16 and 18 months with the MacArthur-Bates CDI: Words and Gestures (CDI: W&G), a reliable and valid parent report instrument appropriate for children 8 to 18 months (Fenson, Marchman, Thal, Dale, Reznick & Bates, 2007). Parents marked words that their child “understands and says” and a total expressive vocabulary score was derived (396 max).

Receptive and expressive language—When the children were 18 months, trained examiners administered the Bayley Scales of Infant and Toddler Development, 3rd Edition (BSID-III, Bayley, 2006), a standardized assessment for children aged 16 days to 42 months. Testing typically occurred over two visits, conducted 1 week apart. The composite Language Scale is commonly applied in clinical protocols and offers a comprehensive evaluation of a broad range of children’s language skills. Scaled scores were summed and converted to a standard score, based on the child’s age at test, corrected for degree of prematurity in the PT group.

Real-time language comprehension—Children’s efficiency in comprehending words in real time was assessed at 16 and 18 months using the Looking-While-Listening (LWL) procedure (Fernald, Zangl, Portillo, & Marchman, 2008). At each age, children were tested in two visits, typically one week apart. Pairs of pictures of familiar objects appeared on a screen and a prerecorded voice named one of the pictures. The child sat on their caregiver’s lap and the caregiver wore opaque sunglasses to block their view of the images. Looking patterns were video recorded and later coded offline. Each session lasted approximately 5 min.

Visual stimuli were color pictures of familiar objects, matched for animacy and visual salience. Target images familiar to children in this age range were presented in yoked pairs: *baby-doggy*, *birdie-kitty*, *ball-shoe*, and *book-car*. Target images were presented four times each as target and distracter, interspersed between four filler trials, yielding 32 experimental trials at each session for a total of 64 experimental trials. Trials were presented in two pseudo-random orders, such that target order and picture position were counterbalanced across participants and across sessions. Pictures were displayed on the screen for 2 sec prior to speech onset and remained onscreen for 1 sec after sound offset. In the auditory stimuli, target nouns were presented in sentence-final position followed by an attention-getter (e.g., *Where's the doggy? Do you like it?*). Mean length of target noun was 639 msec (range = 565–769).

To ensure that children were tested only on words with which they were familiar, parents were asked to indicate whether their child understood each of the target words at each visit. Trials on which the parent reported that the child did not understand the target word were excluded on a child-by-child basis. At 16 months, FT children were reported to understand significantly more target words than PT children (FT $M = 7.4$, PT $M = 6.9$ words), $t(95) = 2.0$, $p < .05$, $d = .41$. However, all children were reported to know at least four of the target words, and just over half (55 of 97, 56.7%) were reported to know all eight words. At 18 months, all children were reported to know at least five words and about 85% (82 of 97, 84.5%) knew all eight target words, with no group difference (FT $M = 7.8$; PT $M = 7.7$), $t(95) = 0.97$, $p = .34$, $d = .20$.

All LWL sessions were prescreened by trained research assistants blind to target side. Trials on which the child was judged to be inattentive or there was parent interference were not coded. Children were excluded from analyses if they did not have at least 16 coded trials, or 25% of all experimental trials. Children's gaze was identified for each 33-msec interval as either fixed on one of the images (*left* or *right*), or shifting between pictures or away (*off*). Based on fixations at target noun onset, trials were later designated as target-initial (T-initial) or distracter-initial (D-initial).

Language processing efficiency is reflected in two measures: First, *accuracy* was the mean proportion looking to the target image from 300–1800 msec after noun onset, divided by the total looking time to either image. At both ages, mean accuracy scores were based on significantly more coded trials for children in the FT than in the PT group (16 months: FT $M = 46.3$, PT $M = 40.2$ trials, $t(95) = 2.5$, $p < .02$, $d = .52$; 18 months: FT $M = 51.1$, PT $M = 44.1$; $t(95) = 3.8$, $p < .001$, $d = .76$). Second, *reaction time (RT)* was the mean latency (msec) to initiate a shift from distracter to target image only on D-initial trials. Shifts prior to 300 msec and later than 1800 msec were excluded because they were less likely to be in response to the target word. Mean RT scores were based on a minimum of two valid shifts with the FT children contributing more valid shifts, on average, than children in the PT group at the younger, but not at the older, age (16 months: FT $M = 17.7$, PT $M = 13.1$ trials, $t(95) = 2.9$, $p < .004$, $d = .60$; 18 months: FT $M = 20.2$, PT $M = 18.3$, $t(95) = 1.4$, $p = .14$, $d = .31$). One child in the PT group had only one valid shift at 16 months and therefore, analyses of RT at this age were based on 40 children in the PT group.

To evaluate coding reliability, 25% of the sessions were randomly selected and fully re-coded. At 16 months, inter-coder agreement was 98% for trial-level accuracy within .05, and the proportion of trials on which first-shift latency agreed within one 33-msec frame was 99%. At 18 months, accuracy reliability was 96% and latency reliability was 98%.

Results

Overview

We first present descriptive statistics of AWC from the LENA™ recordings at 16 months, asking whether there are group differences between FT and PT children. We next explore FT vs. PT group differences in child outcomes. Using a series of multiple regression models, we then examine the shared and unique contributions of AWC to FT and PT children's language knowledge on the BSID-III and lexical processing efficiency in the LWL task. We then explore whether the relation between caregiver talk and children's outcomes is parallel in the two groups, that is, whether birth status moderates these relations. Finally, we explore the unique contribution of medical risk in the PT group, asking whether medical risk moderates the relations of caregiver talk to child outcomes in children born PT. Prior to all analyses, it was determined that all predictor variables had no outliers. We confirmed that the predictors met the assumption for normality based on visual inspection of the residuals from the regression models ($SW = .946-.99$; $p = .05-.92$).

Caregiver talk during naturalistic recordings

As shown in Table 2, AWC/hr was quite similar in the FT and PT groups. At the same time, there was substantial variation in both groups, with some children hearing as few as 500 words/hr, on average, and other children hearing nearly seven times more, on average, close to 3500 words/hr. AWC/hr did not differ as a function of birth order or gender in either group (all $p > .17$). To determine if AWC/hr was influenced by the inclusion of twins, we constructed two sub-samples of PT children which included all singletons plus one randomly selected child from each twin pair. The AWC/hr values were nearly identical in both sub-samples (Sample 1: $M = 1740.2$; Sample 2: $M = 1735.0$), suggesting that including both members of each twin pair did not artificially inflate the AWCs for PT children.

Outcomes in children born FT and PT

Table 2 also reports on several child outcomes. At 16 months, there were no group differences in CVC/hr, a measure of all vocalizations, including non-vegetative sounds, such as growls and squeals, as well as babbling and words, produced by the target child during the naturalistic recording. This suggests that children in the FT and PT groups were similar in their overall frequency of vocalizations which could serve as bids for adult communicative responses. Given the lower reliability of the CVC measure reported in earlier studies (e.g., Oller et al., 2010), this finding should be interpreted with caution. On a parent reported measure of children's production of recognizable vocabulary words, children born FT were reported to produce significantly more different words at 16 months than their PT counterparts. This group difference was no longer reliable at 18 months. On the BSID-III, there were no reliable group differences yet substantial variation in both groups. On the LWL task, both FT and PT children looked to the named picture at above chance levels at

16, FT: $t(55) = 10.9, p < .001, d = 1.5$; PT: $t(40) = 6.4, p < .001, d = .1.0$, and 18 months, FT: $t(55) = 12.6, p < .001, d = 1.7$, PT: $t(40) = 11.8, p < .001, d = 1.8$. As a group, FT children were significantly more accurate than PT children at 16 months, a difference that was no longer reliable at 18 months. Finally, processing speed (RT) was comparable at both 16 and 18 months. Accuracy and RT on the LWL task were initially computed following standard practice (Fernald et al., 2008) of including only those trials on which the parents indicated that the child understood the target word. Because parents may have underestimated children's familiarity with the target words, we also ran the analyses including all target words used in the study. The patterns of findings were unchanged.

Predicting child outcomes from caregiver talk in children born FT

We now ask whether variation in the amount of caregiver talk (AWC/hr) at 16 months accounted for significant variance in FT children's outcomes at 18 months using linear regression models. In each analysis, control factors were SES and child sex, factors that are associated with variation in the amount of caregiver talk. We also controlled for the number of the child's own vocalizations on the recording (CVC/hr) because frequency of child bids for caregiver attention can influence the frequency of adult communicative utterances (e.g., as measured by AWC/hr). We also controlled for the child's prior language performance at 16 months because earlier scores are likely to be associated with later outcomes.

In Table 3, Models 1–3 examined predictors of FT children's scores on the BSID-III. Model 1 shows that both SES and child sex were significant predictors, indicating that boys and children from higher-SES groups tended to score more highly. Model 2 shows that child vocalizations and an earlier measure of child language, the CDI, contributed an R^2 change of approximately 19% variance above sex and SES, $F(2, 51) = 9.71, p < .001$. Finally, Model 3 reveals that AWC/hr contributed an R^2 change of nearly 8% additional variance, with all factors except SES remaining significant predictors, $F(1, 50) = 9.29, p = .004$. The coefficient for AWC in Model 3 can be interpreted as every 100 word increase in AWC/hr is associated with a 0.9 point increase in BSID-III total language score, controlling for SES, sex, and earlier child language skills. The overall model accounted for nearly 60% of the variance in BSID-III scores.

Models 4–6 examined the contribution of our predictors to accuracy in the LWL task. There was a non-significant contribution of SES and child sex (Model 4), and child vocalizations and child accuracy at the younger age together contributed significantly to accuracy at 18 months, accounting for about 18% additional variance, $F(2, 51) = 4.54, p = .015$ (Model 5). In Model 6, the addition of AWC/hr nearly doubled the variance accounted for, adding more than 15% additional variance in predicting processing efficiency, $F(1, 50) = 11.70, p = .001$. This model indicates that every 100 word increase in AWC/hr is associated with a nearly 1% increase in accuracy in LWL for FT children, controlling for SES, sex, and earlier child language skills.

Finally, child sex, but not SES, was a significant predictor of speed of lexical processing (RT) (Model 7), indicating girls were faster in lexical processing than boys. However, this effect was reduced when child vocalizations and RT at 16 months were added, $F(2, 51) = 11.92, p < .001$ (Model 8). Critically, AWC/hr added about 7% additional variance beyond

these control variables, $F(1, 50) = 7.43, p = .009$, with all factors accounting for nearly 50% of the variance (Model 9). Model 9 indicates that every 100 word increase in AWC/hr is associated with a 6.9 msec decrease in RT, controlling for SES, sex, and earlier child language skills. In sum, these results both replicate and extend earlier findings which demonstrate relations between FT children's early learning environments and later language outcomes (Weisleder & Fernald, 2013).

Predicting child outcomes from caregiver talk in children born PT

To examine the relations between caregiver talk and child outcomes in children born PT, three parallel series of models are presented in Table 4. On the BSID-III, neither SES nor child sex contributed (Model 10), but child vocalizations and child expressive vocabulary reported on the CDI at 16 months together added significant additional variance, $F(2, 36) = 7.20, p = .002$ (Model 11). Importantly, Model 12 indicated that AWC/hr accounted for approximately 20% additional variance, $F(1, 35) = 13.94, p = .001$, and that every 100 word increase in AWC/hr is associated with a 1.5 point increase in BSID-III total language scores, controlling for SES, sex, and earlier child language skills.

To directly test whether the patterns of relations between AWC and outcomes on the BSID-III differed by FT/PT group status, we assessed a moderation effect following Hayes (2013) using the PROCESS macro in SPSS. All tests of estimation used bias-corrected 95% confidence intervals with 10,000 bootstrap samples. The results indicated that the coefficient for the interaction between caregiver talk and group did not achieve statistical significance, $b = .002 (.004)$ [CI: $-.007, .010$]. Thus, as shown in Figure 1a, the relations between caregiver talk and child outcomes on the BSID-III, a standardized test of language, were parallel in the FT and PT groups.

Turning now to accuracy, neither demographic (Model 14) nor child variables at 16 months (Model 15) were significant predictors; however, AWC/hr accounted for nearly 16% additional variance, $F(1, 35) = 6.93, p = .013$ (Model 16). Model 16 suggests that every 100 word increase in AWC/hr is associated with nearly a 1% increase in accuracy, controlling for SES, sex, and earlier child language skills. As shown in Figure 1b, the relations between AWC/hr and processing accuracy were comparable in FT and PT children. A test of moderation indicated that the increase in variance accounted for by the interaction term did not reach statistical significance, $b = .001 (.001)$, [CI: $-.001, .001$], suggesting parallel relations in both the FT and PT children.

Finally, none of the control variables accounted for significant variance in RT in the PT group (Models 18 and 19). AWC/hr added a marginally significant 8% additional variance, $F(1, 34) = 3.45, p = .072$ (Model 20), however, the overall model was not significant. The regression coefficient for AWC indicates that every 100 word increase in AWC/hr is associated with an 8.4 msec decrease in RT, controlling for SES, sex, and earlier child language skills. The parallel effects of AWC/hr on residual variance in RT in FT and PT children are illustrated in Figure 1c, again reflected in a non-significant interaction term in a test of moderation, $b = .004 (.04)$, [CI: $-.084, .092$]

Outcomes as a function of medical risk in PT children

We now examine group differences between PT children with Higher vs. Lower medical risk. Notably, AWC/hr was not different in PT children from Higher ($n = 19$; $M = 1753.3$) vs. Lower ($n = 22$; $M = 1806.3$) medical risk groups, $F(1,37) = 0.05$, $p = .82$, $\eta^2_{\text{partial}} = .01$, controlling for SES and child gender. Thus, healthier PT children were not necessarily exposed to more caregiver talk, on average, than PT children who were experiencing more medical complications.

To determine if PT children higher in medical risk produced fewer words during the day-long recordings or scored more poorly on tests of language knowledge and processing than their healthier PT peers, controlling for SES and gender, we conducted a series of comparisons of means. CVC/hr did not differ in children with Higher ($M = 185.2$) vs. Lower ($M = 194.3$) medical risk, $F(1,37) = 0.02$, $p = .89$, $\eta^2_{\text{partial}} = .01$. Differences in reported vocabulary size were not significant at either the 16- (Higher: $M = 15.5$; Lower: $M = 29.1$), $F(1,37) = 2.6$, $p = .12$, $\eta^2_{\text{partial}} = .07$, or 18-month time points (Higher: $M = 45.3$; Lower: $M = 82.1$), $F(1,37) = 3.1$, $p = .09$, $\eta^2_{\text{partial}} = .08$. There was also no reliable group difference in BSID-III (Higher: $M = 98.2$, $SD = 17.4$; Lower: $M = 101.6$, $SD = 18.3$), $F(1,37) = 0.5$, $p = .50$, $\eta^2_{\text{partial}} = .01$.

In contrast, PT children with Higher medical risk were significantly less accurate at 18 months, on average, in identifying a familiar referent during real-time language comprehension ($M = .63$, $SD = .09$), compared to children with Lower medical risk ($M = .68$, $SD = .07$), $F(1, 37) = 5.9$, $p < .02$, $\eta^2_{\text{partial}} = .14$. Moreover, PT children from the Higher Medical Risk group demonstrated slower RTs ($M = 838$, $SD = 152$) than those from the Lower Medical risk group ($M = 720$, $SD = 157$), $F(1,37) = 6.1$, $p < .02$, $\eta^2_{\text{partial}} = .14$, controlling for SES and child gender. Thus, child outcomes that reflect information processing skills during real-time language comprehension appear to be more susceptible than knowledge-based assessments to health-related factors in this PT sample.

To further illustrate the effect of medical risk on child performance in the LWL task, Figure 2 plots the time course of the proportion looking to target over the course of the stimulus sentence at 16 and 18 months for PT children from Lower and Higher Medical Risk groups. For all children at both ages, looking to target is at chance levels at the onset of the target noun and increases as the sentences unfolds in time. However, PT children from the Lower Medical Risk group tended to increase their proportion target looking earlier in the sentence and reach a higher asymptote than PT children from the Higher Medical Risk group. Interestingly, the curve for the Higher Medical Risk children at 18 months essentially overlaps with the curve for the Lower Medical Risk children at 16 months, indicating that 18-month-old PT children at increased risk were performing at the same level overall as lower-risk children who were 2 months younger.

We now ask whether the contribution of medical risk to child outcomes is unique or overlapping in relation to caregiver talk and then test for moderation effects. In Table 4, the final model in each series explored the contribution of medical risk after the control variables and caregiver talk were taken into account. In Model 13, we see little role for medical risk in predicting additional variance in BSID-III, $F(1, 34) = .184$, $p = .671$, while

earlier child vocabulary and AWC/hr remained significant predictors. To assess the moderating effect of medical risk, we again used the PROCESS macro (Hayes, 2013) to evaluate the additional contribution of the interaction between AWC/hr and medical risk to the prediction of BSID-III scores. The interaction term was not significant, $b = .001$ (.003), [CI: $-.007, .007$]. As illustrated in Figure 3a, AWC/hr was associated with BSID-III scores to the same degree at lower (-1 SD), average, and higher ($+1$ SD) values of medical risk.

For accuracy, Model 17 shows that medical risk added nearly 18% additional variance beyond AWC/hr, $F(1, 34) = 9.58$, $p = .004$, nearly doubling the prediction of the entire model, and both remaining significant. Thus, the contributions of medical risk and AWC/hr were additive in predicting to processing accuracy. In the test of moderation, the interaction term between AWC/hr and medical risk was again not significant, $b = .001$ (.001), [CI: $-.001, .001$]. Figure 3b shows that higher risk PT children had lower accuracy scores overall, but there were comparable relations between AWC/hr and accuracy at lower, average, and higher values of medical risk.

Finally, medical risk significantly contributed to variation in PT children's RT scores (Model 21), adding more than 10% additional variance, $F(1, 33) = 4.87$, $p = .034$. Unlike BSID-III and accuracy, AWC/hr did not account for significant unique variance in RT after taking medical risk into account, with medical risk effectively eliminating the modest effect of AWC/hr on child RT. The test of moderation again revealed a non-significant interaction between AWC/hr and medical risk, $b = -.02$ (.035), [CI: $-.089, .054$]. Figure 3c shows that medical risk affects RT scores, but that links between AWC/hr and RT are roughly parallel at lower, average, and higher values of medical risk.

Discussion

This study examined differences in the amount of caregiver talk heard by FT and PT toddlers and assessed the contribution of caregiver talk to later knowledge- and processing-based language skills. This research revealed four main findings. First, there was remarkable similarity in the number of adult words that FT and PT children heard, on average, but there was also substantial variability in both groups. Second, amount of caregiver talk was significantly related to both knowledge-based and processing-based outcomes in FT children. Third, for PT children, caregiver talk was a significant predictor of language outcomes, accounting for substantial variance in BSID-III and in language processing accuracy. However, the relation of caregiver talk to processing speed did not achieve statistical significance. Tests of moderation revealed that relations between amount of caregiver talk and child outcomes were parallel across the FT and PT groups. Fourth, medical risk added significant variance to child outcomes in the PT group for accuracy and processing speed, but not for knowledge-based outcomes. Importantly, medical risk did not moderate the relation between caregiver talk and outcomes in the PT group. Thus, caregiver talk remains an important predictor of language outcomes in children born FT and PT.

FT and PT children hear similar amounts of caregiver talk

In the current study, we first hypothesized that we would find a difference in the amount of caregiver talk as a function of PT status. In contrast, our results suggested that the language

learning environments of SES-matched children born FT and PT are quite similar, on average. On this index of “language nutrition,” many parents of PT children were highly engaged in caregiver-child interactions despite the increased cognitive and social risks associated with PT birth. The LENA™ technology enabled unobtrusive naturalistic recordings of the child’s engagement with one or more caregivers during activities over the course of a full day. These estimates captured not just the number of words heard when a caregiver and child were engaging in one particular activity, but also reflected the frequency with which those interactions occurred during daily activities. As in earlier studies, there was considerable variation among families, with some children hearing nearly seven times more words on average than other children over the day. Importantly, there was much greater within-group than between-group variation in these FT and PT children.

Caregiver talk predicts outcomes in both FT and PT toddlers

Our main focus was the extent to which amount of caregiver talk predicted children’s language outcomes. Looking first in the FT children, we replicated previous findings by Weisleder and Fernald (2013); those FT children who heard more talk were also those who were faster and more accurate in directing their gaze to the object named by a familiar noun. We also documented similar relations on a widely-used standardized test of language knowledge, the BSID-III. Taken together, these results suggest that those FT children who experience more language-rich interactions from caregivers are better positioned to develop both stronger information-processing skills and a greater body of language knowledge, compared to peers who experience fewer interactions.

We asked if quantity of caregiver talk differentially predicted outcomes in FT and PT children, with either increased, decreased or parallel effects. Moderation analyses revealed largely parallel effects of the relations between caregiver talk and outcomes in FT and PT children. Thus we find no support for the hypothesis that the development of language skills by children with increased risk due to PT birth is more or less sensitive to the quantity of caregiver talk than children born FT at lower risk. Our results also suggested some similarities and differences in the patterns of relations between caregiver talk and our three outcomes for FT and PT children. Regardless of birth status, those children who heard more adult words tended to score higher on a test of language knowledge and to comprehend familiar words more accurately in real time, even after controlling for SES, gender, and earlier language skills. However, this relation was only marginally significant for processing speed in the PT group. This weaker overall predictive validity in PT than FT children in this particular measure may be due to weaker stability of RT in the PT group. In general, the models were weaker in predicting to processing speed and accuracy than to scores on the knowledge-based BSID-III.

What is the role of medical risk in PT children?

The next question of interest was the nature of the contribution of medical risk to outcomes in the PT group and whether medical risk moderated the relations between caregiver talk and outcomes. We found that PT children with higher medical risk did not hear less talk from caregivers, on average, than did children with lower medical risk. The regression analyses revealed that medical risk did not account for additional variance in BSID-III scores,

caregiver talk remained a significant predictor after medical risk was taken into account, and medical risk did not moderate the relation between AWC/hr and children's scores. These findings suggest that caregivers of children born PT who talk more may support the acquisition of their child's knowledge-based language skills regardless of the child's level of medical risk. It is possible that untimed, face-to-face measures like the BSID-III are administered in a manner that partially accommodates for behaviors associated with medical complications - for example, by allowing repetitions, breaks, and pauses to reorient the child to the task. Thus, knowledge-based measures may be particularly good at capturing underlying language skills, although they may also attenuate the effects of behavioral or attentional factors that influence performance during timed processing-based language tasks.

In contrast, medical risk accounted for significant variance in language processing efficiency. For accuracy, knowing a child's medical risk in addition to caregiver talk nearly doubled the model's prediction. For RT, medical risk added considerable variance, and eliminated the contribution of caregiver talk. Thus, efficiency of language processing during timed tasks may be particularly influenced by factors included in the medical risk composite used here. At the same time, the relations between caregiver talk and child outcomes were parallel at all levels of medical risk, indicating that medical risk did not moderate these effects.

These results provide support for intervention efforts that increase caregiver talk as a means of facilitating language development in FT and PT children. For PT children, medical risk should be considered an important additional determinant of outcomes in this early stage of development, especially in relation to processing-based measures. Note though that medical factors have been shown to decrease in their effect, while environmental factors, such as quality of caregiver-child interaction, increase in their role in shaping outcomes over the first few years of life (Miceli, Goeke-Morey, Whitman, Kolberg, Miller-Loncar, & White, 2000). It is possible that our results would have revealed a stronger link between caregiver talk and language processing speed in PT children at older ages. Critically, the effects of caregiver talk were not moderated by medical risk, suggesting that improving language environments is beneficial regardless of birth status or health history.

Limitations of the current study

One limitation is that this sample was relatively small, and it is possible that moderation effects would have been identified if we had a larger sample and increased statistical power. A second limitation was that the sample was predominantly higher SES. Because PT birth is disproportionately overrepresented in low-SES compared to higher-SES groups in the U.S. (Kramer, Seguin, Lydon, & Goulet, 2000), our findings may not be characteristic of those seen in a more nationally-representative sample. Moreover, SES is strongly related to the amount of caregiver talk that children hear, and thus the full range of caregiver talk for PT children may not be represented in this sample as would be found in a more diverse sample. Fourth, without additional coding, we did not have access to features beyond adult word count, such as diversity of vocabulary and syntactic complexity, that are known to support child language outcomes (Weizman & Snow, 2001). Such additional coding might also have revealed differences in style of caregiver talk. For example, caregivers may use more

directive talk with those PT children who may require more support for daily routines and be less capable of engaging in reciprocal interactions. More directive talk may result in caregivers devoting less of their interactions to providing names for objects, descriptions, or elaborations on the activities in which the child and caregiver are engaged (Rowe, 2012). A fifth limitation is that the audio recording alone could not provide access to non-verbal features of caregiver-child interactions, such as eye-gaze or tone of speech, which might have revealed group differences (Landry et al., 1997). Future studies should continue to explore both quantitative and qualitative features of caregiver talk using multiple methods that could impact language outcomes in children born PT.

Conclusions

This study led to several new discoveries regarding the effect of caregiver speech on children's language outcomes. We found that children heard similar amounts of caregiver talk over the course of a full-day home recording, regardless of birth status or medical risk. Moreover, variability in the amount of speech that children heard had strikingly similar predictive relations to language function in FT and PT children. Nevertheless, we found both similarities and differences in the patterns of relations between caregiver talk and our three outcomes in FT and PT children. In FT children, caregiver talk consistently predicted knowledge- and processing-based language functions. For PT children, caregiver talk predicted knowledge-based language functioning and processing accuracy, but caregiver talk was less consistently related to processing speed, especially when medical risk was taken into account. Critically, the potential benefits of caregiver talk to PT children's outcomes were equally evident in those children who were more or less vulnerable to the adverse consequences of preterm birth, suggesting that interventions that focus on improving caregiver talk may be beneficial regardless of children's birth status or medical risk. These findings highlight the critical role of caregiver talk in shaping language outcomes in both FT and PT children and offer new insights into the links between neurodevelopmental risk factors and caregiver-child engagement.

Acknowledgments

We are grateful to the children and parents who participated in this research. Special thanks to 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 and Heidi M. Feldman (R01 HD069150).

References

- Adams-Chapman I, Bann C, Carter SL, Stoll BJ. Language outcomes among ELBW infants in early childhood. *Early Human Development*. 2015; 91(6):373–379. DOI: 10.1016/j.earlhumdev.2015.03.011 [PubMed: 25955535]
- Allen MC. Neurodevelopmental outcomes of preterm infants. *Current Opinion in Neurology*. 2008; 21(2)
- Bayley N. Bayley Scales of Infant and Toddler Development. Third. Harcourt; San Antonio, TX: 2006.
- Bilgin A, Wolke D. Maternal Sensitivity in Parenting Preterm Children: A Meta-analysis; *Pediatrics*. 2015. 2014–3570. peds
- Bosch L. Precursors to language in preterm infants: speech perception abilities in the first year of life. *Progress in Brain Research*. 2011; 189:239–257. DOI: 10.1016/B978-0-444-53884-0.00028-2 [PubMed: 21489393]

- Brazy JE, Eckerman CO, Oehler JM, Goldstein RF, O'Rand AM. Nursery Neurobiologic Risk Score: Important factors in predicting outcome in very low birth weight infants. *The Journal of Pediatrics*. 1991; 118(5):783–792. PMID: 80047-2. DOI: 10.1016/S0022-3476(05) [PubMed: 2019935]
- Caskey M, Stephens B, Tucker R, Vohr B. Importance of Parent Talk on the Development of Preterm Infant Vocalizations. *Pediatrics*. 2011; 128(5):910–916. DOI: 10.1542/peds.2011-0609 [PubMed: 22007020]
- Caskey M, Stephens B, Tucker R, Vohr B. Adult Talk in the NICU With Preterm Infants and Developmental Outcomes. *Pediatrics*. 2014; 133(3):e578–e584. DOI: 10.1542/peds.2013-0104 [PubMed: 24515512]
- Cusson RM. Factors Influencing Language Development in Preterm Infants. *Journal of Obstetric, Gynecologic, & Neonatal Nursing*. 2003; 32(3):402–409. DOI: 10.1177/0884217503253530
- De Schuymer L, De Groote I, Beyers W, Striano T, Roeyers H. Preverbal skills as mediators for language outcome in preterm and full term children. *Early Human Development*. 2011; 87(4):265–272. DOI: 10.1016/j.earlhumdev.2011.01.029 [PubMed: 21330069]
- Eckerman CO, Hsu HC, Molitor A, L H, Goldstein RF. Infant arousal in an en-face exchange with a new partner: Effects of prematurity and perinatal biological risk. *Developmental Psychology*. 1999; 35(1):282–293. DOI: 10.1037/0012-1649.35.1.282 [PubMed: 9923482]
- Fenson L, Marchman V, Thal D, Dale P, Reznick S, Bates E. *MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual*. 2nd. Baltimore: Paul H. Brookes; 2007.
- 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(1):203–222. DOI: 10.1111/j.1467-8624.2011.01692.x [PubMed: 22172209]
- Fernald A, Perfors A, Marchman VA. Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*. 2006; 42(1):98–116. DOI: 10.1037/0012-1649.42.1.98 [PubMed: 16420121]
- Fernald A, Zangl R, Portillo AL, Marchman VA. Looking while listening: Using eye movements to monitor spoken language. *Developmental Psycholinguistics: On-Line Methods in Children's Language Processing*. 2008:113–132.
- Field TM. Affective displays of high-risk infants during early interactions. *Emotion and Early Interaction*. 1982:101–125.
- Forcada-Guex M, Pierrehumbert B, Borghini A, Moessinger A, Muller-Nix C. Early dyadic patterns of mother-infant interactions and outcomes of prematurity at 18 months. *Pediatrics*. 2006; 118(1):e107–114. DOI: 10.1542/peds.2005-1145 [PubMed: 16818525]
- Foster-Cohen S, Edgin JO, Champion PR, Woodward LJ. Early delayed language development in very preterm infants: evidence from the MacArthur-Bates CDI. *Journal of Child Language*. 2007; 34(3): 655–675. DOI: 10.1017/S0305000907008070 [PubMed: 17822143]
- Foster-Cohen SH, Friesen MD, Champion PR, Woodward LJ. High prevalence/low severity language delay in preschool children born very preterm. *Journal of Developmental and Behavioral Pediatrics: JDBP*. 2010; 31(8):658–667. DOI: 10.1097/DBP.0b013e3181e5ab7e [PubMed: 20613625]
- Gueron-Sela N, Atzaba-Poria N, Meiri G, Marks K. The Caregiving Environment and Developmental Outcomes of Preterm Infants: Diathesis Stress or Differential Susceptibility Effects? *Child Development*. 2015; 86(4):1014–1030. DOI: 10.1111/cdev.12359 [PubMed: 25875941]
- Hayes AF. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York, NY: Guilford Press; 2013.
- Hart B, Risley TR. *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing; 1995.
- Hoff E. The Specificity of Environmental Influence: Socioeconomic Status Affects Early Vocabulary Development Via Maternal Speech. *Child Development*. 2003; 74(5):1368–1378. DOI: 10.1111/1467-8624.00612 [PubMed: 14552403]
- 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(6):F31–F39. DOI: 10.1111/j.1467-7687.2008.00768.x [PubMed: 19046145]

- Huttenlocher J, Haight W, Bryk A, Seltzer M, Lyons T. Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*. 1991; 27(2):236–248. DOI: 10.1037/0012-1649.27.2.236
- Institute of Medicine (US) Committee on Understanding Premature Birth and Assuring Healthy Outcomes. Behrman RE, Butler AS, editors. *Preterm Birth: Causes, Consequences, and Prevention*. Washington (DC): National Academies Press (US); 2007. Committee on Understanding Premature Birth and Assuring Healthy Outcomes
- Jansson-Verkasalo E, Valkama M, Vainionpää L, Pääkkö E, Ilkko E, Lehtihalmes M. Language Development in Very Low Birth Weight Preterm Children: A Follow-Up Study. *Folia Phoniatrica et Logopaedica*. 2004; 56(2):108–119. DOI: 10.1159/000076062 [PubMed: 15001826]
- Kramer MsSéguin L, Lydon J, Goulet L. Socio-economic disparities in pregnancy outcome: why do the poor fare so poorly? *Paediatric and Perinatal Epidemiology*. 2000; 14(3):194–210. DOI: 10.1046/j.1365-3016.2000.00266.x [PubMed: 10949211]
- Landry SH, Chapieski L, Fletcher JM, Denson S. Three-Year Outcomes for Low Birth Weight Infants: Differential Effects of Early Medical Complications. *Journal of Pediatric Psychology*. 1988; 13(3): 317–327. DOI: 10.1093/jpepsy/13.3.317 [PubMed: 3199292]
- Landry SH, Smith KE, Miller-Loncar CL, Swank PR. Predicting cognitive-language and social growth curves from early maternal behaviors in children at varying degrees of biological risk. *Developmental Psychology*. 1997; 33(6):1040–1053. DOI: 10.1037/0012-1649.33.6.1040 [PubMed: 9383626]
- Landry SH, Smith KE, Swank PR. Environmental effects on language development in normal and high-risk child populations. *Seminars in Pediatric Neurology*. 2002; 9(3):192–200. DOI: 10.1053/spen.2002.35499 [PubMed: 12350040]
- Lee ES, Yeatman JD, Luna B, Feldman HM. Specific language and reading skills in school-aged children and adolescents are associated with prematurity after controlling for IQ. *Neuropsychologia*. 2011; 49(5):906–913. DOI: 10.1016/j.neuropsychologia.2010.12.038 [PubMed: 21195100]
- Madigan S, Wade M, Plamondon A, Browne D, Jenkins JM. Birth Weight Variability and Language Development: Risk, Resilience, and Responsive Parenting. *Journal of Pediatric Psychology*. 2015; 40(9):869–877. DOI: 10.1093/jpepsy/jsv056 [PubMed: 26112158]
- Marchman VA, Adams KA, Loi EC, Fernald A, Feldman HM. Early language processing efficiency predicts later receptive vocabulary outcomes in children born preterm. *Child Neuropsychology*. 2016; 22(6):649–665. DOI: 10.1080/09297049.2015.1038987 [PubMed: 26031342]
- 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(3):F9–F16. DOI: 10.1111/j.1467-7687.2008.00671.x [PubMed: 18466367]
- Miceli PJ, Goeke-Morey MC, Whitman TL, Kolberg KS, Miller-Loncar C, White RD. Brief Report: Birth Status, Medical Complications, and Social Environment: Individual Differences in Development of Preterm, Very Low Birth Weight Infants. *Journal of Pediatric Psychology*. 2000; 25(5):353–358. DOI: 10.1093/jpepsy/25.5.353 [PubMed: 10880066]
- Oller DK, Niyogi P, Gray S, Richards JA, Gilkerson J, Xu D, Warren SF. Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences*. 2010; 107(30):13354–13359. DOI: 10.1073/pnas.1003882107
- Rees CM, Pierro A, Eaton S. Neurodevelopmental outcomes of neonates with medically and surgically treated necrotizing enterocolitis. *Archives of Disease in Childhood - Fetal and Neonatal Edition*. 2007; 92(3):F193–F198. DOI: 10.1136/adc.2006.099929 [PubMed: 16984980]
- Rowe ML. A Longitudinal Investigation of the Role of Quantity and Quality of Child-Directed Speech in Vocabulary Development. *Child Development*. 2012; 83(5):1762–1774. DOI: 10.1111/j.1467-8624.2012.01805.x [PubMed: 22716950]
- Salemi N, Suttora C, D'Odorico L. A comparison of characteristics of early communication exchanges in mother-preterm and mother-full-term infant dyads. *First Language*. 2007; 27(4):329–346. DOI: 10.1177/0142723707081654

- Sansavini A, Pentimonti J, Justice L, Guarini A, Savini S, Alessandrini R, Faldella G. Language, motor and cognitive development of extremely preterm children: Modeling individual growth trajectories over the first three years of life. *Journal of Communication Disorders*. 2014; 49:55–68. DOI: 10.1016/j.jcomdis.2014.02.005 [PubMed: 24630591]
- Siegel AC, Singer LT, Lewis B, Hawkins S, Yamashita T, Baley J. Preschool language outcomes of children with history of bronchopulmonary dysplasia and very low birth weight. *Journal of Developmental and Behavioral Pediatrics*. 2001; 22(1):19–26. [PubMed: 11265919]
- Smith KE, Swank PR, Denson SE, Landry SH, Baldwin CD, Wildin S. The Relation of Medical Risk and Maternal Stimulation with Preterm Infants' Development of Cognitive, Language and Daily Living Skills. *Journal of Child Psychology and Psychiatry*. 1996; 37(7):855–864. DOI: 10.1111/j.1469-7610.1996.tb01481.x [PubMed: 8923228]
- van N der Spek IL, Franken MCJP, Weisglas-Kuperus N. Language Functions in Preterm-Born Children: A Systematic Review and Meta-analysis. *Pediatrics*. 2012; 129(4):745–754. DOI: 10.1542/peds.2011-1728 [PubMed: 22430458]
- Suskind DL, Leffel KR, Graf E, Hernandez MW, Gunderson EA, Sapolich SG, Levine SC. A parent-directed language intervention for children of low socioeconomic status: a randomized controlled pilot study. *Journal of Child Language*. 2016; 43(2):366–406. DOI: 10.1017/S0305000915000033
- Venker CE, Kover ST. An Open Conversation on Using Eye-Gaze Methods in Studies of Neurodevelopmental Disorders. *Journal of Speech, Language, and Hearing Research*. 2015; 58(6):1719–1732. DOI: 10.1044/2015_JSLHR-L-14-0304
- Weisleder A, Fernald A. Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science*. 2013; 24(11):2143–2152. DOI: 10.1177/0956797613488145 [PubMed: 24022649]
- Weizman ZO, Snow CE. Lexical output as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology*. 2001; 37(2):265–279. DOI: 10.1037/0012-1649.37.2.265 [PubMed: 11269394]
- Wickremasinghe AC, Hartman TK, Voigt RG, Katusic SK, Weaver AL, Colby CE, Barbaresi WJ. Evaluation of the ability of neurobiological, neurodevelopmental and socio-economic variables to predict cognitive outcome in premature infants. *Child: Care, Health and Development*. 2012; 38(5):683–689. DOI: 10.1111/j.1365-2214.2011.01281.x
- 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>
- Zauche LH, Thul TA, Mahoney AED, Stapel-Wax JL. Influence of language nutrition on children's language and cognitive development: An integrated review. *Early Childhood Research Quarterly*. 2016; 36:318–333. DOI: 10.1016/j.ecresq.2016.01.015
- Zimmerman FJ, Gilkerson J, Richards JA, Christakis DA, Xu D, Gray S, Yapanel U. Teaching by Listening: The Importance of Adult-Child Conversations to Language Development. *Pediatrics*. 2009; 124(1):342–349. DOI: 10.1542/peds.2008-2267 [PubMed: 19564318]

Figure 1.

a, b and c. In full term (FT, $n=56$) and preterm (PT, $n=41$) groups, models of the relations between Adult Word Count per hour (AWC/hr) and (a) language knowledge measured by the Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III), (b) accuracy and (c) processing speed (RT) controlling for SES, sex, and child vocalizations and child ability at 16 months.

Figure 2.

Mean proportion of looking to the target picture as a function of time in msec from noun onset for Lower Medical Risk and Higher Medical Risk preterm (PT, $n = 41$) children. Circles represent the time course of correct looking at 16 months; squares represent the time course of looking in the same children at 18 months. Error bars represent SE of the mean over participants.

Figure 3.

a, b and c. In preterm children (PT, $n = 41$) at lower (-1 SD), average, and higher (+1 SD) medical risk, model estimates of relations between Adult Word Count per hour (AWC/hr) and (a) language knowledge measured by the Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III), (b) processing accuracy, and (c) processing speed (RT), controlling for SES, sex, and child vocalizations and child ability at 16 months.

Table 1 Demographics of full term (FT, $n = 56$) and preterm (PT, $n = 41$) participants.

	FT		PT	
	<i>n (%) or M (SD)</i>	Range	<i>n (%) or M (SD)</i>	Range
% Male	27 (48.2)	–	21 (51.2)	–
Maternal Education (years)	16.6 (1.5)	12–18	16.9 (1.6)	12–18
SES (HI) ¹	58.7 (8.8)	22–66	58.7 (8.5)	30–66
Birth weight (g)	3566 (444)	2438–4649	1251 (281)	670–1750
Gestational Age (weeks)	40.2 (1.1)	37.6–43.4	29.9 (2.0)	26.9–32.7
Twins	0 (0)	–	14 (34.1%)	–
First Born	34 (60.7)	–	25 (61.0)	–
Medical Risk ²	–	–	3.5 (1.8)	1–7

¹Hollingshead Four-factor Index of Socioeconomic Status (HI, Hollingshead, 1975, possible range 8 = 66);²Composite of 13 medical risk factors (possible range = 0–13).

Table 2

Descriptives of LENA™ naturalistic recordings at 16 months and child outcomes at 16 and 18 months (corrected for prematurity) in full term (FT, $n = 56$) and preterm (PT, $n = 41$) infants.

	FT		PT						
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>t</i> (95)	<i>p</i>	<i>Cohen's d</i>				
Naturalistic recordings (LENA™) ¹									
AWC/hr	1642.8 (592.0)	1781.7 (594.2)	1.1	.26	.23				
CVC/hr	206.9 (83.6)	190.1 (80.6)	0.9	.32	.21				
Vocabulary (CDI) ²									
Words Produced: 16 months	39.5 (43.7)	22.8 (28.1)	2.2	.03*	.45				
Words Produced: 18 months	88.8 (86.7)	65.0 (67.4)	1.5	.15	.31				
Language (BSID-III) ³									
Language Scale: 18 months	105.7 (17.7)	100.0 (17.7)	1.6	.13	.32				
Processing Efficiency (LWL) ⁴									
Accuracy: 16 months	.63 (.09)	.59 (.09)	2.0	.05*	.44				
Accuracy: 18 months	.66 (.09)	.66 (.09)	0.1	.91	.01				
RT: 16 months	814 (144)	873 (209)	1.7	.10	.33				
RT: 18 months	722 (143)	775 (164)	1.7	.09	.34				

Note:

* $p < .05$, All two-tailed independent sample t-tests use $df = 95$, except RT at 16 months ($df = 94$).

¹ Adult Word Count (AWC) and Child Vocalizations (CVC) per hour based on naturalistic recordings using LENA™;

² Total number of words reported as “understands and says” on the MacArthur-Bates CDI: Words & Gestures (Fenson et al., 2007);

³ Standard scores (corrected for prematurity) on the Language Scale of the Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III; Bayley, 2006) at 18 months;

⁴ Mean proportion looking to target (accuracy) and mean latency to shift from distracter to target picture (RT) in the Looking-While-Listening (LWL) task.

Table 3

In full term toddlers (FT, $n = 56$), multiple regression models predicting BSID-III, accuracy and RT at 18 months from caregiver talk (AWC) in the naturalistic recordings at 16 months, controlling for demographic and child variables at 16 months. Unstandardized B (SE).

BSID-III Language ¹	Model 1	Model 2	Model 3
HI: SES ²	.52 (.23) *	.36 (.20)	.28 (.19)
Child Sex	16.70 (4.02) **	11.06 (3.72) **	10.02 (3.46) **
CVC/hr at 16 months ³	–	.05 (.02) *	.05 (.02) *
Vocabulary at 16 months ⁴	–	.13 (.05) **	.12 (.04) **
AWC/hr at 16 months ⁵	–	–	.01 (.01) **
<i>R²-change</i>	–	18.8% **	7.7% **
<i>Total R²</i>	31.8% **	50.6% ***	58.3% ***
LWL Accuracy ⁶	Model 4	Model 5	Model 6
HI: SES ²	.01 (.01)	.01 (.01)	.01 (.01)
Child Sex	.02 (.03)	–.01 (.03)	–.01 (.02)
CVC/hr at 16 months ³	–	.01 (.01)	.01 (.01)
Accuracy at 16 months ⁶	–	.38 (.14) **	.30 (.13) *
AWC/hr at 16 months ⁵	–	–	.01 (.01) **
<i>R²-change</i>	–	14.5% *	15.5% **
<i>Total R²</i>	3.8%	18.4% **	33.9% **
LWL RT ⁷	Model 7	Model 8	Model 9
HI: SES ²	.24 (2.06)	1.15 (1.79)	1.67 (1.70)
Child Sex	-112.0 (36.06) **	-53.53 (33.77)	-50.39 (31.85)
CVC/hr at 16 months ³	–	–.01 (.19)	.05 (.18)
RT at 16 months ⁷	–	.56 (.12) ***	.50 (.11) ***
AWC/hr at 16 months ⁵	–	–	–.07 (.03) **
<i>R²-change</i>	–	26.9% ***	7.5% **
<i>Total R²</i>	15.5% *	42.4% ***	49.8% ***

Note:

* $p < .05$;

** $p < .01$;

*** $p < .001$;

¹ Standard score on the Language Scale of the BSID-III (Bayley, 2006);

²Hollingshead Four-factor Index of Socioeconomic Status (Hollingshead, 1975);

³CVC/hour based on LENA™ recordings;

⁴Reported words produced on the MacArthur-Bates CDI (Fenson et al., 2007);

⁵AWC/hr based on LENA™ recordings;

⁶Mean proportion looking to target (accuracy) in the LWL task at 16 or 18 months;

⁷Mean latency to shift from distracter to target picture (RT) in the LWL task at 16 or 18 months.

Table 4

In preterm toddlers ($n = 41$), multiple regression models predicting BSID-III, accuracy and RT at 18 months from caregiver talk in the naturalistic recordings at 16 months, controlling for demographic and child variables at 16 months. Unstandardized B (SE).

BSID-III Language¹	Model 10	Model 11	Model 12	Model 13
HI: SES ²	.02 (.34)	.01 (.31)	-.04 (.27)	-.02 (.28)
Child Sex	6.07 (5.60)	2.72 (4.97)	.71 (4.29)	.36 (4.43)
CVC/hr at 16 months ³	–	-.02 (.04)	-.06 (.03)	-.06 (.03)
Vocabulary at 16 months ⁴	–	.36 (.10) **	.34 (.08) **	.35 (.09) **
AWC/hr at 16 months ⁵	–	–	.02 (.01) **	.02 (.01) **
Medical Risk ⁶	–	–	–	.58 (1.35)
<i>R² change</i>	–	27.7% **	19.7% **	0.3%
<i>Total R²</i>	3.0%	30.7% **	50.5% **	50.7% **
LWL Accuracy⁷	Model 14	Model 15	Model 16	Model 17
HI: SES ²	.01 (.01)	-.01 (.01)	-.01 (.01)	-.01 (.01)
Child Sex	.02 (.03)	.02 (.03)	.01 (.03)	.02 (.02)
CVC/hr at 16 months ³	–	.01 (.01)	.01 (.01)	.01 (.01)
Accuracy at 16 months ⁷	–	.16 (.16)	.15 (.15)	-.07 (.15)
AWC/hr at 16 months ⁵	–	–	.01 (.01) *	.01 (.01) *
Medical Risk ⁶	–	–	–	-.02 (.01) **
<i>R² change</i>	–	2.8%	15.9% *	17.6% **
<i>Total R²</i>	1.2%	4.0%	19.9%	37.5% **
LWL RT⁸	Model 18	Model 19	Model 20	Model 21
HI: SES ²	-.81 (3.03)	1.27 (3.38)	1.63 (3.27)	2.03 (3.11)
Child Sex	-46.04 (50.72)	-39.65 (50.85)	-27.0 (49.63)	-39.03 (47.34)
CVC/hr at 16 months ³	–	-.10 (.34)	.13 (.35)	.13 (.33)
Child RT at 16 months ⁸	–	.20 (.13)	.21 (.13)	.16 (.12)
AWC/hr at 16 months ⁵	–	–	-.08 (.05) #	-.07 (.04)
Medical Risk ⁶	–	–	–	30.67 (13.9) *
<i>R² change</i>	–	6.4%	8.4% #	10.6% *
<i>Total R²</i>	2.3%	8.7%	17.1%	27.8% #

Note:

# $p < .08$;

* $p < .05$;

^{**}
 $p < .01$;

¹Standard score for the Language Scale of the BSID-III (Bayley, 2006);

²Hollingshead Four-factor Index of Socioeconomic Status (Hollingshead, 1975);

³CVC/hour from LENA™ recordings;

⁴Reported words produced on the CDI: Words & Gestures (Fenson et al., 2007);

⁵AWC/hr from LENA™ recordings;

⁶Scores on a medical risk composite (max = 13);

⁷Mean proportion looking to target (accuracy) in the LWL task at 16 or 18 months;

⁸Mean latency to shift from distracter to target picture (RT) in the LWL task at 16 or 18 months ($n = 40$).