

Clinical Focus

Mapping the Early Language Environment Using All-Day Recordings and Automated Analysis

Jill Gilkerson,^a Jeffrey A. Richards,^a Steven F. Warren,^b Judith K. Montgomery,^c
Charles R. Greenwood,^d D. Kimbrough Oller,^{e,f} John H. L. Hansen,^g and Terrance D. Paul^a

Purpose: This research provided a first-generation standardization of automated language environment estimates, validated these estimates against standard language assessments, and extended on previous research reporting language behavior differences across socioeconomic groups.

Method: Typically developing children between 2 to 48 months of age completed monthly, daylong recordings in their natural language environments over a span of approximately 6–38 months. The resulting data set contained 3,213 12-hr recordings automatically analyzed by using the Language Environment Analysis (LENA) System to generate estimates of (a) the number of adult words in the child's environment, (b) the amount of caregiver–child interaction, and (c) the frequency of child vocal output.

Results: Child vocalization frequency and turn-taking increased with age, whereas adult word counts were age independent after early infancy. Child vocalization and conversational turn estimates predicted 7%–16% of the variance observed in child language assessment scores. Lower socioeconomic status (SES) children produced fewer vocalizations, engaged in fewer adult–child interactions, and were exposed to fewer daily adult words compared with their higher socioeconomic status peers, but within-group variability was high.

Conclusions: The results offer new insight into the landscape of the early language environment, with clinical implications for identification of children at-risk for impoverished language environments.

This research measured quantitative information about the natural language environment of typically developing infants and toddlers by using automated analysis of daylong audio recordings. Decades of research have established the importance of the early language environment for optimizing developmental trajectories (Hart & Risley, 1992, 1995; Hoff, 2003; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Rowe, 2008), and

recent advancements in audio recording and automated voice labeling through the Language Environment Analysis (LENA) System now allow for the efficient collection and analysis of large amounts of naturalistic language samples (Caskey, Stephens, Tucker, & Vohr, 2011; Caskey & Vohr, 2013; Warren et al., 2010; Weisleder & Fernald, 2013; Xu et al., 2008b; F. J. Zimmerman et al., 2009). Here, we report automated analysis results from a database designed to establish first-generation reference information about the landscape of talk and interaction experienced by typically developing American children. This research examined the validity of criterion LENA measures compared against professionally administered language assessments and parent questionnaires. As well, data collected in this relatively larger scale study allowed us to extend research on the early language environment previously reported by Hart and Risley (1992, 1995).

Why Measure the Early Language Environment?

Over the past two decades, a number of studies have shown that quantitative measures of language environment factors present before age three, including interaction behaviors and vocabulary exposure, can predict vocabulary use and lexical development (Hoff, 2003; Huttenlocher,

^aLENA Research Foundation, Boulder, CO

^bLife Span Institute, University of Kansas, Lawrence

^cDepartment of Communication Sciences and Disorders, Chapman University, Orange, CA

^dJuniper Garden's Children's Project, University of Kansas, Kansas City

^eSchool of Communication Sciences and Disorders and Institute for Intelligent Systems, University of Memphis, TN

^fKonrad Lorenz Institute for Evolution and Cognition Research, Klosterneuburg, Austria

^gSchool of Engineering and Computer Science, University of Texas at Dallas

Correspondence to Jill Gilkerson: jillgilkerson@lenafoundation.org

Editor: Krista Wilkinson

Associate Editor: Cynthia Cress

Received October 29, 2015

Revision received April 5, 2016

Accepted October 31, 2016

https://doi.org/10.1044/2016_AJSLP-15-0169

Disclosure: The authors have declared that no competing interests existed at the time of publication.

et al., 1991; Rowe, 2008) and that quantity and type of caregiver–child interaction play a significant role in language and cognitive development (Bornstein, Tamis-LeMonda, Hahn, & Haynes, 2008; Landry, Smith, & Swank, 2006; Tamis-LeMonda, Bornstein, & Baumwell, 2001). However, the amount and quality of interaction children experience early in life can vary widely between families, resulting in cumulative disparities that may negatively affect development. Hart and Risley (1992, 1995) reported that infants and toddlers living in poverty experience significantly fewer adult words addressed to them than their higher socioeconomic status (SES) peers, projecting a 30-million-word deficit for low-income children by age four. These findings are supported by subsequent research showing that infants and toddlers from higher SES households experience more complex language input in terms of sentence length and vocabulary richness compared with lower SES children (Hoff, 2003; Hoff & Naigles, 2002; Rowe, 2008). Note that the longitudinal research of Hart and Risley demonstrated a relationship between early child-directed language and IQ and academic success in grade school (Walker, Greenwood, Hart, & Carta, 1994).

Turn-taking interactions may be even more important than vocabulary exposure, because language, social, emotional, and cognitive development are influenced cumulatively through bidirectional, reciprocal interactions throughout childhood (Sameroff & Fiese, 2000). Children experiencing a high degree of parental responsivity in the first few years of life reach language milestones earlier (Tamis-LeMonda, et al., 2001), demonstrate higher cognitive functioning (Kochanska, Forman, & Coy, 1999; Landry, Smith, Swank, & Miller-Loncar, 2000), exhibit more adaptive social skills, and display fewer emotional and behavioral problems (Goldberg, Corter, Lojkasek, & Minde, 1990). Further, children with developmental delays, social-emotional disorders, and intellectual disabilities may themselves evoke fewer interactions with caregivers, which can negatively affect the amount of adult language input they experience over time (van Ijzendoorn, et al., 2007; Warren & Brady, 2007; Wheeler, Hatton, Reichardt, & Bailey, 2007; Yoder, Warren, Kim, & Gazdag, 1994). Thus, the developmental influence of early communicative interaction coupled with observed variability in experiences strongly suggests that children could benefit from direct measurement of their environment.

This relationship between the early language environment and child language and cognitive outcomes has motivated the development of a number of intervention programs directed toward enhancing talk and interaction in the home (see Roberts & Kaiser, 2011, for an overview). These parent-focused interventions assume that time and resources devoted specifically to training the caregiver to adjust behaviors within daily routines can be more impactful over the long term than the use of short and infrequent sessions spent directly with the child (Delaney & Kaiser, 2001; Hemmeter & Kaiser, 1994; Kaiser & Hancock, 2003; Landry, et al., 2006; Landry, Smith, Swank, & Guttentag, 2008). The growing number of these intervention

programs points to the need for effective measures of naturalistic talk and interaction to inform the approach for individual families, evaluate its effectiveness, and, perhaps most important, to facilitate identification of children experiencing impoverished language environments and parents who would crucially benefit from such programs.

How Can the Early Language Environment Be Measured?

Measures of the natural language environment have traditionally involved collecting audio or video recordings in the home followed by labor-intensive transcription and coding. As one of the first intensive attempts to record and document the naturalistic home language environment of young children, the research of Hart and Risley (1992, 1995) was vital in establishing the link between the quantity and nature of caregiver talk and early language and cognitive development. Investigators visited the homes of 42 families once a month for three years, recording talk and interaction for 1 hr each session. The resulting database included more than 1,300 hr of audio that took an additional 4 years to transcribe, code, and analyze. The logistics and resources necessitated by their methodology, unfortunately, limited their sample size, both in terms of recording length and number of children observed. Their lowest SES group included only six children and their highest SES group 13, reducing the extent to which within-group variation could be assessed. Further, families recorded only 1 hr per month, and this single hour was extrapolated to generate daily count estimates. It is unclear whether a single hour recorded once per month can be considered representative of a child's total language experience. Adding to this concern, their 1-hr, in-person sampling approach was susceptible to intrusive observer effects, which might have elevated counts for some families.

Despite the fact that Hawthorne effects are even more problematic for shorter time samples, 15–30 min recordings are common in laboratory research and clinical settings because they are simply more practical in terms of data collection, transcription, and coding. To minimize travel and provide better control, it is also common for researchers to simulate home environments in a laboratory setting. However, the generalizability of these laboratory situations to naturalistic interactions experienced between parents and children throughout the day is debatable. Thus, although the need for measuring the natural language environment exists, the logistical burdens associated with completing and analyzing naturalistic recordings of sufficient duration have made it impractical until recently to scale or sustain traditional data collection and transcriptional analysis approaches for larger studies or ongoing interventions.

These issues have been addressed by the development of LENA hardware and software technology, facilitating daylong audio recordings in the child's natural language environment and by development of an automated approach to labeling vocalizations occurring in the recordings. Children 2–48 months of age wear a digital recorder in the front

pocket of provided clothing, and then all-day recordings are automatically analyzed by using specially adapted speech recognition technology, generating estimates of adult word exposure (including both child-directed and overheard speech), as well as adult-child conversational turn-taking and child vocalization frequency throughout the day.

This type of automated analysis has been shown to yield useful results and has led to significant research using the methodology (e.g., Oller et al., 2010; Weisleder & Fernald, 2013) to address questions that require large data sets and that would essentially be impossible to conduct with traditional human-coding approaches. We do not, however, underestimate limitations of the automated method. Its algorithms are modeled on human coding, and although the resulting estimates have proven useful, human coding is more reliable and may be preferable for cases in which there are sufficient resources to permit it (see Apparatus section for a more complete description).

Although the system cannot, at present, evaluate aspects of adult speech related either to addressee or to semantic content of words and sentences, it has the potential to augment studies that evaluate such issues and to contribute large-scale data on quantity of vocalization and interaction that could not otherwise be obtained. Over the past several years, the LENA technology has been used to characterize the language environments of children across a diversity of populations. Warren et al. (2010) observed deficits in child vocalization and turn-taking among young children with autism spectrum disorders (ASD), while other research with ASD populations has reported correlations between automated language environment variables and standard language assessments (Dykstra et al., 2012) and shown impressive performance for early attempts at ASD classification by using automated approaches (Oller, et al., 2010). The LENA technology has been used extensively by researchers for the deaf and hard of hearing to assess the early language environments of children with hearing loss (Aragon & Yoshinaga-Itano, 2012; Wiggin, Gabbard, Thompson, Goberis, & Yoshinaga-Itano, 2012) and has been applied in studies investigating the early acoustic environment of premature infants (Caskey, et al., 2011). Also, the system has been used to investigate patterns of talk in other languages and cultures (Pae, 2013; Zhang et al., 2015).

Findings from these and other studies suggest that automated language environment analysis can provide meaningful information about the frequency with which children vocalize, as well as the quantity of adult language input and adult-child interaction that occurs throughout the day. However, most research to date has focused on relatively small, specialized populations and has not been explicitly designed to generate population-level standards. Nevertheless, to compare results across specialized and descriptive intervention studies effectively requires a large reference data set. Thus, the first purpose of the research described here was to characterize the distributions of automatically generated LENA language measures within a large reference

sample and to provide a first-generation standardization for monolingual English-speaking households. The second was to explore relations between LENA automated measures and scores on language assessments administered during professional evaluations and to assess whether there is overlap between information collected on standard language assessments and criterion LENA measures. Third, the current database allowed us to reevaluate the adult word count (AWC) results of Hart and Risley (1992, 1995) with a much larger sample and all-day recordings and to estimate the amount of naturalistic talk and assess its relationship to SES. Our research questions were as follows:

1. What are the overall by-age distributions observed for automated estimates of child vocalization frequency, AWCs, and conversational turn-taking from naturalistic recordings within a large sample of children 2–48 months of age living in monolingual English-speaking households?
2. Do LENA measures correlate with standard language assessments?
3. Do observable differences exist in the automated estimates of language behaviors of families across socioeconomic groups?

Method

Data for the present study were collected over 38 months in two phases between 2006 and 2009. Phase I was conducted over 6 months, and a subsample of Phase I participants continued for an additional 32 months in Phase II. The purpose of the second phase was to conserve resources by collecting additional longitudinal data with a smaller sample. Data from both phases were merged for the analyses presented here, except where noted. While prior studies using the LENA Research Foundation's database have included subsets of the data to be considered here (e.g., Warren et al., 2010), this is the first report that includes the full set of semilongitudinal data covering the age range from 2 to 48 months.

Participants

Phase I participants were 329 monolingual English-speaking families with typically developing children 2–48 months of age. Families were recruited by using direct mail postcards sent to approximately 6,000 households in the Denver metro area that offered families \$75 per monthly, daylong recording session for 6 months. Roughly one third responded and were briefly screened by phone. Parents were excluded who were uninterested after hearing further study details, reported existing language or developmental disabilities for their child, or whose children were outside the recruitment demographics. An effort was made to balance the final sample across child age and gender and to match mother's attained education to that of the (all races combined) U.S. Census distribution (U.S. Census Bureau, 2005). Overall, 364 families who passed the initial

screen received informed consent materials, and 329 families chose to participate and completed Phase I.

Near the end of Phase I, a subset of families ($n = 80$) were invited to participate in Phase II, which was designed to increase the number of recordings contributed by individual children at each age level. This cohort was selected to achieve the same proportional demographic distributions as Phase I and to approximate a normal distribution with respect to children's language development based on standard assessments administered during Phase I. Fifty-nine families (74%) completed Phase II, recording for an additional 32 months. Table 1 summarizes participant family characteristics and participation details for Phases I and II. Although the sample was balanced with respect to mother's attained education, it is a sample of convenience and is not ethnically diverse, though the distribution approximates that of the United States overall at the time (see Table 1; U.S. Census Bureau, 2007). Further, because children were recruited from monolingual English-speaking families, the sample is not necessarily representative of the U.S. immigrant population or other nonnative English-speaking families.

Apparatus

Participants' natural language environments were characterized using the LENA language environment analysis system (Xu et al., 2008a, 2008b). The LENA System includes a digital recorder that children wear in the front pocket of clothing designed to minimize friction and optimize microphone placement. Phase I recorders held a maximum of 12 hr, increasing to 16 hr during Phase II. The resulting audio file is transferred to a computer, where LENA processing algorithms incorporating pattern recognition software and speech signal processing technology parse the audio stream into sound categories (i.e., distinguish human speech activity from other environmental sounds) and thus generate a segment map of the language environment by its different acoustic features. Eight categories are identified: female and male adult, key child (the child wearing the recorder) and other child, overlapping human speech, television or electronic media, noise, and silence. For the current study, only the first three categories were used.

Segmentation reliability compared on 70 hr of human transcription and coding has been assessed and reported previously, with sensitivity and precision for adult and key child segments between 68% to 82% (see Xu, Richards, & Gilkerson, 2014; Xu et al., 2008a, 2008b; F. J. Zimmerman, et al., 2009). Discrepancies derived primarily from cases in which the system identified human-coded adult or child segments as overlapping speech or other noises and sounds that are eliminated from analysis and not included in the automated adult word and conversational turn (CT) counts. Only 2% of human-coded adult speech segments were mislabeled by the system as child, and only 7% of human-identified child vocalization segments were misidentified as originating from an adult speaker. In a similar way, only 4%

Table 1. Family, child, and recording characteristics for Phase I and Phase II samples.

Category	Phase I		Phase II	
	<i>n</i>	%	<i>n</i>	%
Mother's education ^a				
<High school (22%)	45	13.7	12	15.0
High school or GED ^b (26%)	108	32.8	21	26.2
Some college (29%)	92	28.0	29	36.3
Bachelor's degree or higher (23%)	84	25.5	18	23.5
Family ethnicity ^a				
Caucasian (66%)	260	79.0	62	77.5
Latino or Hispanic (14%)	28	8.5	7	8.7
African American (13%)	14	4.3	5	6.2
Native American (1%)	4	1.2	1	1.3
Asian (4%)	2	0.6	0	0.0
Other (2%)	14	4.3	4	5.0
Not specified	7	2.1	1	1.3
Child gender				
Female	162	49.2	43	53.8
Male	167	50.8	37	46.2
Child age ^c				
2–12	115	35.0	30	37.5
13–24	103	31.3	23	28.8
25–36	96	29.2	18	22.5
36–48	15	4.5	9	11.2
Child language percentile ^d (%)				
<2	5	1.5	2	2.5
2–16	13	4.0	4	5.0
17–50	87	26.4	33	41.3
51–84	152	46.2	31	38.8
85–98	48	14.6	8	10.0
>98	4	1.2	1	1.2
Not specified	20	6.1	1	1.2
Total	329		80	
Audio recordings				
Contributed	1,772	100	1,843	100
Excluded	271	15	131	7
Included	1,501	85	1,712	93

^aPercentages reflect population distributions for sample based on 2007 census. ^bGED - General Educational Development. ^cChild age in months at the first recording. ^dLanguage percentiles based on composite *Preschool Language Scale–Fourth Edition* and *Receptive Expressive Emergent Language Test–Third Edition* total language scores.

of segments identified by LENA as child vocalizations were actually adult speech according to transcribers, and 3% of segments labeled as adult were really child.

To produce estimates of adult speech, clear adult segments are parsed to distinguish consonant from vowel sounds, from which approximate word counts are generated by using a previously validated regression model. The resulting AWC estimates were designed to be unbiased and in 70 hr of transcribed recording data were, on average, only 2% lower than human estimates (Xu et al., 2008a, 2008b). For child vocalization estimates, regions of potential speech-related vocalization in key child segments are distinguished from nonspeech sounds (which are not meant to be included in any of the LENA measures) and counted. Age-based modeling by using independent transcriptions of child vocalizations across a range of ages facilitated the identification of features common to speech-related

output and their differentiation from nonspeech sounds. More specifically, *child vocalizations* as identified by the LENA System include any articulations that originate from the vocal tract of the child, except for fixed signals (instinctive reactions, including screams, cries, etc.) and sounds related to respiration (e.g., breaths) or digestion (e.g., burps). Thus, the system is designed to filter fixed signals, as well as respiratory and digestive sounds, from LENA counts. What then remain are potentially communicative speech and speechlike sounds (see Oller, 2000) operationally defined as breath groups of vocalizations separated by 300 ms of silence or other intervening sounds. The value of 300 ms was chosen because it represents the high end of the distribution for silence or low-energy periods that occur within utterances (Oller & Lynch, 1992; Rochester, 1973). This automated approach to defining vocalizations necessarily contrasts with traditional methods for identifying utterances based on semantic content (e.g., Brown, 1973), which may permit long breaks within a single sentence or multiple sentences without pauses in between. Thus, child vocalizations as operationally defined by the LENA System cannot be directly compared with traditional utterance counts derived from semantic analysis and human judgment about what constitutes a phrase or full proposition. However, the consistency inherent to the automated method allows for reliable comparisons between recording samples, which we argue has its own utility. For more detail on LENA System processing, see Oller et al. (2010) and Xu et al. (2008a, 2008b).

Measures

The LENA technology, as described, analyzes automatically identified adult male and adult female segments and key child voice segments to generate estimates of the three environmental behavioral measures used in the current study: (a) AWCs, (b) child vocalization frequency, and (c) conversational turn-taking counts. AWC is an estimate of the number of adult words spoken loudly enough to register clearly in the LENA recorder, but it does not differentiate child-directed speech from overheard speech. In practice, we estimate these words typically occur within a 10-ft radius of the child wearing the recorder. The child vocalization count (CVC) reflects the number of speech-related vocalizations produced by the child as identified by the automated procedure. For example, because child vocalizations are separated by 300 ms of silence and may be of varying length, the vowel “a” spoken in isolation or the babble “babababa,” with no pauses or breaks, would both be assigned a count of one vocalization. Likewise, the string “mommy I want a cookie” would also count as one vocalization, provided that there was not more than 300 ms of silence between words or syllables. CT counts are the number of alternations within a conversation between clear, speech-related adult and key child vocalizations, as labeled by the automated procedure. A conversation was defined as a sequence of vocalizations bounded by at least 5 s of nonvocal material, based, in part, on rules suggested by Hart and Risley

(1995). In this formulation, either child or adult may initiate a turn, and responses may not serve as the initiation of a subsequent turn. Thus, both of the sequences child–adult and child–adult–child are counted as one and only one turn. Note that conversations can sometimes be identified in error because some alternating adult and child utterances can occur in the absence of conversation, as when a parent is talking on the phone while holding a child who may also be vocalizing. Turns may include protophones, babbles, or words produced by the child. For example, if a 2-month-old infant says “aa” or a 9-month-old says “bababa,” and the parent responds within 5 s, this would count as a turn. If the parent or child interrupts the initiator, as is often the case in spontaneous speech interactions, the system will identify that section as overlapping speech, but the vocalization segment immediately following the overlap segment will be coded as a turn (given that the overlap section is not more than 5 s).

In addition to the automated language environment measures, a variety of developmental language assessments were obtained over Phases I and II. These include both self-report questionnaires completed by parents at home and standard assessments administered by a certified speech-language pathologist (SLP) in a dedicated evaluation room at the LENA Research Foundation. The current study analyzed only those assessments collected on the majority of Phase I or Phase II participants.

Parent Questionnaires

To measure vocabulary development during Phase I, parents of children ages 8–30 months were asked to complete (depending on age) either the Words and Gestures or Words and Sentences version of the MacArthur-Bates Communicative Development Inventory (MB-CDI) every other month (Fenson et al., 2007). Verbal production vocabulary scores were included here. To obtain a broader range of developmental information, the Child Development Inventory was incorporated during the first year of Phase II and completed at approximately 4-month intervals (Ireton, 1992). The Child Development Inventory is a 300-item questionnaire covering a number of developmental domains, including expressive and receptive language subscales. It was included in the test battery because the expressive language component contains questions about grammatical complexity and sentence structure, thus providing a supplement to the vocabulary information collected from the MB-CDI. For the present study, we report on the expressive language scores from the first administration of the Child Development Inventory, completed by parents on average 4.0 months ($SD = 0.5$) after their final Phase I recording. Parents were sent no more than one of these questionnaires per recording month.

Professional Language Evaluations

During Phase I, parents were asked to bring their children to the research center to be evaluated by a certified SLP. Approximately half of participating families were randomly selected to visit every 2 months, and the others

were seen once over the 6-month period. (The assessment procedure for Phase II participants was similar, except that evaluations were completed only at study months 6, 12, 20, and 30) The number and type of standard assessments administered during each session varied with child age and attention span and with time constraints. The core battery that participants completed typically included the Receptive Expressive Emergent Language Test–Third Edition (REEL-3; Bzoch, League, & Brown, 2003), the Preschool Language Scale–Fourth Edition (PLS-4; I. L. Zimmerman, Steiner, & Pond, 2002), and the Cognitive Adaptive Test and Clinical Linguistic and Auditory Milestone Scale (CAT/CLAMS; Accardo & Capute, 2005). The REEL-3 is an observational assessment of language development for children birth through 36 months of age. Like the REEL-3, the Preschool Language Scale–Fourth Edition is a widely used language development assessment but covers a broader age range (birth–83 months). The CAT/CLAMS is an observational measure of language (CLAMS) and cognitive (CAT) development for children up to 36 months of age. This relatively short assessment (6–20 min) is commonly used by developmental pediatricians and was included because it provides a subscale specific to cognitive development. Parents did not receive feedback on assessment results but were compensated \$100 for time and travel to assessment sessions.

Procedure

Phase I recording data collection protocols were maintained across Phase II. Participant families completed one full-day audio recording each month, with approximately 4 weeks intervening between recording sessions. To simplify parent instruction and ensure a representative distribution of recording data across days of the week, participants were randomly assigned a recording day number (e.g., the 10th of the month) and asked to record on that specific day each month. Recording materials were sent to participants' homes to arrive at least 1 day before their scheduled recording sessions. Recording packets included step-by-step mostly picture instructions, clothing, a recorder, a prepaid return envelope, and a session-specific recording log asking detailed information about the day (e.g., Was the child sick? Did the child attend preschool? What time was the TV on?).

Parents were instructed to activate the recorder when the child first awoke, secure it in the clothing pocket, dress the child, and then go about their usual daily routines. The recorder was attached to the child and so measured the natural language environment and experience of the child wherever he or she was continuously throughout the day. As such, the recordings could occur in a variety of settings from morning until bedtime (e.g., home, child-care, library, playground, restaurant, grocery store). The recorder was programmed to shut off automatically when it reached full capacity and remained on the child at all times with the exception of car rides, naps, and baths. When the child was not wearing the recorder, parents were asked to keep it nearby and to continue recording. At the end of a recording day, parents completed session logs and packed

all materials into the included postage-paid envelope, which was retrieved from their home the following day by courier. Although the recorder could be turned off manually by parents, to simplify procedures, they were simply told to turn it on at the beginning of the day and let it turn off by itself after it had reached recording capacity. If the parents returned a recorder that had been paused or contained breaks, they were compensated anyway. If parents turned in a recording that was shorter than 12 hr, they were asked to rerecord. All parents were told that if they were not comfortable with sharing private information that was recorded during the session, we would delete the audio file, and they could rerecord on a different day. During the course of the study, parents were not provided recording-based feedback. Phase I participants recorded once per month for 6 months between January and June 2006, and Phase II participants (a subset of Phase I) continued to complete monthly recordings for an additional 32 months. Thus, Phase II participant children who were older at initial recruitment were up to 68 months of age at the end of Phase II. Because the LENA System technology has been validated only up to 48 months, recordings for older participants were not included on the current analyses and the recording age range is limited to children between 2 and 48 months of age. During Phase I, families contributed, on average, 4.6 recordings ($SD = 1.4$; range 1–7), and Phase II participants, on average, provided an additional 21.4 recordings ($SD = 9.1$; range 3–32). Combining phases, the average recording contribution was 9.8 ($SD = 10.2$; range 1–39) recordings per family.

Participants were compensated \$75 per recording (roughly \$6/hr for daylong recordings). To encourage retention and facilitate communication, parents were provided a \$200 bonus for completing the study, with potential deductions to the bonus amount if commitments were not met. Although parents were told that deductions would be applied each time they failed to communicate about canceling or rescheduling sessions or did not return equipment at the scheduled times, such deductions were rarely applied, except in the case of repeated instances.

Statistical Approach

Analyses of the collected recording data were conducted by using Pearson correlation, *t* test, and analysis of variance (ANOVA) with contrasts. For many analyses of measures presented here, child age was an important covariate, and statistical analyses were correspondingly adjusted either by controlling for age or by conducting analyses on age-standardized measures. Our analytic approach was selected to address three goals. The first was to characterize the distributions of LENA measures across the full range of child ages. Given that participant families varied considerably in their recording contributions and to maximize the sampling at a given age, these data were initially summarized cross-sectionally within each month of child age. Analyses of change across child ages were conducted at the level of the family, either by sampling from within their available recordings or by averaging across them. In

addition, we examined child-related measures (CVC and CT) in empirically motivated age groupings. This family-level approach was intended also to reduce variance in and confer a greater degree of stability to LENA values within families. Test-retest reliability and variability on LENA measures within families was assessed by correlational comparisons across both shorter and longer time spans. Toward the second goal, LENA measures were validated via correlation against a variety of criterion assessment measures of each child's language development. Third, LENA measures were compared across families from different socioeconomic groups (based on mother's attained education) and also within groups. We note that the scale of audio sampling of natural language environments made possible through the use of the automated approach used here, especially when combined with a longitudinal data collection design, provides opportunities to apply statistical methods (e.g., multilevel modeling) that may offer greater interpretive depth and fewer limitations compared with those used in much prior research. Nevertheless, consistent with our goals for this presentation of research that incorporates a complex and still somewhat novel methodology, we opted to characterize results with standard and easily interpreted statistical analyses.

Results

A total of 49,765 hr of data were collected over 3,615 daylong recording sessions from 329 families during Phases I and II (see Table 1). Although during some stages of data collection the recorders held up to 16 hr of audio, for the current analysis, recordings were truncated at exactly 12 elapsed hours, which defines the recording day for the present report. Parents made use of their ability to pause recordings rarely, in only 90 recordings overall (2.5%), and contiguous 12-hr sections were included for 58 of these cases. Across both phases, a total of 402 recordings were excluded. The majority of these ($n = 273$, 68%) were dropped due to contiguous recording sections being <12 hr in duration. Recordings affected by mechanical problems or parent error (e.g., not adhering to recording protocols) or for other reasons yielding extreme atypical values on one or more language environment measures were excluded as well ($n = 53$, 13%). During Phase II, a subset of recordings demonstrated significantly elevated friction noise due to a clothing manufacturing error. Recordings completed using this batch of clothing with language environment measures that deviated significantly from those for other recordings within the same family ($n = 77$, 19%) were excluded. Table 1 summarizes recording exclusions by phase of study. The final data set for the current study contains 3,213 12-hr recording sessions (38,556 hr) from 329 families.

Distributions of LENA Measures Across Child Age

Descriptive statistics for the distributions of the three primary LENA measures (AWC, CVC, and CT) across

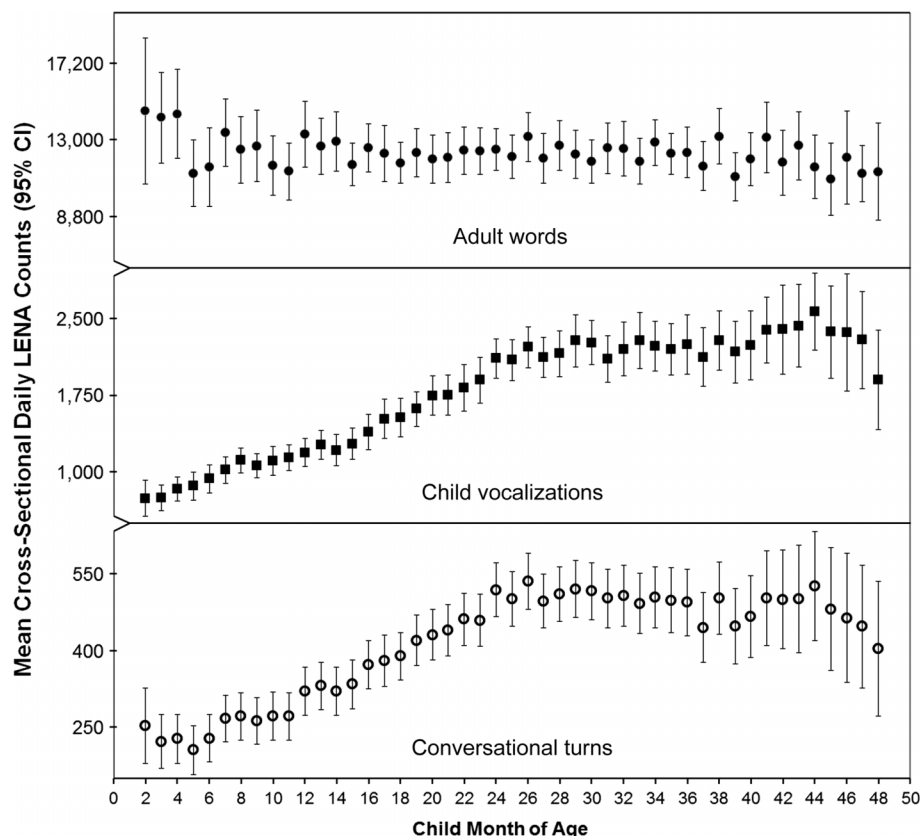
Table 2. Daily LENA measures: distribution by child age.

Month of age	<i>n</i>	Adult words		Child vocalizations		Conversational turns	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
2	16	15,439	8,234	776	324	271	142
3	34	14,993	6,481	787	348	233	95
4	42	15,315	7,249	874	375	239	106
5	45	11,872	6,084	913	413	221	118
6	50	12,503	7,336	987	478	242	115
7	56	14,148	6,390	1,067	469	283	138
8	57	13,248	6,427	1,158	440	287	136
9	60	13,578	7,226	1,108	452	276	126
10	60	12,301	5,813	1,168	516	289	144
11	64	11,986	5,895	1,187	481	285	124
12	65	14,136	6,880	1,245	477	341	170
13	71	13,351	6,296	1,326	578	355	193
14	71	13,695	6,497	1,285	605	347	202
15	73	12,122	4,575	1,347	612	355	175
16	78	13,167	5,701	1,481	755	404	233
17	74	12,991	6,456	1,614	758	412	236
18	87	12,262	5,053	1,634	824	423	250
19	81	12,923	5,803	1,696	758	451	261
20	83	12,592	5,574	1,841	830	468	263
21	83	12,722	5,928	1,846	838	480	281
22	83	13,010	5,605	1,950	990	502	285
23	88	13,038	5,700	2,021	962	497	274
24	88	12,977	4,953	2,197	839	547	243
25	87	12,622	5,117	2,189	872	533	255
26	84	13,800	5,735	2,305	848	568	265
27	95	12,750	6,581	2,213	867	530	269
28	96	13,425	6,092	2,277	987	545	268
29	91	12,910	5,690	2,419	1,113	556	268
30	92	12,405	5,470	2,367	963	552	280
31	92	13,291	6,240	2,223	971	545	298
32	87	13,310	6,530	2,347	1,072	555	298
33	91	12,469	5,591	2,444	1,205	531	281
34	91	13,453	5,707	2,362	1,042	540	267
35	84	12,729	4,928	2,325	1,039	533	275
36	82	12,990	6,025	2,401	1,170	533	274
37	76	12,197	5,466	2,274	1,078	478	245
38	74	13,893	6,373	2,430	1,137	533	246
39	70	11,574	5,199	2,323	1,134	487	273
40	61	12,542	5,333	2,400	1,246	509	299
41	47	13,851	6,013	2,500	975	536	265
42	44	12,382	6,101	2,569	1,363	546	371
43	38	13,239	5,546	2,563	1,107	538	290
44	38	12,060	5,351	2,683	1,030	561	271
45	31	11,437	5,084	2,506	1,250	523	328
46	29	12,822	6,617	2,563	1,550	504	298
47	33	11,531	4,102	2,454	1,198	473	211
48	28	12,128	6,837	2,071	1,147	449	306

Note. LENA = Language Environment Analysis.

child age at recording are presented in Table 2. Although, in general, different families contributed data from varying numbers of recordings across different ranges of child age, each family is represented only once in any one age month. When more than one recording was completed in the same age month, LENA measures for that family were averaged within that month. Note that the sample sizes for the months between 41 and 48 months were typically lower than those occurring at the younger ages. Figure 1 depicts these cross-sectional monthly averages across the range of child ages.

Figure 1. Daily Language Environment Analysis (LENA) measures by child age. Values shown are daily (12-hr) mean counts with 95% confidence interval (CI), computed cross-sectionally within child month of age.



The average within-family relationship between child recording age and daily AWC (averaged to each family's mean AWC and child age) was not significant, $r(327) = .05$, $p = .33$. To examine this relationship further, we next selected values from the first and last recording for each family (i.e., sampling each family at the child's minimum and maximum recording ages). No significant correlation between AWC and age was observed for either the first, $r(327) = .04$, $p = .51$, or last recording, $r(327) = .08$, $p = .17$. Consistent with these results and as depicted in Figure 1, cross-sectional daily AWC means were relatively stable across child recording ages, at least after 5 months. However, average daily AWCs for families with children age months 2–4 ($M = 15,071$, $SD = 5,864$) were 19% higher on average than for those with older children ($M = 12,622$, $SD = 4,281$), $t(327) = 3.45$, $p = .001$. A subset of 42 families contributed recordings that spanned a range younger and older than 4–5 months. On average for these families, mean daily AWCs for the younger ages ($M = 15,435$, $SD = 6,080$) were significantly higher by roughly 24% than at later ages ($M = 12,459$, $SD = 4,649$), $t(41) = 3.78$, $p = .001$.

In contrast to AWC, on average, child age increased with both mean daily CVC, $r(327) = .60$, $p < .001$ and mean daily CT, $r(327) = .46$, $p < .001$ (each again averaged

within family). However, the cross-sectional means (see Figure 1) suggested that the relationship of both measures to age attenuated to some degree at approximately 26 months, after which the rates of increase slowed. Thus, we conducted additional analyses splitting the sample at that point. Using the first recording for each family, in the 2–26 months range, child age correlated strongly with daily CVC, $r(231) = .66$, $p < .001$, and with daily CT, $r(231) = .59$, $p < .001$. But for older children, the relationship was reduced to non-significance for both measures, CVC $r(94) = .04$, $p = .67$, CT $r(94) = -.06$, $p = .59$. Repeating these comparisons using the last recording for each family achieved more balanced sample sizes but with similar outcomes. In the younger sample, child age again correlated with daily CVC, $r(143) = .55$, $p < .001$, and with daily CT, $r(143) = .48$, $p < .001$, and in the older sample, neither CVC, $r(182) = .09$, $p = .22$, nor CT, $r(182) = .004$, $p = .96$, correlated with age.

To further illustrate this age-related change, we averaged recordings within family while maintaining the split between age groupings. Recordings for families spanning the age split threshold were included only in the age group in which they were more frequent. For families in the 2–26 months range, mean child age correlated highly with mean daily CT, $r(190) = .57$, and CVC, $r(190) = .65$, both

$p < .001$. For those in the 27–48 month range, correlations with age were not significant for either CT, $r(135) = -.10$, $p = .24$, or CVC, $r(135) = .05$, $p = .57$. In summary, up to 26 months estimated CVCs increased an average of 65.3 ($SE = 5.5$) per month of age, after which the rate of increase slowed to a nonsignificant additional 9.0 ($SE = 15.6$) vocalizations per month. CT estimates grew at an average rate of 17.2 ($SE = 1.8$) per month for the younger period, after which the rate dropped to a nonsignificant decrease of -4.9 ($SE = 4.2$) turns per month at older ages.

Patterns of Variability Within Families

The within-family stability of LENA measures across multiple monthly recordings contributed during Phase I is summarized in Table 3. We compared count values collected from 1 to 4 months apart. Because not all participants contributed usable recordings each month, averaged Pearson correlations were weighted by sample size to compare counts between recordings over different time spans. Bonferroni correction values for alpha based on the number of comparisons of each time span are shown in Table 3. Correlations between recordings collected from 4 to 16 weeks apart decrease with time but indicate that caregiver and child vocal behavior as characterized by the automated system is moderately consistent across these time intervals. Although AWCs correlate at lower levels than the other two measures, suggesting greater month-to-month variability, paired sample t tests comparing mean count values for all contiguous months (e.g., 1 vs. 2, 2 vs. 3) revealed only one significant month-to-month difference (unadjusted for multiple testing and out of five possible comparisons) between months 4 and 5, mean difference = 1,657 ($SD = 5,946$), $t(244) = 4.36$, $p < .001$.

To examine daily variation within families on LENA metrics, during Phase I a subset of 52 participants was randomly selected to record a second time in addition to their once monthly schedule on the day following a regular recording. As shown in Table 3, AWCs were positively correlated to a stronger degree than for recordings one or more months apart. However, average daily AWCs varied approximately 12% between the first ($M = 13,626$,

$SD = 6,494$) and second ($M = 12,006$, $SD = 4,575$) recording days, $t(51) = 2.38$, $p = .02$. CVCs for successive days correlated similarly to those from recordings taken 1–3 months apart and, on average, differed only marginally ($M_1 = 1,557$, $SD_1 = 807$; $M_2 = 1,723$, $SD_2 = 828$) between days, $t(51) = 1.80$, $p = .08$. Average daily turn counts for the 2 days ($M_1 = 382$, $SD_1 = 259$; $M_2 = 387$, $SD_2 = 206$) differed only by 1%, $t(51) = 0.14$, $p = .89$, though the correlation between days was slightly lower than for recordings collected over longer spans. Analyses conducted controlling for child age yielded similar results. In summary, these results indicate that LENA measures on average are stable to a moderately high degree over time, but individual families may demonstrate day-to-day variability in their language behavior. We did not examine the impact of situational factors that would be likely to have an effect.

Criterion Validity of Automated Measures

Criterion validity of automated language estimates was examined by comparing the three daily totals to scores from SLP-administered assessments and parent report questionnaires. Although these assessments are broader in scope (e.g., incorporating behavioral and semantic content) compared with the vocalization frequency-based automated measures, some overlap between volubility and expressive language development could be expected. For these comparisons, scores from all available assessments and recordings contributed for Phase I were averaged within family. Correlations between automated and other measures (controlling for child age) are provided in Table 4. Child vocalizations predicted 7%–16% and CT frequencies 9%–14% of the variance in language and cognitive scores. AWC effects were somewhat lower, contributing up to 8%. Given that vocabulary growth may be nonlinear across the age range under examination, an additional analysis was conducted on MB-CDI vocabulary scores for children at an average age of 15 months or older. This age threshold was selected empirically at the point at which average vocabulary increased sharply, from a mean of 11.4 ($SD = 14.5$) words at 14 months to 77.6 ($SD = 40.5$) at 15 months. For the 55 children ages 5–14 months, no significant age-controlled

Table 3. Test-retest reliability of LENA measures: average correlations across recordings.

Weeks between recordings	No. of comparisons (α)	Sample size	Child vocalizations		Conversational turns		Adult word counts	
			r_{Mean}	95% CI	r_{Mean}	95% CI	r_{Mean}	95% CI
4	5 (.01)	92–258	.69	[0.61, 0.76]	.71	[0.64, 0.77]	.44	[0.32, 0.54]
8	4 (.013)	110–248	.64	[0.56, 0.72]	.66	[0.58, 0.74]	.40	[0.28, 0.51]
12	3 (.017)	110–236	.67	[0.58, 0.74]	.66	[0.57, 0.74]	.40	[0.27, 0.52]
16	2 (.025)	101–185	.60	[0.48, 0.69]	.64	[0.53, 0.73]	.42	[0.27, 0.54]
0	1 (.05)	52	.67	[0.48, 0.80]	.56	[0.34, 0.72]	.66	[0.47, 0.79]

Note. LENA = Language Environment Analysis; CI = confidence interval. For comparisons made on monthly recordings, up to six recordings were included for each participant, collected at 4-week intervals. Sample size and average correlation (r_{Mean}) are based on the number of possible comparisons for a given interval duration between recordings (e.g., 4 weeks or 8 weeks). For the within-week comparison, participants contributed two recordings collected on contiguous days. Correlations provided are averages for all available comparisons weighted by sample sizes. Comparison alpha based on Bonferroni corrected value. All Pearson correlations $p < .001$.

Table 4. Concurrent validity of LENA measures: correlations with criterion measures.

Measure	Child vocalizations			Conversational turn counts		Adult word counts	
	<i>n</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
SLP-administered							
PLS-4 ^a	306	.32	<.001	.30	<.001	.21	<.001
REEL-3 ^a	265	.38	<.001	.35	<.001	.22	<.001
CLAMS ^b	258	.37	<.001	.34	<.001	.21	.001
CAT ^b	258	.27	<.001	.30	<.001	.18	.005
Parental report							
MB-CDI ^c	182	.35	<.001	.35	<.001	.28	<.001
Child Development Inventory ^d	203	.40	<.001	.38	<.001	.16	.03

Note. LENA = Language Environment Analysis; PLS-4 = Preschool Language Scale–Fourth Edition; REEL-3 = Receptive Expressive Emergent Language Test–Third Edition; CLAMS = Clinical Linguistic and Auditory Milestone Scale; CAT = Cognitive Adaptive Test; MB-CDI = MacArthur–Bates Communicative Development Inventory. All correlations controlled for age.

^aExpressive language scale standard score. ^bDevelopmental quotient. ^cVerbal production score. ^dExpressive language score.

correlations between vocabulary score and LENA measures were found. However, for the 126 children aged on average between 15 and 31 months, vocabulary size controlled for age correlated with CVC, $r(123) = .38$, CT, $r(123) = .39$, and AWC, $r(123) = .34$, all $p < .001$.

Distributions of LENA Measures Across Demographic Groups

Overall, no significant differences on LENA measures were observed related to child gender. Averaging values within family over a mean of 9–10 recordings, at an average age of 23 months for both genders, neither AWC, $t(327) = 0.74$, $p = .46$, CT, $t(327) = 0.48$, $p = .63$, nor CVC, $t(327) = 0.46$, $p = .65$, evidenced any significant differences between the two groups. Table 5 provides basic statistics for LENA measures for families with boys versus girls, as well as across SES groups.

The four SES groupings used here were based on mother's attained level of education. Patterns of comparison varied from measure to measure. Figure 2 illustrates

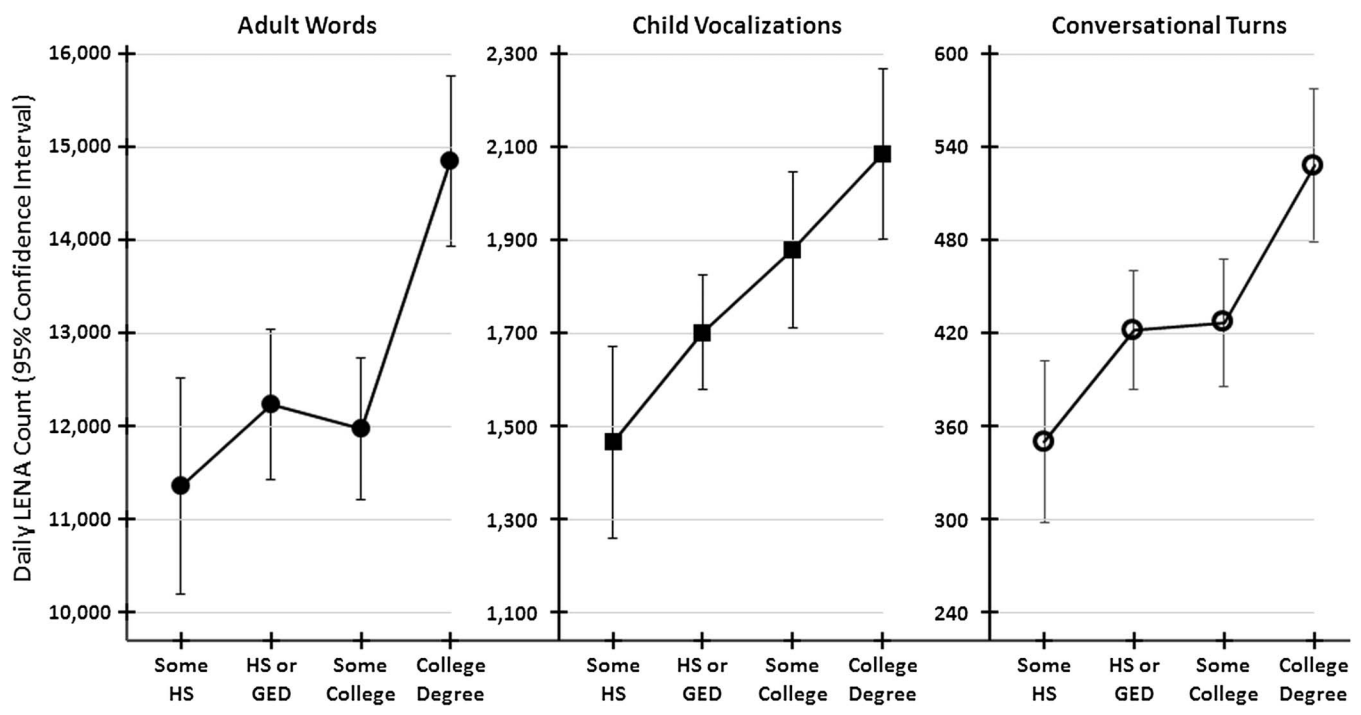
count patterns across SES groups. Regarding AWC, in a one-way ANOVA, the group with the most advanced education (at least a bachelor's degree) had average daily word counts significantly higher (24%) than those of the three groups with less education (combined $M = 11,975$, $SD = 4,017$), contrast $t(325) = 5.67$, $p < .001$. Those three groups did not differ significantly from each other (mean AWC differences between groups 2% to 8%). An alternate analysis using polynomial contrasts over SES groups demonstrated significant linear, quadratic, and cubic effects, as could be expected given the observed lack of differentiation among the first three groups. We further investigated whether the similarity among these groups with respect to AWC could be attributed to differential effects of household size. The number of adults living in the house at the time of recording was available for 266 families (81%). Of these, 25 (9%) described themselves as single-parent households, 195 (73%) reported dual-parent households, and 46 (17%) indicated three or more adults (anyone past puberty) living in the residence. A small correlation was found overall between AWC and number of

Table 5. Daily LENA measures: distribution by gender and maternal education.

Maternal education	Child gender	<i>n</i>	Month of age	Adult words		Child vocalizations		Conversational turns	
			<i>M</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
All combined	Female	162	23	12,895	4,103	1,795	747	434	202
	Male	167	23	12,528	4,439	1,839	826	447	226
Some high school		45	21	11,358	3,987	1,466	704	350	179
	Female	23	21	11,690	4,150	1,389	646	346	149
	Male	22	20	11,012	3,875	1,546	767	354	209
High school or GED		108	23	12,239	4,264	1,702	649	422	205
	Female	50	22	12,391	4,021	1,705	637	411	194
	Male	58	23	12,107	4,494	1,699	665	431	215
Some college		92	23	11,969	3,731	1,879	825	427	201
	Female	47	23	12,355	3,741	1,773	717	413	175
	Male	45	23	11,565	3,720	1,989	919	442	227
College degree		84	24	14,848	4,307	2,085	860	528	230
	Female	42	25	14,759	4,101	2,146	824	533	231
	Male	42	23	14,937	4,552	2,024	899	524	231
Total		329	23	12,709	4,274	1,817	787	441	214

Note. LENA = Language Environment Analysis; GED = General Educational Development.

Figure 2. Daily Language Environment Analysis (LENA) measures by mother's attained education (high school [HS]; General Educational Development [GED]). Values shown are daily (12-hr) mean counts with 95% confidence intervals averaged within participant families and socioeconomic strata.



adults, $r(264) = .12$, $p = .04$, but no significant differences were found between single parent versus other households within SES groups. Controlling for number of adults in the ANOVA also did not significantly affect the overall or group contrasts. Restricting the sample to the majority dual parent families did not substantially alter the prior results. No significant differences on AWC were found among the lower three SES groups, and AWCs for all three were significantly lower than those of the college graduate group.

With respect to child vocalizations, a linear increase was observed by education group, with children of the most educated mothers producing the most vocalizations, linear contrast $t(325) = 4.65$, $p < .001$. On average, these children produced 42% more vocalizations than children of the least educated mothers. No significant higher order polynomial effects were found. The pattern for CTs was linear by SES group as well but with little difference between the two middle groups, those with a high school degree or GED and those with some college, linear $t(325) = 4.56$, $p < .001$. Children in the most educated group typically engaged in over 50% more turns than those in the lowest group. As for vocalizations, no significant higher order effects were found for CTs.

Note that although the average recording age for children in the lowest SES group was 2–3 months younger than that of children in the other groups, these differences did not reach statistical significance. The most disparate group difference was between the lowest SES group (age

$M = 21.0$, $SD = 11.4$ months) and the highest SES group (age $M = 24.3$, $SD = 11.7$ months), $t(127) = 1.57$, $p = .12$. As well, we repeated the SES group comparisons reported here controlling for child age to similar results. Age was not a significant predictor in the overall model, and age-adjusted marginal means differed from those presented in Table 5 by $<1\%$.

Patterns of Variability Within SES Groups

To further highlight the relevance of the variability in counts observed within education groups, we conducted an additional analysis of CT comparing SES extremes, those with a college degree versus those who did not complete high school. For this illustration, we limited each family to their first six recordings (i.e., covering Phase I) and excluded families with fewer than three recordings (total $N = 302$) and applied a median split to age-standardized CT values within each group. We compared the resulting upper half of the lower education group ($N = 19$) with the lower half of the upper education group ($N = 39$). Parents and caretakers in the upper half of the least educated group on average engaged with their children significantly more ($M = 451$, $SD = 197$) than the parents in the lower half of the most educated group ($M = 334$, $SD = 124$), $t(56) = 2.76$, $p = .008$. Although the two groups did not significantly differ on child age ($M_{\text{Low Ed}} = 17.6$, $SD = 12.7$ months; $M_{\text{High Ed}} = 19.9$, $SD = 11.9$ months), $t(56) = 0.68$, $p = .50$, we also compared these same groups

by using age-standardized CT values to equivalent results. Analyses of SES median groups selected on AWC, $t(56) = 5.24, p < .001$, and CVC, $t(56) = 2.43, p = .02$, yielded patterns of results similar to those for CT. Figure 3 plots age-adjusted AWC and CT for individual families in the two groups.

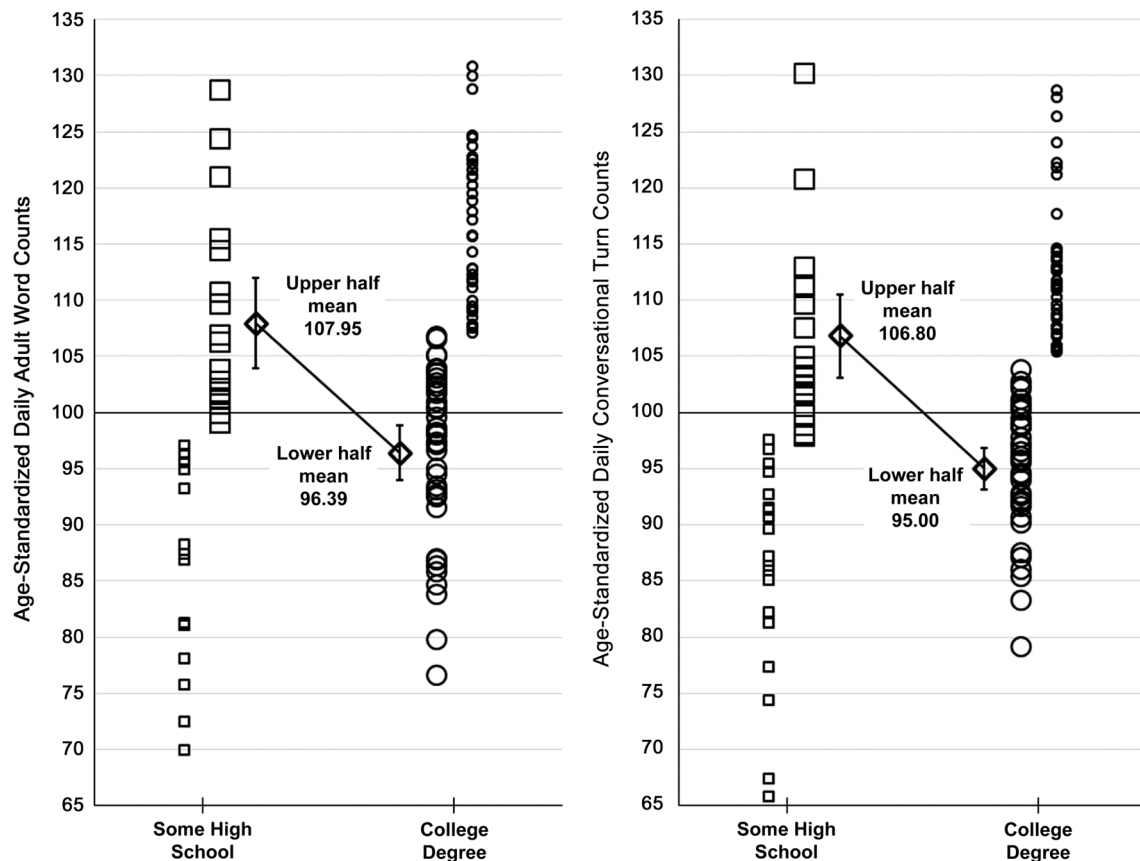
Discussion

The purpose of this research was to characterize the language environment of infants and toddlers with respect to three behaviors sampled continuously throughout the day: adult word production, adult-child conversational turn-taking, and child vocalization frequency. These results represent a new quantitative mapping of the language environment experienced by a reference sample of typically developing monolingual English-speaking American children, provide information about the criterion validity of LENA automated measures compared with child language assessment scores, and describe observed variability related to child age and SES in this first-generation reference sample.

Age-Related Developmental Changes and Variability

Age-related patterns in the early environment varied, depending on the language behavior under investigation. There was no relationship between amount of adult word exposure and child age after 5 months of age, and substantial variability was observed at all ages. The children in our sample were exposed to approximately 12,300 adult words over the course of a 12-hr day, with standard deviations ranging from 4,575 to nearly 7,000 words. Caregivers talked significantly more around 2- to 4-month-olds compared with children at the older ages, which might be attributed to the relative lack of mobility at early infancy coupled with more close face-to-face interactions in the earliest months of life (e.g., during feeding). This finding is in line with research in the area of early vocal development, showing differences in vocalizations directed toward infants before and after 6 months of age, suggesting that maturational changes in early vocal babbling behavior can influence the amount and type of caregiver vocal output at different stages of infancy (Snow, 1977). Consistent with

Figure 3. Median group comparison of daily AWC and CT for least versus most educated parents. Values shown are age-standardized (to $M = 100, SD = 15$), daily (12-hr) adult word and conversational turn counts averaged within families to illustrate the full range of values within socioeconomic groups. Parents in the upper half of the least educated group demonstrate significantly more talk and engagement with their children than parents in the lower half of the most highly educated group.



previous research (Hart & Risley, 1995), the number of words spoken to or near toddlers generally did not increase with age.

Child vocalization frequency according to automated analysis increased steadily (60 more each month on average) up to approximately 26 months. This finding could have been influenced to some extent by reduction in sleep time across the age span, providing more opportunity for vocal output at later ages. It should also be noted that age-specific modeling contributed to identification of child vocal segments, so there is a possibility that measurement differences between ages could be influenced to some degree by differential modeling techniques. After the early linear increase, CVCs began to plateau around 26 months, with children on average adding only nine more vocalizations per month thereafter. Given processing limitations placed on communicative interactions (e.g., sentences cannot be infinitely long), it makes sense that the amount of individual vocal output should plateau at some point, though why this occurred at the observed age is a matter for further investigation. Vocalization frequency also was found to be more variable over time. Reduced variability at the younger ages may be reflective of a stronger biological or maturational influence earlier on, and increased variability for older children could be attributed to the cumulative effects of environmental influences over time.

Similar patterns were observed for conversational turn-taking. Infants engaged in 200–300 back-and-forth interactions with adults in their environment per day during the first year of life, with CTs increasing steadily thereafter (on average around 15 more turns added per month), plateauing around 500 turns per day by the second birthday. Although the sheer number of words spoken by caregivers each day generally did not increase with age, the frequency with which parents interacted in terms of conversational turn-taking did. This result is not surprising if we consider children to be active participants in the social dance of conversation (Jaffe et al., 2001; Stern, 1974). We have seen that as children age and their spoken language becomes more sophisticated, they vocalize more. The associated increase in CTs suggests that as children are able to express more elaborate concepts and needs more often, they elicit more back and forth interaction with their caregivers. Given the well-documented importance of interaction for early development (Landry, et al., 2006; Tamis-LeMonda, et al., 2001; F. J. Zimmerman, et al., 2009), a relative deficit in turn-taking could be an important flag for clinical referral.

Correlations With Standard Language Assessments

Comparisons with standard child language assessments allowed us to establish criterion validity for LENA measures, in a first attempt to determine whether these measures can provide useful clinical information about a child's language skills. At the most basic level, the LENA System simply provides macro-level statistics about the frequency with which a child vocalizes and engages in

interaction throughout the day. Although the measures reported here do not evaluate the sophistication or complexity of the child's vocal output, we conjecture that gross quantitative measures of daylong activity could offer useful if more general information about a child's overall language ability or developmental trajectory. It is certainly the case that widely used standard developmental assessments of children younger than 48 months of age frequently incorporate components that are focused on vocal output, such as word production, grammatical representation, and social responsiveness.

Child vocalization frequency and conversational turn-taking accounted for 7%–16% of the variance observed in other assessments, which may not be surprising because relatively brief SLP observations and parent questionnaires can be considered a proxy for direct measure of natural behavior. On the other hand, one might consider these results to be remarkable, given that the automated measures are derived from completely objective processing, free of human evaluation and subjectivity. At the very least, the construct validity of the LENA measures suggests that the technology offers a viable supplement to standard early evaluation and intervention practices. In particular, LENA can provide professionals ongoing macro-level information about a child's vocal production more efficiently than could be accomplished via the administration of standard assessments. The clinical utility of this type of information for ongoing progress monitoring is a matter for further investigation.

Correlations with standard assessments were somewhat weaker for AWC estimates. One possible reason is that the CTs and child vocalizations measures both involve the child voice, so they may be less prone to error than AWC given the close proximity of the child's mouth to the recorder microphone and the typically greater distance of adult speakers. Furthermore, the LENA AWC count measure is void of content information and currently does not differentiate child-directed speech from overheard speech, potentially reducing correlations with developmental assessments because child-directed speech may be the primary driver of the correlations (see the following discussion of Weisleder and Fernald, 2013). A transcriptional analysis of adult word content focusing on quality of adult-child interactions may yield stronger associations with standard assessments. Still, the significant correlations between LENA AWCs and standard measures of child language skills reported here do support the contention that the quantity of language input children experience early in life is positively related to child language development, albeit perhaps due primarily to the quantity of child-directed speech.

Patterns of Behavior Among SES Groups

These results generally corroborate the Hart and Risley (1995) findings, demonstrating quantitative differences among SES groups on some language behaviors measured over the course of the day. Children whose mothers

graduated from college were exposed to 3,000 or so more words per day, translating into a four-million-word gap by 4 years of age between the highest and lowest SES groups in our sample. Although this average overall number is considerably smaller than the 30-million-word gap reported by Hart and Risley, it is important to note that their methodology differed from that of the current work in a number of ways. The estimate of Hart and Risley was based on 1-hr recordings extrapolated out to a 14-hr day. Given that they typically recorded during the early evening hours, which is a time of relatively high talk and interaction (Greenwood, Thiemann-Bourque, Walker, Buzhardt, & Gilkerson, 2011), their extrapolations likely resulted in inflated daily estimates. They also compared more extreme groups than here, as their lowest SES group was on public assistance, and their highest SES group consisted largely of academic professionals. However, we do note that for the AWC distribution reported for our sample (ignoring SES), the average daily 14-hr count difference between parents above the 98th percentile versus those below the 2nd percentile is approximately 20,500 words. Multiplied over 4 years, that daily difference equates to a word gap of 29.95 million words. That is, the 30-million-word gap observed decades ago by Hart and Risley does indeed apply here, albeit only when comparing the top versus bottom 2% of families.

As Figure 2 illustrates, there was little difference in AWC among the first three SES groups, whereas children living in households where the mother had graduated from college were exposed to significantly more adult words per day. This pattern stands in contrast to Hart and Risley's (1992, 1995) research, which reported increasing and significant differences in quantity of adult word exposure as a continuous function of SES. There are a number of factors that could account for this discrepancy, including differing methodologies. Alternately, compared with what Hart and Risley observed some 20 years earlier, that we did not see a difference within our first three SES groups could be reflective of a sort of merging in language behavior patterns for caregivers who have not completed college over that period, a genuine social change. In essence, the differences between the SES outcomes since the initial Hart and Risley data were collected (in the mid-1980s) could indicate that the divide between the language experience of children whose mothers have a college degree and those whose mothers do not could be widening, while within the group whose caregivers have not completed college the language experience of children could be narrowing. Recent research on the relationship between income and educational attainment for lower and higher income children supports this notion (Duncan & Murnane, 2011; Ramey & Ramey, 2009). Although the measurable shift among SES groups in daily talk with infants and toddlers could represent a shift in parenting culture, this speculation should be a focus of future research.

The current research allows us to expand on Hart and Risley's SES findings; with a larger number of children in the high and low SES groups, it is possible to look at within-group variation. Figure 3 shows that there were

children living in low SES households experiencing a rich language environment in terms of quantity of adult word exposure and interaction. Although on average, SES group differences were large and significant, our data provide new perspectives on the findings of Hart and Risley, elucidating important variation within SES groups. Full-day in situ measurement of the naturalistic home language environment suggests that many high SES children experience a relative paucity of adult language input, and many lower SES parents can and do talk with their children at above-average levels. This finding is in line with the Hart and Risley data, which also demonstrated variation within SES groups. However, the depth of information provided about the daily behavior of individual families within the current sample suggests that intervention is probably unnecessary for many low-income parents, while some children living in higher SES households could benefit from language environment enrichment. Direct measures of interaction in the home may provide a better indication of the need for this type of intervention compared to SES alone.

Limitations

These results notwithstanding, limitations inherent to the automated method are noteworthy. For one, currently the implementation provides no information about the content of words spoken to or near the child. In this sense, automated analysis explicitly does not characterize the quality of speech or interaction, and it has rightly been suggested that a focus on quantity should neither ignore nor minimize the importance of the quality of early language input to child development (Cartmill et al., 2013; Rowe, 2012). However, studies using transcriptional analyses have reported significant correlations between the quantity of adult input and the quality of the speech in terms of lexical diversity, as well as grammatical and syntactic complexity (Hoff, 2003; Rowe, 2008). Furthermore, Hart and Risley (1995) found that increases in adult talk were associated with increases in quality features of interaction, including number of affirmations, as well as caregiver questions directed to the child.

In addition to lacking content information for adult speech, the technology currently does not identify the addressee. Although research has shown that children can and do learn vocabulary from overheard speech (Akhtar, Jipson, & Callanan, 2001; Oshima-Takane, Goodz, & Derevensky, 1996), the extent to which it impacts various aspects of communicative development at different ages has been challenged (Oller, 2010). A recent study using LENA technology investigated the developmental influence of child-directed compared with overheard speech for 29 low-income, Spanish-learning Latino children recorded at 19 months of age (Weisleder & Fernald, 2013). Analysis of coded audio showed that the quantity of child-directed adult speech correlated with children's vocabulary and language processing measures, whereas the quantity of adult-directed overheard speech did not. Applied to the current research, this finding suggests that the proportion of automated adult word estimates

attributable to overheard speech may be significantly less impactful than the portion attributable to child-directed speech. Although coding for addressee was outside the scope of this study, we note that Hart and Risley (1995) reported approximately 50%–75% of utterances spoken around 11- to 18-month-olds included child-directed speech, with variation among socioeconomic subgroups. Regardless of the degree of impact or relative proportion of overheard speech it contains, significant correlations reported here between AWC and developmental assessment scores do indicate that the measure is informative. More work is needed to parse out the child-directed speech from AWC estimates and to establish its relative impact on language acquisition at different ages.

Analysis of variability over time showed that child vocalizations and CTs demonstrated higher within-family correlations than AWCs. Because behaviors can vary day-to-day, we do not expect perfect concordance between measures derived from different recordings. But given that the child is wearing the recorder, the metrics more closely associated with their vocal output (child vocalizations and turns) demonstrate relatively greater stability compared to AWCs, which in part depend on the physical proximity of the adult to the child. Thus, the clinical use of single recordings may be more limited for AWCs than for CTs or child vocalizations, and it is advisable to collect multiple recordings when possible. Still, to the extent that the word counts are representative of the child's language experience on a particular day, the AWC metric can be useful for providing feedback to caregivers about that day's language behavior.

There are other limitations associated with automated analysis. Age-related differences in behavioral measures could, to some extent, be attributed to changing sleep patterns or modeling techniques used at different ages or to differences in the reliability of automated voice labeling across age. Further, quantitative measures do not include information about emotional tone or nonverbal communication associated with gestures and smiles or other semantic, quality-focused measures. Despite the unavailability of context-related language environment estimates, we believe that the empirical data reported here show that meaningful information about the early language environment can be attained through automated methods.

Limitations associated with the current research sample should be addressed in future work. Although participant families approximated the U.S. population on mother's education level, which can be considered a proxy for SES, the sample was neither geographically nor ethnically diverse. Thus, while this study informs our understanding of language behaviors in monolingual English-speaking households in the Denver metro area, future research with a more diverse population is needed to investigate the extent to which these patterns generalize to a broader range of languages, ethnicities and cultures, especially with respect to immigrant populations. That is, until a more suitably comprehensive recording effort may be undertaken, it is advisable not to consider the language environment distributions presented

here as definitively normative for typical parent-child interactions in either research or clinical applications with demographic subsets that differ markedly from the included sample. As well, the large number of audio recordings across ages used in the current research are suggestive of developmental changes over time, but the mostly cross-sectional analyses do not provide information about child-specific longitudinal development over time. Thus, patterns of within-family change seen here may not necessarily apply between families generally. These issues can be addressed more appropriately, for example, by applying multilevel modeling and other statistical methods to these data and reported in more depth in the future.

Clinical Practice Implications

Acknowledging such limitations, correlations observed between child-related LENA measures and standard assessments seem to reveal the potential for the use of automated language environment analysis as a Level 1 screen to identify children at risk for impoverished language environments and to support parent-focused intervention programs designed to increase talk and interaction in the home. While established means for flagging children at risk for developmental issues already exist (e.g., parent questionnaires), no other direct measures currently exist that identify children at risk for impoverished language environments. Thus, the AWC and CT reference data reported here provide a crucial benchmark for identifying children experiencing low levels of talk and interaction in the home. Given the established relationship between a child's early environment and cognitive, social, and emotional development, an automated screening and referral approach could have a profound impact for children who, though not necessarily at clinically low levels with respect to language development, need more stimulation for healthy development. Further, given that children with language delays and disabilities may elicit less talk from caregivers and consequently experience deficits in environmental input (Warren & Brady, 2007), they could also benefit from interventions focused on remediating conditions of impoverished language environments. For now, clinical applications should be limited to families demographically representative of the current sample. Further research is needed to test the effects of this type of screening and referral at different ages, given that very early parental behavior change in the natural environment could have maximal impact on development over time.

Future Directions

The ability to directly and objectively measure the natural language environment offers a new window into the experience of young children and families and shows enormous potential for future research. An important next step is to investigate the relationship between quantitative measures derived from automated analysis and the content-related quality of language input. Although previous research has shown that as the number of adult words increases,

lexical variety, syntactic complexity, and positive valence also increase (Hart & Risley, 1995; Hoff, 2003; Rowe, 2008), these results should be empirically verified using all-day recordings and automated analysis. Future efforts should explore the clinical utility of automated measures for both supplementing initial developmental evaluations and for continuous progress monitoring in early intervention and should investigate the extent to which children with developmental language disabilities from various diagnostic subsets can benefit from interventions focused on enhancing the early language environment. It will also be important to study the language environments experienced by children living in other countries and cultures and to validate the system and establish normative standards for other languages. We believe that the data presented here can serve to provide initial normative reference expectations for clinical use as a Level 1 screen in populations that resemble the demographic distribution of the current sample, and in the future, this technology could be used clinically with a wider reach when more broadly representative normative data are available. Future research efforts may also explore the possibility of automatically identifying child-directed speech and emotional tone.

Conclusions

The current research provides an expanded view into the landscape of the early language environment of young American children. Automatically generated language estimates from daylong audio recordings revealed patterns consistent with a number of phenomena reported in the literature and uncovered new information about age-related developmental changes, as well as variability within socioeconomic strata. Although there are limitations to the types of information available through automated analysis technology, the advantages offered in terms of representative sampling and conservation of human resources for transcription and coding are substantial. Given the important role of the early home environment to a child's language, social, cognitive, and other development and the potential benefits of early identification of deficits for successful intervention, the clinical potential for an automated data collection and analysis approach is far reaching.

Acknowledgments

This research was funded by the LENA Research Foundation, a nonprofit 501(c)(3) public charity. We gratefully acknowledge the Paul family for their wisdom and philanthropy, without which none of this work would have been possible; the LENA parents and children; the members of the LENA Scientific Advisory Board; and LENA Research Foundation employees past and present who contributed to this study.

References

Accardo, P. J., & Capute, A. J. (2005). *The Capute Scales: Cognitive Adaptive Test/Clinical Linguistic & Auditory Milestone Scale (CAT/CLAMS)*. Baltimore, MD: Brookes.

- Akhtar, N., Jipson, J., & Callanan, M. A. (2001). Learning words through overhearing. *Child Development*, 72, 416–430. <https://doi.org/10.1111/1467-8624.00287>
- Aragon, M., & Yoshinaga-Itano, C. (2012). Using Language ENvironment Analysis to improve outcomes for children who are deaf or hard of hearing. *Seminars in Speech and Language*, 33, 340–353. <https://doi.org/10.1055/s-0032-1326918>
- Bornstein, M. H., Tamis-LeMonda, C. S., Hahn, C. S., & Haynes, O. M. (2008). Maternal responsiveness to young children at three ages: Longitudinal analysis of a multidimensional, modular, and specific parenting construct. *Developmental Psychology*, 44, 867–874. <https://doi.org/10.1037/0012-1649.44.3.867>
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Bzoch, K. R., League, R., & Brown, V. L. (2003). *Receptive-Expressive Emergent Language Test* (3rd ed.). Austin, TX: Pro-Ed.
- Cartmill, E. A., Armstrong, B. F., III, Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences*, 110, 11278–11283. <https://doi.org/10.1073/pnas.1309518110>
- Caskey, M., Stephens, B., Tucker, R., & Vohr, B. (2011). Importance of parent talk on the development of preterm infant vocalizations. *Pediatrics*, 128, 910–916. <https://doi.org/10.1542/peds.2011-0609>
- Caskey, M., & Vohr, B. (2013). Assessing language and language environment of high risk infants and children: A new approach. *Acta Paediatrica*, 102, 1–11. <https://doi.org/10.1111/apa.12195>
- Delaney, E. M., & Kaiser, A. P. (2001). The effects of teaching parents blended communication and behavior support strategies. *Behavioral Disorders*, 26, 93–116.
- Duncan, G. J., & Murnane, R. J. (Eds.). (2011). *Whither opportunity? Rising inequality, schools, and children's life chances*. New York, NY: Russell Sage Foundation.
- Dykstra, J., Sabatos-DeVito, M. G., Irvin, D. W., Boyd, B. A., Hume, K. A., & Odom, S. L. (2012). Using the Language Environment Analysis (LENA) system in preschool classrooms with children with autism spectrum disorders. *Autism*, 17, 582–594. <https://doi.org/10.1177/1362361312446206>
- Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories*. Baltimore, MD: Brookes.
- Goldberg, S., Corter, C., Lojkasek, M., & Minde, K. (1990). Prediction of behavior problems in 4-year-olds born prematurely. *Development and Psychopathology*, 2, 15–30. <https://doi.org/10.1017/S0954579400000560>
- Greenwood, C. R., Thiemann-Bourque, K., Walker, D., Buzhardt, J., & Gilkerson, J. (2011). Assessing children's home language environments using automatic speech recognition technology. *Communication Disorders Quarterly*, 32, 83–92. <https://doi.org/10.1177/1525740110367826>
- Hart, B., & Risley, T. R. (1992). American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Developmental Psychology*, 28, 1096–1105. <https://doi.org/10.1037/0012-1649.28.6.1096>
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.
- Hemmeter, M. L., & Kaiser, A. P. (1994). Enhanced milieu teaching: Effects of parent-implemented language intervention. *Journal of Early Intervention*, 18, 269–289. <https://doi.org/10.1177/105381519401800303>

- Hoff, E. (2003). The specificity of environmental influence: Socio-economic status affects early vocabulary development via maternal speech. *Child Development*, 74, 1368–1378. <https://doi.org/10.1111/1467-8624.00612>
- Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, 73, 418–433.
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, 27, 236–248. <https://doi.org/10.1037/0012-1649.27.2.236>
- Ireton, H. (1992). *Child Development Inventory*. Minneapolis, MN: Behavior Science Systems.
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., Jasnow, M. D., Rochat, P., & Stern, D. N. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development*, 66(2), i–viii, 1–132.
- Kaiser, A., & Hancock, T. (2003). Teaching parents new skills to support their young children's development. *Infant & Young Children*, 16, 9–21.
- Kochanska, G., Forman, D. R., & Coy, K. C. (1999). Implications of the mother-child relationship in infancy for socialization in the second year of life. *Infant Behavior & Development*, 22, 249–265. [https://doi.org/10.1016/S0163-6383\(99\)00009-0](https://doi.org/10.1016/S0163-6383(99)00009-0)
- Landry, S. H., Smith, K. E., & Swank, P. R. (2006). Responsive parenting: Establishing early foundations for social, communication, and independent problem-solving skills. *Developmental Psychology*, 42, 627–642. <https://doi.org/10.1037/0012-1649.42.4.627>
- Landry, S. H., Smith, K. E., Swank, P. R., & Guttentag, C. (2008). A responsive parenting intervention: The optimal timing across early childhood for impacting maternal behaviors and child outcomes. *Developmental Psychology*, 44, 1335–1353. <https://doi.org/10.1037/a0013030>
- Landry, S. H., Smith, K. E., Swank, P. R., & Miller-Loncar, C. L. (2000). Early maternal and child influences on children' later independent cognitive and social functioning. *Child Development*, 71, 358–375. <https://doi.org/10.1111/1467-8624.00150>
- Oller, D. K. (2000). *The emergence of the speech capacity* (Vol. XVII). Mahwah, NJ: Lawrence Erlbaum.
- Oller, D. K. (2010). All-day recordings to investigate vocabulary development: A case study of a trilingual toddler. *Communication Disorders Quarterly*, 31, 213–222. <https://doi.org/10.1177/1525740109358628>
- Oller, D. K., & Lynch, M. P. (1992). Infant vocalizations and innovations in infraphonology: Toward a broader theory of development and disorders. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 509–536). Timonium, MD: York Press.
- Oller, D. K., Niyogi, P., Gray, S., Richards, J. A., Gilkerson, J., Xu, D., . . . Warren, S. F. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences*, 107, 13354–13359. <https://doi.org/10.1073/pnas.1003882107>
- Oshima-Takane, Y., Goodz, E., & Derevensky, J. L. (1996). Birth order effects on early language development: Do secondborn children learn from overheard speech? *Child Development*, 67, 621–634. <https://doi.org/0009-3920/96/6702/0021>
- Pae, S. (2013, April). *An investigation of the language environment of young Korea children and the impact of feedback on language development*. Paper presented at the LENA International Conference 2013, Denver, CO.
- Ramey, G., & Ramey, V. A. (2009). *The rug rat race* (Working Paper 15284). Cambridge, MA: National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w15284.pdf>
- Roberts, M. Y., & Kaiser, A. P. (2011). The effectiveness of parent-implemented language interventions: A meta-analysis. *American Journal of Speech-Language Pathology*, 20, 180–199. [https://doi.org/10.1044/1058-0360\(2011/10-0055\)](https://doi.org/10.1044/1058-0360(2011/10-0055))
- Rochester, S. R. (1973). The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research*, 2, 51–81. <https://doi.org/10.1007/BF01067111>
- Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, 35, 185–205. <https://doi.org/10.1017/S0305000907008343>
- Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, 83, 1762–1774. <https://doi.org/10.1111/j.1467-8624.2012.01805.x>
- Sameroff, A. J., & Fiese, B. H. (Eds.). (2000). *Transactional regulation: The developmental ecology of early intervention* (Vol. 2). Cambridge, United Kingdom: Cambridge University Press.
- Snow, C. E. (1977). The development of conversation between mothers and babies. *Journal of Child Language*, 4, 1–22. <https://doi.org/10.1017/S0305000900000453>
- Stern, D. N. (1974). Mother and infant at play: The dyadic interaction involving facial, vocal, and gaze behaviors. In M. R. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. xxiv, 264). New York, NY: Wiley.
- Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development*, 72, 748–767. <https://doi.org/10.1111/1467-8624.00313>
- U.S. Census Bureau. (2005). *Educational attainment of the population 15 years and over, by single years of age, sex, race, and Hispanic origin: 2004*. Retrieved from <https://www.census.gov/hhes/socdemo/education/data/cps/2004/tables.html>
- U.S. Census Bureau. (2007). *Population profile of the United States: Dynamic version. Race and Hispanic origin in 2005, Table 1 population by race and Hispanic origin: 2000 and 2004*. Retrieved from <http://www.census.gov/population/pop-profile/dynamic/RACEHO.pdf>
- van Ijzendoorn, M. H., Rutgers, A. H., Bakermans-Kranenburg, M. J., Swinkels, S. H., van Daalen, E., Dietz, C., . . . van Engeland, H. (2007). Parental sensitivity and attachment in children with autism spectrum disorder: Comparison with children with mental retardation, with language delays, and with typical development. *Child Development*, 78, 597–608. <https://doi.org/10.1037/0033-2909.133.1.1>
- Walker, D., Greenwood, C., Hart, B., & Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic factors. *Child Development*, 65, 606–621. <https://doi.org/10.1111/j.1467-8624.1994.tb00771.x>
- Warren, S. F., & Brady, N. C. (2007). The role of maternal responsiveness in the development of children with intellectual disabilities. *Mental Retardation and Developmental Disabilities Research Reviews*, 13, 330–338. <https://doi.org/10.1002/mrdd.20177>
- Warren, S. F., Gilkerson, J., Richards, J. A., Oller, D. K., Xu, D., Yapanel, U., & Gray, S. (2010). What automated vocal analysis reveals about the vocal production and language learning environment of young children with autism. *Journal of Autism and Developmental Disorders*, 40, 555–569. <https://doi.org/10.1007/s10803-009-0902-5>

- Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24, 2143–2152. <https://doi.org/10.1177/0956797613488145>
- Wheeler, A., Hatton, D., Reichardt, A., & Bailey, D. (2007). Correlates of maternal behaviours in mothers of children with fragile X syndrome. *Journal of Intellectual Disability Research*, 51, 447–462. <https://doi.org/10.1111/j.1365-2788.2006.00896.x>
- Wiggin, M., Gabbard, S., Thompson, N., Goberis, D., & Yoshinaga-Itano, C. (2012). The school to home link: Summer preschool and parents. *Seminars in Speech and Language*, 33, 290–296. <https://doi.org/10.1055/s-0032-1326919>
- Xu, D., Richards, J. A., & Gilkerson, J. (2014). Automated analysis of child phonetic production using naturalistic recordings. *Journal of Speech, Language, and Hearing Research*, 57, 1638–1650. https://doi.org/10.1044/2014_JSLHR-S-13-0037
- Xu, D., Yapanel, U., Gray, S., Gilkerson, J., Richards, J. A., & Hansen, J. (2008a, October). *Signal processing for young child speech language development*. Paper presented at the Workshop on Child, Computer and Interaction, Chania, Crete, Greece.
- Xu, D., Yapanel, U., Gray, S., Gilkerson, J., Richards, J. A., & Hansen, J. (2008b). *Signal processing for early child speech and language development*. Paper presented at the Workshop on Child, Computer and Interaction, Chania, Crete, Greece
- Yoder, P. J., Warren, S. F., Kim, K., & Gazdag, G. E. (1994). Facilitating prelinguistic communication skills in young children with developmental delay II: Systematic replication and extension. *Journal of Speech and Hearing Research*, 37, 841–851. <https://doi.org/10.1044/jshr.3704.841>
- Zhang, Y., Xu, X., Jiang, F., Gilkerson, J., Xu, D., Richards, J. A., ... Topping, K. J. (2015). Effects of quantitative linguistic feedback to caregivers of young children: A pilot study in China. *Communication Disorders Quarterly*, 37, 16–24. <https://doi.org/10.1177/1525740115575771>
- Zimmerman, F. J., Gilkerson, J., Richards, J. A., Christakis, D. A., Xu, D., Gray, S., & Yapanel, U. (2009). Teaching by listening: The importance of adult-child conversations to language development. *Pediatrics*, 124, 342–349. <https://doi.org/10.1542/peds.2008-2267>
- Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2002). *Preschool Language Scale* (4th ed.). San Antonio, TX: The Psychological Corporation.