Developmental Science

Developmental Science 17:6 (2014), pp 880–891 DOI: 10.1111/desc.12172

PAPER

Look who's talking: speech style and social context in language input to infants are linked to concurrent and future speech development

Nairán Ramírez-Esparza,[1] Adrián García-Sierra[2] and Patricia K. Kuhl[2]

1. Department of Psychology, University of Connecticut, USA
2. Institute for Learning & Brain Sciences, University of Washington, USA

Abstract

*Language input is necessary for language learning, yet little is known about whether, in natural environments, the speech style and social context of language input to children impacts language development. In the present study we investigated the relationship between language input and language development, examining both the style of parental speech, comparing 'parentese' speech to standard speech, and the social context in which speech is directed to children, comparing one-on-one (1:1) to group social interactions. Importantly, the language input variables were assessed at home using digital first-person perspective recordings of the infants' auditory environment as they went about their daily lives (*N *=26, 11- and 14-months-old). We measured language development using (a) concurrent speech utterances, and (b) word production at 24 months. Parentese speech in 1:1 contexts is positively correlated with both concurrent speech and later word production. Mediation analyses further show that the effect of parentese speech-1:1 on infants' later language is mediated by concurrent speech. Our results suggest that both the social context and the style of speech in language addressed to children are strongly linked to a child's future language development.*

Introduction

There is substantial research showing the impact of language input to children on language learning (e.g. Hurtado, Marchman & Fernald, 2008; Huttenlocher, Haight, Bryk, Seltzer & Lyons, 1991; Rowe, 2012). In the classic longitudinal study by Hart and Risley (1995, 1999) one-hour recordings were made of 42 families once a month for two and a half years as they interacted in a natural setting. Results suggested that the quantity as well as the quality of the language used by parents interacting with their children was related to the children's vocabulary. One finding was that vocabulary grew faster in children exposed to a greater quantity of language regardless of socioeconomic status (SES). Hart and Risley (1999) also found that conversations associated with parent–child interaction during a joint activity (e.g. while working on a puzzle) were more powerful predictors of later language than interactions during everyday activities (e.g. while feeding or dressing the child).

The influence of Hart and Risley's findings has been substantial. However, there is little experimental data examining natural language interactions in home settings and their associations with concurrent and later measures of speech development. The style of language used to address children in the home, and the social context in which language interactions occur at home, and the interaction of both factors on language development, warrant examination.

For example, the use of a 'parentese' speech style (as opposed to standard speech) has been the focus of many studies. Parentese is simplified at the grammatical and lexical levels, and has a unique acoustic and visual signature: acoustically, it is characterized by higher pitch, slower tempo, and exaggerated intonation contours (Fernald, 1985; Grieser & Kuhl, 1988); visually, parentese exaggerates articulatory gestures and social affect

Address for correspondence: Nairán Ramírez-Esparza, Department of Psychology, University of Connecticut, Bousfield Psychology Building, 181, 406 Babbidge Road, U-1020, Storrs, CT 06269, USA; e-mail: nairan.ramirez@uconn.edu

(Weikum, Vouloumanos, Navarra, Soto-Faraco, Sebastián-Gallés & Werker, 2007). Parentese contains particularly good phonetic exemplars – sounds that are clearer, longer, and more distinct from one another – acoustically 'exaggerated' when compared to standard speech (Burnham, Kitamura & Vollmer-Conna, 2002; Kuhl, Andruski, Chistovich, Chistovich, Kozhevnikova, Ryskina, Stolyarova, Sundberg & Lacerda, 1997). The degree to which mothers show this acoustic exaggeration when talking to their infants is associated with their infants' performance in the laboratory on the discrimination of difficult synthetic speech contrasts (Liu, Kuhl & Tsao, 2003).

Parentese is hypothesized to be an implicit social response that facilitates communication (Hurtado *et al.*, 2008; Kuhl, 2004; see also Hoff, 2006), and many have hypothesized that parentese is beneficial to language learners (Fernald, 1985; Fernald & Kuhl, 1987; Hirsh-Pasek, Kemler Nelson, Jusczyk, Cassidy, Druss & Kennedy, 1987; Karzon, 1985; Kemler Nelson, Hirsh-Pasek, Jusczyk & Cassidy, 1989). However, while parentese speech is linked to language development, its use by caregivers in the home and its relation to infants' early speech development has not been studied. Moreover, the prevalence of parentese speech in the everyday lives of infants is unknown.

Of equal importance, the social context of language addressed to children in the home – whether adults talk to infants in one-on-one vs. group social interactions – and its impact on language development, has not thus far been studied. There is evidence from laboratory studies that parental social feedback in one-on-one settings affects infant speech development. Goldstein and colleagues (Goldstein, King & West, 2003) had caregivers respond to their infants' babbling by smiling at, moving closer to, and touching their infants. These conditions were compared with 'yoked' controls in which caregivers responded similarly, but responses were timed by the experimenter's instructions, rather than infants' vocalizations. Infants in the contingent group not only produced more vocalizations than those in the yoked group; their vocalizations were more mature than those in the yoked group. Goldstein and Schwade (2008) used the same design to show that when caregivers respond to infant babbling with either fully resonant vowels or words, infants increased production of the types of speech they heard from caregivers, and also extended their vocalizations to new phonetic forms. Yoked control infants did not change the phonological characteristics of their babbling. These studies suggest that social responses to infants' early speech may advance infants' speech production capabilities.

The impact of a social context on language learning is also seen in studies in this laboratory demonstrating that second-language learning in infancy at both the phonetic and word levels requires social interaction; infants learn from 'live' tutors but not from the same information delivered via a TV monitor (Conboy & Kuhl, 2011; Kuhl, 2007; Kuhl, Tsao & Liu, 2003). In addition, infants' social responses during language exposure sessions predict the degree to which individual children learn phonemes and words (Conboy, Brooks, Meltzoff & Kuhl, under revision; Kuhl, 2011).

Taken together, previous work led us to hypothesize that there may be an advantage to speech directed toward infants in a one-on-one social context. Adults interacting with infants in a 1:1 context would have ample opportunity to respond to infants contingently, therefore contributing to the advancement of speech production and language learning. In contrast, infants interacting with adults as a group may experience fewer instances of contingent social reactions. More frequent group language interactions, in which more than one adult interacts with the infant, as opposed to one-on-one language interactions, would be expected to be less effective in advancing language learning.

The purpose of the present study was to investigate two variables regarding language input to infants in a natural environment – speech style and social context – and their association with the infant's language development. We were interested in the impact of specific speech styles (parentese vs. standard speech) and specific social contexts (one-on-one vs. group social interactions) in the home on infants' concurrent speech and on their future language development.

In order to accomplish this aim, we used a recently developed language environment analysis system (LENA Foundation, Boulder, CO) to make first-person perspective recordings of the auditory environment in the everyday lives of 11- and 14-month-old infants. The LENA system provides a digital language processor (DLP) that can store up to 16 hours of digitally recorded sound. The DLP weighs 3 oz and can be snapped into a chest pocket in children's clothing, allowing the recorder to be 'out of sight, out of mind'. The audio recordings are downloaded to a computer, analyzed by LENA software to characterize the acoustic environment over time and efficiently identify segments containing adult speech. Identified segments containing adult speech directed to infants were then coded for social language behaviors, assessing the speech style used by adults, contrasting parentese speech with standard speech, and assessing the social context of language interactions, contrasting segments with one adult voice (1:1 social context) and more than one adult voice (group social context). Our goal was to examine, in a natural setting, the associations between these social interaction variables (speech style and social context) and concurrent infant speech utterances as well as later word production.

The LENA approach provides advantages over other observational techniques: (1) first-person digital recordings are made as infants go about their everyday lives, allowing quantification of behaviors not observable in relatively short interactions recorded in the laboratory or in the household; (2) the method allows assessment of speech from all adults (i.e. mom, dad, and/or other) interacting with the infants, instead of the typical situation in which one caregiver's interactions are measured in settings predetermined by researchers; (3) recordings are made unobtrusively since no direct observers are necessary to record and quantify behavior in real time; and (4) ambient sounds, as well as speech, are recorded and together they provide a snapshot of language input in real social contexts, allowing complex codification of the social scene in which language input to the infant occurs.

Infants' concurrent speech development was measured by assessing infant's speech utterances using eight categories coded from the digital recordings. In addition, we assessed later speech development, asking parents to complete a survey reporting word production when their child was 24 months old. Based on previous work, we predicted that infants' speech utterances would be related to later word production (e.g. Oller, Eilers, Neal & Schwartz, 1999; Stoel-Gammon & Sosa, 2009; Warren, Gilkerson, Richards, Oller, Xu, Yapanel & Gray, 2010). Consequently, our approach investigated the relationship between concurrent speech utterances and longitudinal word production measures via mediation analyses.

In summary, our goal in this investigation was to measure the association between the speech style and social context in which language input occurs in a natural home environment and both concurrent and future language development. We examined two social language variables: speech style (parentese speech vs. standard speech) and social context (1:1 interaction vs. group interaction). We were interested in both the overall *quantity* of adult speech (e.g. number of words) as well as the *quality* of speech in terms of speech style and social context that young infants experience in their homes over a four-day period during which recordings were made.

Method

Participants

The participants were 26 infants (15 females, 11 males). Two age groups, 11 months ($N = 13$ age range 10 months and 30 days to 11 months and 22 days) and 14 months ($N = 13$ age range 13 months and 24 days to 14 months and 20 days), were recruited as part of an ongoing large-scale study at the Institute for Learning & Brain Sciences. All infants were full-term (37–43 weeks), normal birth weight (2.5–4.5 kg) and had no major birth or postnatal complications. English was the only language spoken in the home. The sample was 80.8% white and 19.2% other (i.e. three participants were white and Hispanic, one was white and Native American/ Eskimo and one was white, black and Native American). Socioeconomic status (SES) was assessed using the Hollingshead index (Hollingshead, 1975), a broadly used SES measure producing an overall SES score based on parental education level and occupation ($M = 54.7$, $SD = 9.63$, Range = 21–66). Social language interaction data, including infants' speech utterances, were coded based on recordings from all participants, and parental reports of later word production (at 24 months) were obtained from 23 of 26 families who participated in the study.

Social context and language activity assessment

Data collection

Parents received two digital language processors (DLPs) and vests with a chest pocket designed to hold the DLP, allowing digital first-person perspective recordings of the infants' auditory environment at home as they went about their daily lives. They were instructed to record eight continuous hours each day for four consecutive days (including two weekdays and two weekend days), yielding approximately 32 hours of recorded audio data from each infant. Parents were also asked to complete a daily activity diary, noting the most relevant activities for each day.

Data preparation

LENA software was used to quantify language input and to efficiently locate intervals with the language activity of interest (i.e. adult speech) in each participant's large dataset of recorded audio for further analysis of language input, social context, and infant speech utterances. The audio data were transferred from the DLP to a computer and analyzed employing advanced speech-identification algorithms that automatically analyze audio files and produce reports of language activity. The LENA algorithms produced a total adult word count across all four days for each participant in the study. These data provided an estimate of the *quantity* of language input to the child.

The audio files were then further processed using the LENA Advanced Data Extractor Tool (ADEX) in order to efficiently identify intervals with the language activity of interest (i.e. adult speech), and eliminate intervals that did not qualify for analysis. This tool provides outputs of individual speech segments as short as a fraction of a second and was used to segment each participant's large dataset of recorded audio into 30-second intervals,[1] and to automatically calculate an adult word count for each interval. The goal was to identify 40 intervals for each participant on each of the four days – selected across the entire day and chosen from those with the highest adult word counts, thus yielding a total of 160 intervals for each participant. However, some participants failed to record eight hours per day for four days as instructed, and their recorded data yielded fewer than 160 intervals for coding. Consequently, an average of 156.73 (SD = 9.06) intervals per participant were coded, a total of 4,075 over the entire study.

Adapting the Social Environment Coding of Sound Inventory (SECSI) for infants

Mehl, Gosling and Pennebaker (2006) designed the Social Environment Coding of Sound Inventory (SECSI) to assess moment to moment naturalistic social behaviors, environments and interactions in adult populations (e.g. Mehl, Vazire, Ramírez-Esparza, Slatcher & Pennebaker, 2007; Ramírez-Esparza, Mehl, Álvarez-Bermúdez & Pennebaker, 2009). We adapted the SECSI, creating an infant version that focused on social context and language input. The Infant SECSI was designed to be a broad system and coded behaviors beyond the scope of the present study for use in future analyses. The Infant SECSI coded 73 categories organized into six clusters: 'speech partners', 'speech style', 'social context', 'infant speech utterances', 'activities', and 'infant mood'. A subset of 12 categories within these clusters was used to code each of the intervals selected for study. The 12 categories are shown in Table 1, and code 'speech partners' – mom speaks to infant, dad speaks to infant, other adult speaks to infant; 'speech style' – parentese speech is used to address the infant, standard speech is used to address the infant; 'social context' – infant is with one adult, infant is with two or more adults; and 'infant speech utterances' – utterances that qualify as speech (Kent & Murray, 1982): fully resonant vowels, consonant-vowel (CV) syllables, speech utterances intermixed with nonspeech, and word-like strings (see Table 2).

[1] Previous research demonstrated that a 30-second snapshot of ambient sounds provides sufficient information for judgments of behaviors (Mehl *et al*., 2006; Ramírez-Esparza *et al*., 2009).

Table 1 *Infant SECSI categories (N = 12) coded in language input: intercoder reliability*

Infant SECSI items	Intercoder reliability	Relative time use % intervals	
		Mean	*SD*
Speech partners cluster			
1. Mom speaks to infant	0.96	47.95	18.18
2. Dad speaks to infant	0.98	16.71	15.24
3. Other adult speaks to infant	0.91	7.94	9.42
Speech style cluster			
4. Mom using parentese speech	0.87	44.68	18.54
5. Dad using parentese speech	0.9	15.24	14.39
6. Other adult using parentese speech	0.88	7.15	9.48
7. Mom using standard speech	0.82	46.4	17.23
8. Dad using standard speech	0.79	25.63	13.86
9. Other adult using standard speech	0.95	30.7	22.21
Social context cluster			
10. Infant with one adult	0.95	48.17	20.28
11. Infant with two or more adults	0.95	48.17	19.21
Infant speech utterances			
12. Infant Utterances (Table 2)	0.92	65.41	14.33

Note: For all variables $N = 26$. Intercoder reliabilities were computed as intraclass correlations, ICC (2, k) from a training set of 120 intervals that were independently coded by five coders.

Table 2 *Concurrent speech development: infant speech utterances*

Infant speech utterance categories	Relative time use % intervals	
	Mean	*SD*
Single fully resonant vowel: /i/	20.2	12
Repetition of a fully resonant vowel: /i/, /i/, /i/	18.7	9.8
Variegated string of fully resonant vowels: /i/, /a/, /ae/	39.3	15.6
Single consonant-vowel syllable: /gu/	12.0	10.7
Repetition of consonant-vowel syllable: /ninini/	10.6	8.7
Variegated string of consonant-vowel syllables: /gamida/	26.1	14.6
Consonant-vowel syllables intermixed with squeals, growls, yells, raspberries	13.6	8.5
Consonant-vowel syllables with real words and stress and intonation patterns like whole sentences, without meaning	10.2	9.7

Coding selected intervals using 12 Infant SECSI categories

Five research assistants were trained to code the selected intervals for each participant (M = 156.73, SD = 9.06) using the 12 Infant SECSI categories described in Table 1. Coders were provided with basic information about each selected interval (date, day of the week, time of day, and the time stamp of the audio recording). Coders were also provided with the participants' end-of-day diaries to supplement audio recordings. Transcribing software played the specific 30-second interval for coding based on the time stamp entered. The coders listened to each 30-second interval and entered a 'YES' for each Infant SECSI category associated with the interval. For example, in a given 30-second interval the coders would listen and code all Table 1 behaviors that occurred. The resulting matrix of YES and NOs indicated that a specific Infant SECSI category occurred or did not occur in that interval. Infant SECSI categories are non-exhaustive and non-mutually exclusive; that is, several Infant SECSI categories could be coded within a single interval.

The Infant SECSI categories were easy to code but some required training to assure accuracy. All coders had experience coding parentese and infant vocalizations, but coders took part in additional training. During training, we defined and played examples of 'parentese' and standard speech, as well as examples of the eight infant speech utterances that were coded. After training in the use of the Infant SECSI, all coders were tested independently with a training file, which was used to evaluate inter-coder reliability. The 12 categories used in the analysis produced an average intra-class correlation of .91 – indicating effective training and reliable coding – based on a two-way random effects model (ICC [2, k]; Shrout & Fleiss, 1979). See Table 1 for the intercoder reliability for the 12 Infant SECSI categories employed in the present study.[2]

After coding, intervals were examined and excluded from further analysis if they met either of two criteria: (1) parents were instructed to read experimental materials to their children as part of the larger research project and beyond the scope of this study, and intervals including this experimental activity were excluded from analysis (M = 2.9 intervals, SD = 2.0); and (2) infants were sleeping during some intervals as evidenced by information coders gleaned from audio recordings and/or from the daily activity diary completed by caregivers, and these intervals were also excluded from analysis (M = 2.7 intervals, SD = 4.55). The remaining intervals then served as the full set of coded intervals for calculation of the relative time use estimates for each participant (M = 151.23, SD = 9.84).

Relative time use estimates of Infant SECSI categories

The coded data matrices containing YES and NO responses for each participant were aggregated to provide relative time use data by calculating the percentage of intervals coded for each category. For example, a relative time use estimate of 47.5% for the Infant SECSI category 'Mom speaks to infant' indicated that for a participant with 160 intervals, this category was coded YES in 76 of the 160 selected intervals for that participant. The Infant SECSI categories are not mutually exclusive, and a single 30-s interval may be coded YES for multiple categories. Consequently, the relative time use estimates of Infant SECSI categories are not expected to add to 100%. Table 1 provides the mean and standard deviations of coded interval percentages for the Infant SECSI categories employed in the present study.

Social interaction scenarios analyzed in the study

To test our hypothesis, we examined four different social interaction scenarios based on 11 of the 12 categories in the Infant SECSI (i.e. excluding infant speech utterances): (1) *Parentese speech-1:1* – mother, father, or other adult spoke directly to the infant, parentese speech was used, and only one adult voice was recorded during the interval, (2) *Parentese speech-group* – mother and/or father and/or other adult spoke directly to the infant, parentese speech was used, and two or more adult voices were recorded during the interval, (3) *Standard speech-1:1* – mother, father, or other adult spoke directly to the infant, standard speech was used, and only one adult voice was recorded during the interval, (4) *Standard speech-group* – mother and/or father and/or other adult spoke directly to the infant, standard speech was used, and two or more adult voices were recorded during the interval.[3]

[2] We independently verified that the samples coded as parentese vs. standard speech contained the acoustic differences characteristic of these two speech styles. Sixty occurrences of the word 'you' from the audio recordings of nine randomly selected study participants, representing 30 pairs produced by the same adult addressing the same infant (one coded as parentese, one as standard speech), were isolated for acoustic analysis. PRAAT, a computer software package for the analysis of speech, yielded measures of pitch range (i.e. maximum pitch–minimum pitch) and mean pitch for each sample. Repeated measures analysis of variance compare mean pitch and pitch range for parentese speech vs. standard speech. Mean pitch was significantly higher for parentese speech (M = 310 Hz) than standard speech (M = 219 Hz), $F(1, 29) = 36.650$, $p = .000$, $\eta_p^2 = .544$. The pitch range was also significantly larger for parentese speech (M = 68 Hz) than standard speech (M = 19 Hz), $F(1, 29) = 14.006$, $p = .001$, $\eta_p^2 = .326$.

[3] Although the mother is likely to spend more time with the child, all caregivers are part of the everyday life of the infant. For example, some participants were enrolled in daycare. Also, during the weekends, fathers are very involved in the everyday activities of the child. In sum, collapsing across speakers provided a more complete assessment of the child's natural everyday social interactions.

Relative time use estimates for these four social interaction scenarios were calculated for each participant and served as input to our statistical analyses. For each interval for each participant, the pattern of YES and NO responses was evaluated to determine whether that interval was an instance of one of the four social interaction scenarios defined above. The number of intervals identified as instances of a specific social interaction scenario for each participant was divided by the full set of intervals for that participant and multiplied by 100 to yield a percent relative time use estimate. For example, a relative time use estimate of 53.1% for parentese speech-1:1 in an individual participant with a full set of 160 intervals indicated that during 85 of 160 intervals, three Infant SECSI categories were coded YES: (1) the mother, father, *or* other adult spoke directly to the infant, (2) the adult used parentese speech, and (3) only one adult voice was recorded. Similarly, a relative time use estimate of 15.2% for standard speech-group in an individual participant with a full set of 145 intervals indicated that during 22 of 145 intervals, three Infant SECSI categories were coded YES: (1) the mother, father, *or* other adult spoke directly to the infant, (2) the adult used standard speech, and (3) more than one adult voice was recorded. Mean and standard deviation for percent of coded intervals for each of the four social interaction scenarios are provided in Table 3. Relative time use estimates for the social interaction scenarios were neither mutually exclusive nor exhaustive, and were therefore not expected to add to 100%.

Speech development assessment

Speech development was assessed concurrently and when participants were 24 months old. The concurrent measure was the relative time use estimate for infant speech utterances. It was calculated by dividing the number of intervals coded YES for infant speech utterances by the full set of intervals included in the analysis for each participant and multiplying by 100 as described above. For example, a relative time use estimate of 64.9% for infant speech utterances in an individual participant with a full set of 151 intervals indicated that 98 of 151 intervals were coded YES for the Infant SECSI category infant speech utterances.

The longitudinal speech development measure at 24 months was collected using the MacArthur-Bates Communicative Development Inventory (CDI; Fenson, Marchman, Thal, Dale, Reznick & Bates, 2007), a reliable and valid parent survey for assessing language and communication development from 8 to 30 months. Specifically, parents reported the number of words produced based on the 680-word check-list section of the CDI when the infants were 24 months old ($N = 23$, age range 1 year, 11 months, and 28 days to 2 years, 0 months and 9 days). At 24 months, the children produced an average of 376.04 words ($SD = 198.89$).

Table 3 *Relative time use for social interaction scenarios*

Social interaction scenarios	Relative time use % intervals	
	Mean	SD
Parentese Speech-1:1	41.12	19.62
Parentese Speech-group	19.21	8.16
Standard Speech-1:1	7.92	5.12
Standard Speech-group	19.36	8.6

Results

The initial step in analysis was evaluation of the overall effects of (1) age group at enrollment (i.e. 11 months vs. 14 months old), (2) SES (i.e. Hollingshead, 1975), and (3) the LENA generated measure of the total adult word count, which represented the *quantity* of adult speech in the environment across the four days. The LENA generated total word count for each participant across the four days ranged from 16,591 to 56,224 (Mean = 31,111.51, $SD = 9{,}885.52$).

We examined associations among these three variables as well as relationships to the other experimental variables (i.e. social interaction scenarios derived from the Infant SECSI, concurrent infant speech utterances derived from the Infant SECSI, and words produced at 24 months).

Participants enrolled in the study at 11 months or at 14 months showed no significant age group effects for SES, the LENA generated adult word count, social interaction scenarios, or speech development (infant speech utterances, words produced at 24 months). Participants were collapsed across age at enrollment for the remaining analyses.

Previous investigations have reported links between SES and speech development (e.g. Hart & Risley, 1995; Hoff, 2003) and we examined the relationships between SES (Hollingshead, 1975), the LENA generated adult word count, social interaction scenarios, and speech development. SES was significantly correlated only with percent intervals coded for parentese speech in a 1:1 social context ($r = .602$, $p = .001$, $n = 26$) (Table 4). Remaining analyses control for SES.

The adult word count provided by LENA did not correlate significantly with the social interaction scenarios

Table 4 *Correlations among socioeconomic status, adult word count, speech development measures and social interaction scenarios*

	Correlations with covariates	
Target variables	Hollingshead (SES)	LENA Adult Word Count
Hollingshead (SES)	1	
Lena Adult Word Count	0.085	1
Infant Speech Utterances	0.308	0
Words Produced at 24 months	0.151	0.267
Parentese Speech-1:1	.602*	−0.131
Parentese Speech-Group	−0.177	0.344
Standard Speech-1:1	0.252	0.248
Standard Speech-Group	−.314	.365^

^p = .073; *p < .001.
Note: LENA adult word count was generated using the LENA software feature that indicates the average number of words used by adults across the four days. Correlations between LENA adult word count, speech development measures and social interaction variables are controlled for SES.

or speech development, controlling for SES.[4] These analyses indicate that, at least in this study, the LENA measure of the *quantity* of words in an infant's environment is not related to speech development or to the social interaction scenarios targeted for study in this investigation.

We evaluated intercorrelations among social interaction scenarios and speech development, controlling for SES (Table 5). We found a pattern of significant correlations between standard speech-group context and parentese speech. Infants who experienced more standard speech in a group context also experienced less parentese speech in a 1:1 context ($r = -.608$, $p = .001$, $df = 23$), and more parentese speech in a group context ($r = .826$, $p = .000$, $df = 23$). In other words, in families in which standard speech in a group context is frequent, parentese speech occurs in a group context and there is less parentese speech in a 1:1 context.

Do social interaction variables relate to infant speech utterances?

We evaluated the associations between infant speech utterances and four social interaction scenarios: (1) parentese speech-1:1, (2) parentese speech-group, (3) standard speech-1:1 and (4) standard speech-group, controlling for SES. The results show that only increased parentese speech-1:1 was significantly related to increased infant speech utterances ($r = .54$, $p < .01$, $df = 23$) (Table 5, Figure 1B). Infants who experience more parentese speech-1:1 social interactions show increased percentages of infant speech utterances.

Do social interaction variables predict later word production?

The pattern of results for later word production is similar to that for the concurrent speech measures. Results show that the percentage of coded intervals for parentese speech-1:1 is positively related to later word production, $r = .429$, $p < .05$, $df = 20$ (Table 5, Figure 1A), controlling for SES. Other social interaction variables are neutral, unrelated to later word production.

Since concurrent speech is itself positively correlated with later word production, controlling for SES, $r = .718$, $p < .001$, $df = 20$ (Table 5, Figure 1C), we examined whether the relationship between parentese speech-1:1 and later word production is mediated by infant speech utterances using statistical mediation analyses. Following the guidelines of Baron and Kenny (1986), we found: (1) a significant relationship ($r = .429$, $p < .05$, $df = 20$) between the predictor variable (parentese speech-1:1) and the outcome variable (words produced at 24 months), (2) a significant relationship ($r = .539$, $p < .01$, $df = 23$) between the predictor variable (parentese speech-1:1) and the potential mediator (concurrent speech), and (3) a significant relationship ($r = .718$, $p < .001$, $df = 20$) between the mediator variable (concurrent speech) and the outcome variable (words produced at 24 months), controlling for SES (Figure 1). Using the moderated mediation macro developed by Preacher and Hayes (2008) with 5000 bootstrapping re-samples and introducing SES into the model as a covariate, the results show that the relationship between parentese speech-1:1 and words produced at 24 months is reduced in magnitude when concurrent speech is included in the model (i.e. from .031, $p < .05$ to .011, $p = .31$). Concurrent speech was deemed a significant mediator because the 95% bias-corrected confidence interval did not include zero (i.e. .0034 to .0403). The partial effect of the control variable SES on word count at 24 months was not significant (−.025, $p = .429$). The final model explained 55% of the variance of word count at 24 months (Figure 1).

Discussion

The goal of the current investigation was to examine the relationship between social characteristics of language input in the natural environment and speech

[4] Note that higher overall adult word count was associated with higher standard speech-group context, and this relationship is marginally significant ($r = .365$, $p = .073$, $df = 23$) (Table 3).

Table 5 *Correlations among social interaction scenarios and speech development measures, controlling for SES*

Experimental variables	Parentese Speech-1:1 $n = 26$	Parentese Speech-Group $n = 26$	Standard Speech-1:1 $n = 26$	Standard Speech-Group $n = 26$	Infant Speech Utterances $n = 26$	Word Count at 24 months $n = 23$
Parentese Speech-1:1	1					
Parentese Speech-Group	−0.318	1				
Standard Speech-1:1	0.019	−0.100	1			
Standard Speech-Group	−0.608**	0.826***	0.198	1		
Infant Speech Utterances	0.539**	−0.126	0.173	−0.321	1	
Words Produced at 24m	0.429*	0.068	−0.212	−0.159	0.718**	1

*$p < .05$; **$p < .01$; ***$p < .001$.

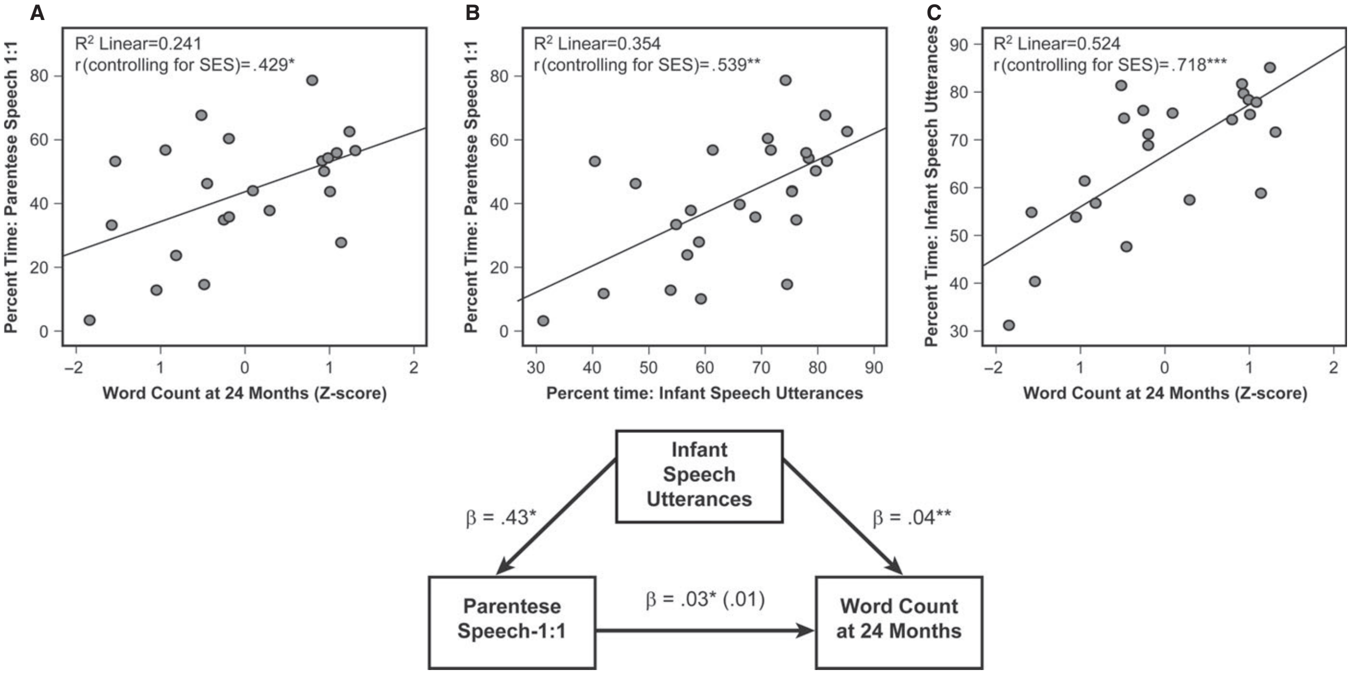

Figure 1 *Scatter plots showing the relationships between (A) Word Count at 24-months and Parentese Speech-1:1, (B) Infant Speech Utterances and Parentese Speech-1:1, (C) Word Count at 24 months and Infant Speech Utterances, and (D) Mediation analyses showing that Infant Speech Utterances mediate the relationship between Parentese Speech-1:1 and Word Count at 24 months. Sample size = 23 infants; β = indicates the regression coefficient; *p < .05; **p < .01; ***p < .001.*

development in 11- and 14-month-old infants. We focused on speech style – whether parentese or standard speech was more prevalent in language addressed to the child – and the social context in which language occurred – whether adult–infant interactions in a one-on-one or group context were more prevalent. To measure language input from a first-person perspective, we employed LENA technology to record the child's auditory world for four consecutive days, including two weekend days. The speech style and social context measures of language input in intervals identified by LENA and coded by trained listeners were evaluated to determine their association with both infants' concurrent language, as assessed by our measures of infant speech utterances, and their future language, as assessed by the number of words produced at 24 months. Moreover, we examined the role of socioeconomic status (SES) on these relationships.

Our results support three conclusions. First, analyses indicate that the raw *quantity* of language input to the child had no effect on our measures of language development, but that the *quality* of speech input to the child was strongly related to the child's developing language skills. The relationships between the LENA generated adult word count, concurrent infant speech, and word count at 24 months were not significant, whereas strong and significant correlations were observed between parentese speech-1:1 and measures of speech development. Links between early speech input and later speech development have also been reported in other investigations. For example, Hoff (2003) reported that vocabulary size in children is related to the quality of parental speech assessed by diversity of vocabulary and syntactical complexity. In a more recent study Rowe (2012) reported that the quality of speech – as measured by the sophistication of the vocabulary used by caregivers, while controlling for SES – predicts later language learning. Similar findings have been reported among Spanish learning children (see Hurtado *et al.*, 2008). Our focus was on the more social aspects of language – both the style in which speech is delivered to children and the social context in which language occurs. Moreover, we measured language input in the child's home environment. Taken together with previous findings, the results of the current study support the idea that the quality of speech experienced by children in their daily lives may make a difference in their language development.

Second, our results regarding speech style and social context demonstrate that infants' speech development is positively linked to the social environments they experience. Infants who interact with a single individual producing the acoustically exaggerated sounds of parentese speech produced greater concurrent speech and also more advanced later language. For example, when compared to children with the lowest amount of parentese speech-1:1 (<2 *SD*, $N = 6$), those with the highest amount of parentese speech-1:1 (>2 *SD*, $N = 6$) produced substantially more words (Mean word count = 168.55 vs. 433.33, respectively) at 24 months of age. The mediation analyses indicate that concurrent speech may also play a role in the observed association between parentese speech-1:1 and later word production. Our analyses suggest that the combination of speech style and social context relates to development through concurrent speech. This finding is in accord with other published findings in which a relation between early speech (canonical babbling) and later word production has been reported (e.g. Oller *et al.*, 1999; Stoel-Gammon & Sosa, 2009; Warren *et al.*, 2010). The present findings extend previous results by showing an association between infants' speech utterances and the frequency with which the infant experiences one-on-one social interaction with an adult speaking parentese in the home.

Third, the current study indicates that the effects of speech style and social context on language development persist when controlling for SES. Our finding that parentese speech-1:1 is positively related to SES is consistent with the Rowe and Goldin-Meadow (2009) finding, which reported that high SES parents use more gestures to communicate when talking to their 14-month-old children than low SES parents, and that at 14 months of age, children from high SES households also use more gestures to communicate meanings than children from low SES households. In the Rowe and Goldin-Meadow study, the use of gesture by the children partially mediated the association between SES and vocabulary size at 54 months of age. Thus, the quality of the interactions between parents and their children, measured with gesture and speech, is associated with advanced language development. This discrepancy between language input measures across SES may also help to explain the gap in vocabulary size between low and high SES families (see also Raizada, Richards, Meltzoff & Kuhl, 2008).

It is also interesting to consider the possibility that caregivers engage in 1:1 interactions and use parentese speech with infants who babble more, and are less likely to engage in these behaviors with infants who babble less (Locke, 2006). The current study did not select intervals on the basis of infant vocalizations, and therefore we cannot compare infant vocalizations in the presence or absence of adults to address, but this question is of interest for future work.

Social factors influencing language development

The findings of the current study suggest that social variables that accompany speech directed to young infants are correlated with the children's future speech development. These findings support other results indicating that the styles and settings that make up the social surround of an infant during social language exchanges may influence language learning. Specifically, our findings are consistent with the idea that infants' language learning is enhanced in one-on-one social contexts, perhaps because these contexts allow more contingent social interaction between adult and child, as suggested by the work of Goldstein and his colleagues. Moreover, our results indicate that the use of a socially natural speaking style used in many cultures to address infants – parentese speech – may contribute to language learning. Both of these variables were shown to be associated with increased infant babbling, and with the number of words produced at 24 months. Our analyses suggest that the number of words produced at 24 months is also associated with infants' speech utterances. Future work can examine children's early speech utterances with and without a conversational partner so that the amount of infant babbling *per se* can be taken into account.

Evidence is accumulating to indicate a greater role for social interaction and social contingency in language learning. A large number of studies now suggest an important role for social behavior in language learning, among them: (1) the necessity of a social context for phonetic learning from exposure to a new language (Conboy & Kuhl, 2011; Kuhl *et al*., 2003); (2) the role of contingent response to infants' babbling, which in turn results in more sophisticated vocalizations that are matched to the adults' vocalizations (Goldstein *et al*., 2003; Goldstein & Schwade, 2008), (3) the role of social behavior (eye gaze) as a tactic that provides vital information for language learning in infants (Brooks & Meltzoff, 2008). These studies underscore the interaction between social learning and language acquisition, a relationship that has been proposed regarding lexical learning (Tomasello, 2003). These findings contribute to our understanding of the social foundations of language, as observed in both typically developing and developmentally challenged young children, suggesting that language acquisition fundamentally *requires* social interaction to instigate learning (Kuhl, 2007; Kuhl, 2011; Kuhl, Coffey-Corina, Padden, Munson, Estes & Dawson, 2013).

Limitations and future directions

The current study was based on recorded sound, and previous studies have shown that language learning is related to infants' appreciation of others' communicative intentions (see Baldwin, 1995). Using video records, future observational studies could quantify the level of attention, habituation, and other variables to elucidate mechanisms that may accompany parentese speech in one-on-one social contexts. Future studies could also include measures on child and parent temperament, as well as other variables that could be related to social interaction in their everyday lives. There are data to suggest that mothers who are depressed do not produce the quantity and quality of speech observed in non-depressed mothers, and that this may be linked to language development in their children (e.g. Murray, Halligan & Cooper, 2010).

The present study is correlational in nature, and therefore we cannot determine causal relationships. The associations we observed, however, show for the first time an interesting link between a combination of speech style and social context (e.g. parentese speech-1:1) and both concurrent and future measures of language development. Our findings are consistent with the idea that infants' early speech and later word production may be related to the social context and the style of speech directed toward the child.

Conclusions

We show that the *quality* of social interactions – defined by both the social context and the style of speech – are strongly associated with concurrent and future infant speech development, regardless of the socioeconomic status of the family. Our data suggest that higher percentages of one-on-one social interaction between adults and their 11- and 14-month-old infants, and the use of 'parentese' by the adult during these 1:1 interactions are more potent predictors of concurrent and future language than the sheer amount of language input to the child. Parentese language input in one-on-one social contexts is strongly and positively associated with concurrent infant speech and also with word production at 24 months. This investigation reflects an important step forward in generating an ecologically valid understanding of the relationship between natural everyday social interactions among caregivers and infants and speech development. These speech style and social context characteristics can be implemented at home and in early care centers (Phillips & Lowenstein, 2011). Caregivers could attempt to provide quality care by fomenting short one-on-one conversations, helping to bridge the gap on vocabulary size among low and high SES children. In summary, caregivers who provide quality speech input to their infants will likely observe who's talking in the near future.

Acknowledgements

This research was supported by a grant from the National Science Foundation's Science of Learning Program to the University of Washington LIFE Center (SBE-0354453), and by grants from the National Institutes of Health (HD37954, DC04661). Special thanks to Elina Sanchez, Lindsay Klarman, Pat Stock, and Denise Padden for their assistance and input.

References

Baldwin, D.A. (1995). Understanding the link between joint attention and language. In C. Moore & P.J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 131–158). Hillsdale, NJ: Lawrence Erlbaum Associates.

Baron, R.M., & Kenny, D.A. (1986). The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, **51**, 1173–1182.

Brooks, R., & Meltzoff, A.N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: a longitudinal, growth curve modeling study. *Journal of Child Language*, **35**, 207–220.

Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new pussycat? On talking to babies and animals. *Science*, **296**, 1435.
Conboy, B.T., Brooks, R., Meltzoff, A.N., & Kuhl, P.K. (under revision). A role for social interaction in infants' learning of second-language phonetics.
Conboy, B.T., & Kuhl, P.K. (2011). Impact of second-language experience in infancy: brain measures of first- and second-language speech perception. *Developmental Science*, **14**, 242–248.
Fenson, L., Marchman, V.A., Thal, D., Dale, P.S., Reznick, J.S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual* (2nd edn.). Baltimore, MD: P.H. Brookes.
Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, **8**, 181–195.
Fernald, A., & Kuhl, P.K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, **10**, 279–293.
Goldstein, M.H., King, A.P., & West, M.J. (2003). Social interaction shapes babbling: testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences, USA*, **100**, 8030–8035.
Goldstein, M.H., & Schwade, J.A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, **19**, 515–523.
Grieser, D.L., & Kuhl, P.K. (1988). Maternal speech to infants in a tonal language: support for universal prosodic features in motherese. *Developmental Psychology*, **24**, 14–20.
Hart, B., & Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: P.H. Brookes.
Hart, B., & Risley, T.R. (1999). *The social world of children learning to talk*. Baltimore, MD: P.H. Brookes.
Hirsh-Pasek, K., Kemler Nelson, D.G., Jusczyk, P.W., Cassidy, K.W., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, **26**, 269–286.
Hoff, E. (2003). The specificity of environmental influence: socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, **74**, 1368–1378.
Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, **26**, 55–88.
Hollingshead, A.B. (1975). Four Factor Index of Social Status. Unpublished working paper, Department of Sociology, Yale University.
Hurtado, N., Marchman, V.A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, **11**, 31–39.
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.
Karzon, R.G. (1985). Discrimination of polysyllabic sequences by one- to four-month-old infants. *Journal of Experimental Child Psychology*, **39**, 326–342.
Kemler Nelson, D.G., Hirsh-Pasek, K., Jusczyk, P.W., & Cassidy, K.W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language*, **16**, 55–68.
Kent, R.D., & Murray, A.D. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *Journal of the Acoustical Society of America*, **72**, 353–365.
Kuhl, P.K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, **5**, 831–843.
Kuhl, P.K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science*, **10**, 110–120.
Kuhl, P.K. (2011). Social mechanisms in early language acquisition: understanding integrated brain systems supporting language. In J. Decety & J. Cacioppo (Eds.), *The Oxford handbook of social neuroscience* (pp. 649–667). Oxford: Oxford University Press.
Kuhl, P.K., Andruski, J.E., Chistovich, I.A., Chistovich, L.A., Kozhevnikova, E.V., Ryskina, V.L., Stolyarova, E.I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, **277**, 684–686.
Kuhl, P.K., Coffey-Corina, S., Padden, D., Munson, J., Estes, A., & Dawson, G. (2013). Brain responses to words in 2-year-olds with autism predict developmental outcomes at age 6. *PLoS ONE*, **8**, e64967.
Kuhl, P.K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, USA*, **100**, 9096–9101.
Liu, H.-M., Kuhl, P.K., & Tsao, F.-M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, **6**, F1–F10.
Locke, J.L. (2006). Parental selection of vocal behavior: crying, cooing, babbling, and the evolution of language. *Human Nature*, **17**, 155–168.
Mehl, M.R., Gosling, S.D., & Pennebaker, J.W. (2006). Personality in its natural habitat: manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology*, **90**, 862–877.
Mehl, M.R., Vazire, S., Ramírez-Esparza, N., Slatcher, R.B., & Pennebaker, J.W. (2007). Are women really more talkative than men? *Science*, **317**, 82.
Murray, L., Halligan, S.L., & Cooper, P.J. (2010). Effects of postnatal depression on mother–infant interactions, and child development. In T. Wachs & G. Bremner (Eds.), *Handbook of infant development* (pp. 192–220). New York: Wiley-Blackwell.
Oller, D.K., Eilers, R.E., Neal, A.R., & Schwartz, H.K. (1999). Precursors to speech in infancy: the prediction of speech and language disorders. *Journal of Communication Disorders*, **32**, 223–246.
Phillips, D.A., & Lowenstein, A.E. (2011). Early care, education, and child development. *Annual Review of Psychology*, **62**, 483–500.
Preacher, K.J., & Hayes, A.F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator modes. *Behavior Research Methods*, **40**, 879–891.

Raizada, R.D.S., Richards, T.L., Meltzoff, A., & Kuhl, P.K. (2008). Socioeconomic status predicts hemispheric specialisation of the left inferior frontal gyrus in young children. *NeuroImage*, **40**, 1392–1401.

Ramírez-Esparza, N., Mehl, M.R., Álvarez-Bermúdez, J., & Pennebaker, J.W. (2009). Are Mexicans more or less sociable than Americans? Insights from a naturalistic observation study. *Journal of Research in Personality*, **43**, 1–7.

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.

Rowe, M.L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science*, **323**, 951–953.

Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: uses in assessing reliability. *Psychological Bulletin*, **86**, 420–428.

Stoel-Gammon, C., & Sosa, A.V. (2009). Phonological development. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 238–256). New York: Wiley-Blackwell.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

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.

Weikum, W.M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker, J.F. (2007). Visual language discrimination in infancy. *Science*, **316**, 1159.

Received: 7 February 2012
Accepted: 8 December 2013