

ARTICLE

# Parentese in infancy predicts 5-year language complexity and conversational turns

Naja FERJAN RAMÍREZ<sup>1,2</sup>, Yael WEISS<sup>2</sup>, Kaveri K. SHETH<sup>1</sup> and Patricia K. KUHL<sup>2,3</sup>

<sup>1</sup>Department of Linguistics, University of Washington, Seattle, WA, USA

<sup>2</sup>Institute for Learning & Brain Sciences, University of Washington, Seattle, WA, USA

<sup>3</sup>Department of Speech and Hearing Sciences, University of Washington, Seattle, WA, USA

Corresponding author. Naja Ferjan Ramirez, Email: [naja@uw.edu](mailto:naja@uw.edu)

(Received 14 June 2022; revised 20 December 2022; accepted 06 January 2023)

## Abstract

Parental input is considered a key predictor of language achievement during the first years of life, yet relatively few studies have assessed its effects on longer-term outcomes. We assess the effects of parental quantity of speech, use of parentese (the acoustically exaggerated, clear, and higher-pitched speech), and turn-taking in infancy, on child language at 5 years. Using a longitudinal dataset of daylong LENA recordings collected with the same group of English-speaking infants (N=44) at 6, 10, 14, 18, 24 months and then again at 5 years, we demonstrate that parents' *consistent* (defined as stable and high) use of parentese in infancy was a potent predictor of lexical diversity, mean length of utterance, and frequency of conversational turn-taking between children and adults at Kindergarten entry. Together, these findings highlight the potential importance of a high-quality language learning environment in infancy for success at the start of formal schooling.

**Keywords:** language input; parentese; conversational turns; LENA; language development predictors

## Introduction

Language acquisition is one of the most critical developmental milestones in early childhood. It is central to learning, socializing and forming relationships. In the United States, Kindergarten language scores have repeatedly been shown to be the single best predictor of school achievement in third and fifth grade (Durham, Farkas, Hammer, Tomblin & Catts, 2007; Pace, Burchinal, Alper, Hirsh-Pasek & Golinkoff, 2019), and children who have better language skills are more successful in regulating their emotions (Cole, Armstrong & Pemberton, 2010). One factor known to influence language development is a child's language learning environment. The seminal work by Hart and Risley (1995) demonstrated that the total amount of speech heard by an infant is highly correlated with their language outcomes. Children whose parents talk less tend to have smaller vocabularies by the time they are three years old. This difference, known as the "30-million-word gap", predicts children's IQ scores and academic success in grade school (Golinkoff, Hoff, Rowe, Tamis-LeMonda & Hirsh-Pasek, 2019; Hart & Risley, 1995).

© The Author(s), 2023. Published by Cambridge University Press.

Until recently, studies examining children's language input relied on time-consuming transcription of parental and child language, limiting the amount of data that could be collected and analyzed. Technological advances now allow for longer, more ecologically valid recordings of children's naturalistic language environments. In recent years, one of the main approaches for measuring children's language input is the Language Environment Analysis (LENA; Greenwood, Thiemann-Bourque, Walker, Buzhardt & Gilkerson, 2011). An important advantage of LENA is that it facilitates audio recordings in children's natural environments on a day-long timescale, and is supplemented by automated speech analyses. Studies using LENA have confirmed significant variation in the amount of language children experience in association with parental language input, though the size of the "word gap" has recently been proposed to be substantially smaller than 30-million words (i.e., around 4 million words; see Gilkerson, Richards, Warren, Montgomery, Greenwood, Oller, Hansen & Paul, 2017). In addition, studies following these key discoveries have noted that parental language behaviors are influenced by a variety of social and cultural factors, such as policies, beliefs, values and political systems, among others (see Rowe & Weisleder, 2020 for a recent review). Together, this large and growing body of research has come to a more fine-tuned conclusion: Language develops in context. Children learn the language(s) that are used around them, and their early social interactions with language shape their language learning trajectories. Studies conducted in industrialized countries, such as the United States, demonstrate that the number of words that infants hear alone is insufficient to account for the observed variation in children's language development; the quality of language input also needs to be considered (see Hirsh-Pasek, Adamson, Bakeman, Owen, Golinkoff, Pace, Yust & Suma, 2015; Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola & Nelson, 2008; Rowe, 2012; Tamis-LeMonda, Kuchirko & Song, 2014; Weisleder & Fernald, 2013).

Unlike input quantity, which is straightforward to measure (i.e., typically, researchers use adult word counts), the quality of caregiver language input can be measured in various ways, and specific aspects of input may be more or less important depending on the child's age and/or level of language development. Rowe and Snow (2020) have recently conceptualized the features of high-quality caregiver input that facilitate language development in terms of three dimensions: linguistic, interactive, and conceptual. Linguistic features include levels of linguistic complexity, repetition, and redundancy that are adapted to the child's age or developmental stage. Interactive features include periods of joint attention, interactive play, parental responsiveness, and reciprocity. Conceptual features include topics of conversation that provide appropriate challenges for the child's age or developmental stage. According to this model, learning is optimal when each of the three dimensions is maximized, and may be hindered if any dimension is minimized. Within this framework, an important goal of developmental language research is to identify aspects of language input that work across these three dimensions, as they may be particularly potent predictors of later language abilities.

Social-interactionist and sociocultural theories have long-emphasized the importance of children's early social experiences for language development, showing that infants (children under age 2 years) benefit enormously from the social and interactional features of language input (e.g., Bruner, 1981; Kuhl, 2007; Snow, 1977a; Vygotsky, 1979). One key feature of social language interactions that could serve as an "ideal" language learning signal in the first 24 months is parentese, the acoustically exaggerated, clear, and higher-pitched speech produced by adults when they address infants. Initially termed "baby talk" (Ferguson, 1964), parentese is distinguished from adult-directed speech (ADS) by a variety of segmental and prosodic features, including higher overall pitch and wider pitch

range, slowed speech rate, exaggerated intonation contours, fewer and simpler lexical items, shorter utterances, and longer pauses between phrases (Fernald, 1985; Fernald & Simon, 1984; Fernald, Taeschner, Dunn, Papousek, de Boysson-Bardies & Fukui, 1989; Garnica, 1977; Grieser & Kuhl, 1988; Stern, Spieker, Barnett & MacKain, 1983). Parentese is used across cultures in spoken and signed languages by parents, grandparents, siblings, teachers, and adults who do not have their own children (Ferguson, 1964; Jacobson, Boersma, Fields, Olson & David, 1983; Kuhl, Andruski, Chistovich, Chistovich, Kozhevnikova, Ryskina, Stolyarova, Sundberg & Lacerda, 1997; Reilly & Bellugi, 1996). For example, one study showed that mothers in the United States, Sweden, and Russia produce acoustically more extreme vowels when addressing their infants than they did when addressing adults, resulting in an expanded vowel space, providing exceptionally well-specified information about the building blocks of words (Kuhl et al., 1997). Other studies have noted the existence of prosodic features of parentese in languages such as German, Mandarin, Tamil, Tagalog, and Korean (Fernald & Simon, 1984; Grieser & Kuhl, 1988; Narayan & McDermott, 2016). From early on, it was noted that a similar speaking style is also used in other circumstances, such as when addressing a family dog (Hirsh-Pasek & Treiman, 1982; Mitchell, 2001) or in conversations with foreigners (Snow, van Eeden & Muysken, 1981), leading some to wonder whether parentese is a misnomer. However, there is now clear evidence that parentese has some unique properties that distinguish it from other speech registers. For instance, parentese, but not dog-directed speech, is characterized by vowel hyperarticulation (Burnham, Kitamura & Vollmer-Conna, 2002; Gergely, Faragó, Galambos & Topál, 2017; Kuhl et al., 1997); along similar lines, foreigner-directed speech lacks the high pitch and positive affect characteristics of parentese (Singh, Morgan & Best, 2002; Uther, Knoll & Burnham, 2007).

It has also long been known that infants prefer parentese over standard ADS from as early as two days after birth (Cooper & Aslin, 1990; see also Fernald, 1985; Fernald & Kuhl, 1987). Infants' preference for parentese over ADS has recently been confirmed in two large-scale studies across cultures, procedures, languages, and laboratories, in monolingual and bilingual infants (Byers-Heinlein, Tsui, Bergmann, Black, Brown, Carbajal, Durrant, Fennell, Fiévet, Frank, Gampe, Gervain, Gonzalez-Gomez, Hamlin, Havron, Hernik, Kerr, Killam, Klassen, Kosie, Kovács, Lew-Williams, Liu, Mani, Marino, Mastroberardino, Mateu, Noble, Orena, Polka, Potter, Schreiner, Singh, Soderstrom, Sundara, Waddell, Werker & Wermelinger, 2021; The ManyBabies Consortium, 2020). While early scholars warned that caregiver use of parentese may be damaging to children's language development (McCarthy, 1954), further investigation demonstrated that parentese was fully grammatical, used phonology that avoided complex clusters of consonants (Newport, Gleitman & Gleitman, 1977; Phillips, 1973; Snow, 1977b) and vowels that were temporally and spectrally expanded (Burnham et al., 2002; Kuhl et al., 1997). Later laboratory research demonstrated that parentese facilitated infants' word segmentation (Thiessen, Hill & Saffran, 2005), word recognition (Singh, Nestor, Parikh & Yull, 2009), and fast mapping (Ma, Golinkoff, Houston & Hirsh-Pasek, 2011).

Studies using daylong recordings have allowed researchers to study parentese as it occurs naturally in infants' day-to-day lives. Studies conducted in the United States with monolingual English-speaking families and bilingual Spanish-English speaking families have shown that most parents use parentese; however, its frequency varies widely, even in families where infants experience high rates of adult talk (Ramírez-Esparza, García-Sierra & Kuhl, 2014, 2016, 2017; see also Ferjan Ramírez, Hippe, Correa, Andert & Baralt, 2022; Shapiro, Hippe & Ferjan Ramírez, 2021). That is, some infants receive most of their language input through parentese, while other infants experience parentese relatively

infrequently. Importantly, these studies also demonstrated that higher rates of parentese use in the homes of monolingual and bilingual infants are associated with higher rates of child babbling at one year of age and greater productive vocabularies at 24 and 33 months of age (Ramírez-Esparza et al., 2014, 2016, 2017; see also Ferjan Ramírez et al., 2022; Shapiro et al., 2021). A recent intervention study (Ferjan Ramírez, Lytle, Fish & Kuhl, 2018; Ferjan Ramírez, Lytle & Kuhl, 2020) suggests that the links between caregiver use of parentese and child language learning may be causal. In the Ferjan Ramírez et al. (2020) study, parent coaching increased the rates of caregiver parentese use from 6 to 18 months, and this increase was associated with enhanced growth in infant babbling from 6 to 14 months (Ferjan Ramírez et al., 2018), and greater word production at 18 months (Ferjan Ramírez et al., 2020). Of note, turn-taking as measured automatically by the LENA technology was also enhanced in this intervention.

Taken together, a large and growing body of research suggests that parentese may represent an ideal high-quality signal for language learning in infancy. This is not surprising, considering that parentese has been demonstrated to maximize all three key quality dimensions proposed in the model by Rowe and Snow (2020). Linguistically, key features of parentese are its distinct segmental and prosodic features, which are adjusted to the child's level in real-time and in accordance with the child's responses. Interactive features of parentese include the behaviors that frequently co-occur with its use, such as eye-gaze, joint attention, interactive play, reciprocity, contiguity, and connectedness. Conceptually, caregivers tend to use parentese to talk about what is happening in the "here and now", with frequent pointing and reference to objects or events present in the child's immediate environment.

Recent studies also suggest an association between parental use of parentese and the frequency of caregiver-infant back-and-forth exchanges (conversational turns), another key mechanism supporting infants' language uptake (Levinson, 2016). Unlike overheard speech or speech from an electronic source (Kuhl, Tsao & Liu, 2003; Shneidman, Arroyo, Levine & Goldin-Meadow, 2013; Weisleder & Fernald, 2013), turn-taking allows caregivers to provide contingent feedback adjusted to their infant's linguistic needs. For example, research has shown that mothers adjust their speech in accordance with their infants' responses (Braarud & Stormark, 2008; Smith & Trainor, 2008). Infants, in turn, adjust their vocalizations, thereby creating a feedback loop that supports language growth (Warlaumont, Richards, Gilkerson & Oller, 2014). Turn-taking provides opportunities for temporal contiguity and contingency and joint engagement between parents and children, which are critical in word learning and predict children's subsequent language skills (Bornstein, Tamis-LeMonda & Haynes, 1999; Conboy, Brooks, Meltzoff & Kuhl, 2015; Hirsh-Pasek et al., 2015; Tamis-LeMonda et al., 2014). Finally, recent brain studies propose that, through contiguity, contingency, connectedness, and social feedback, turn-taking shapes the social circuitry of the language-related brain areas (Merz, Maskus, Melvin, He & Noble, 2020; Romeo, Leonard, Robinson, West, Mackey, Rowe & Gabrieli, 2018a; Romeo, Segaran, Leonard, Robinson, West, Mackey, Yendiki, Rowe & Gabrieli, 2018b).

While there is evidence supporting the short-term benefits of parentese and turn-taking to infant and toddler language learning, the extent to which their consistent use in infancy is associated with longer-term language outcomes remains less clear. While we know that most parents in the United States use parentese when their infants are between 6 and 24 months of age, we also know that there is quite a bit of variability from family to family, and potentially, also within families from one developmental time point to the next (see Shapiro et al., 2021). Within a cascade model of development (Landry,

Smith & Swank, 2003; Tamis-LeMonda, Luo, McFadden, Bandel & Vallotton, 2019), a potential hypothesis is that language outcomes observed in later childhood (i.e., at Kindergarten entry) are a reflection of the CUMULATIVE impact of language input across early childhood. For example, one study examined the relation between consistency of maternal responsiveness across early childhood (birth to preschool years) and children's language outcomes at age 8 years in a sample of preterm and term children. Results demonstrated that children who experienced consistently high levels of responsivity across the first four years of life scored higher on language measures at the age of 8 years compared to those children whose maternal responsivity scores were less consistent (Landry et al., 2003). With these findings in mind, one hypothesis is that CONSISTENT use of high rates of parentese and/or turn-taking that are stable over time could contribute to longer-term positive outcomes. In another study, Gilkerson and colleagues (2018) show that parent-infant conversational turns during a narrow time-window of 18-24 months predicted children's language scores 10 years later. However, a recent study demonstrates a significant association between parental use of parentese in the first year of life (6-14 months) and parent-infant turn-taking at the age of 18 months (Ferjan Ramírez et al., 2018), suggesting that the stepping stone to later language may be either consistent use of parentese in infancy, consistent turn-taking in infancy, or a combination of the two. Within a cascade model of development, one can hypothesize that when infants are raised in consistently rich, high quality language learning environments, these early experiences instigate a drive on the infant's part to respond and join in on the conversation, activating developmental cascades. This would suggest that parental consistent use of high parentese rates and/or turn-taking in infancy may benefit children's language development to the extent that if a child develops sufficiently robust early language skills, this can continue to "drive" language-learning interactions, furthering the child's lexical and grammatical growth. In children aged 4-6 years, the quantity of parent-child conversational turns has been linked not only to children's cognitive performance, but also to the function and structure of their language-related brain networks (Romeo et al., 2018a, 2018b; Romeo, Leonard, Grotzinger, Robinson, Takada, Mackey, Scherer, Rowe, West & Gabrieli, 2021). However, it is currently unknown whether turn-taking in the preschool years (4-6 years) can be linked to consistent parental use of parentese and/or turn-taking in infancy.

Taken together, the developmental cascades model supports the idea that consistent caregiver use of frequent parentese and/or turn-taking in infancy could serve as a stepping stone to language, catalyzing positive developmental cascades, and supporting longer term robust language development beyond toddlerhood. However, no studies thus far have examined whether consistent parental use of parentese and/or turn-taking in infancy is prospectively associated with children's later language outcomes, such as grammatical complexity and lexical diversity at Kindergarten entry, or parent-child conversational exchanges at Kindergarten entry, the age at which such exchanges have been shown to predict children's cognitive skills, brain structure and brain function.

### The present study

In the present study, we examine whether consistent parental use of parentese and/or turn-taking in infancy is associated with children's language complexity and parent-child turn-taking at Kindergarten entry (age 5 years). To answer this question, we collected longitudinal data from 44 children that included:

## 1) In infancy (from 6-24-months of age):

measures of parental input quantity (adult word counts, AWC);  
 measures of parentese;  
 measures of parent-infant turn-taking;  
 measures of infants' spoken language,  
 from daylong home language recordings collected at 6, 10, 14, 18 and 24 months;

## 2) From the same 44 children at the age of 5 years:

measures of parent-child turn-taking;  
 measures of grammatical complexity (utterance length);  
 measures of lexical diversity, manually coded from daylong home language recordings.

Based on previous research (Ferjan Ramírez et al., 2018, 2020; Gilkerson et al., 2018), we hypothesized that *CONSISTENCY* in parental use of parentese and turn-taking, but not their overall volume of speech (AWC), would predict measures of speech complexity and parent-child turn taking at Kindergarten entry. By testing these hypotheses, the goal of the present study is to highlight specific aspects of early interactions associated with later language success.

## Methods

### *Participants*

Seventy-nine families were recruited through the University of Washington Subject Pool as part of a previously published home language intervention study between 6 and 24 months of age (Ferjan Ramírez et al., 2018; recruitment period: September 2016–January 2017). In this intervention, families provided informed consent and were then randomly assigned to either a Coaching Group or a passive Control Group when the babies were 6 months of age. The intervention parents then received “coaching” to enhance their use of parentese speaking style and turn-taking between 6 and 18 months of age, while the control group did not receive such coaching. All families completed audio recordings of their children and environment at five time points in infancy (when infants were 6, 10, 14, 18, and 24 months old). The recordings were employed with the LENA system (Version 3.4.0; LENA, 2015), which provides audio recordings and measures of different components in children’s natural environments (Ferjan Ramírez et al., 2018, 2020; Ferjan Ramírez, Hippe & Kuhl, 2021). All parents in this study (Coaching Group parents and Control Group parents) used parentese and turn-taking between 6 and 24 months of age, though their frequency varied from family to family, and was enhanced in the Coaching Group (Ferjan Ramírez et al., 2020). About 2.5 years after completing the infancy home language intervention, families in both the Coaching and Control groups who agreed to be re-contacted for future research were invited to return for a new, follow-up LENA study when their child was 5 years old (re-recruitment period: February–June 2021). All families who agreed to participate in the follow-up (N=70) completed a phone screening interview to determine whether their children met the following criteria:

- (1) Child not yet enrolled in Kindergarten, aged between 5 years and 5 years and 4 months;
- (2) English is still the primary language of communication in the home;
- (3) Child had no apparent congenital, neurological or physical abnormalities.

Exclusion criteria included any brain injury and medications that impact cognition; intellectual disability; Autism Spectrum Disorder; mood disorders; significant and permanent hearing impairments. After the initial screening process, 52 eligible participants were invited to take part in the follow-up study. Of these, 49 families completed both days of the follow-up LENA recording at 5 years. Of these 49 families, 44 also had all infancy language recordings (i.e., at 6, 10, 14, 18, and 24 months, with no missing data), and were included in the present sample. Of these 44 families (25 with girls, 19 with boys), 33 were in the “Coaching Group” in the Infancy Intervention, and 11 were in the “Control Group” (Ferjan Ramírez et al., 2018, 2020). Note that data collection at age 5 years was NOT pre-planned (i.e., families had to be re-contacted and re-enrolled). As might be expected, this resulted in the present study enrolling a disproportionately low number of participants who were part of the “Control Group” in the Infancy intervention, parents who were contacted less over the course of the study due to no coaching (i.e., 11 participants, 25% of the sample from the “Control Group”, vs. 33 families, 75% of the sample from the original “Coaching Group”). Importantly however, the two parental language behaviors that were manipulated by the Infancy Intervention (parentese and turn-taking between 6 and 24 months) did not differ between the 33 participating children enrolled in the infancy study as “Intervention” and the 11 participating children originally enrolled in the infancy study as “Control” (both  $ps > .05$ ), even though the same two behaviors were demonstrated to be enhanced by the Infancy Intervention when the whole sample of 77 participants was considered (Ferjan Ramírez et al., 2018, 2020). For this reason, and because the goal of the present study was to examine the correlation between language input in infancy and language outcomes at age 5 years, the data were analyzed for the whole available sample as a single group of 44 families. Socio-economic status (SES) was measured with the Hollingshead Index and ranged from 30 to 66 in the final analyzed sample, ( $M = 49.4$ ,  $SD = 10.5$ ) (i.e., working- to upper-middle class families).

### *Data collection, preparation, and annotation*

The home Language Environment Analysis System (LENA) was used to collect naturalistic first-person recordings from all families over two weekend days when children were 6-, 10-, 14-, 18-, 24- months old (Infancy) and then again at 5 years of age. At each timepoint, all families received two LENA recorders in the mail and were instructed to use one recorder on each day of a “typical” weekend, defined as two consecutive days when both parents were home and not working. Parents were asked to start each recording in the morning when the child woke up, go about their day as usual, and turn off the recorder at night when the child went to sleep. Recordings across all ages varied in length between 8:00 and 16:00 hours, with an average of 12 hrs and 47 mins at 6 months, 13 hrs and 4 mins at 10 months, 13 hrs at 14 months, 12 hrs and 42 minutes at 18 months, 12 hours and 50 minutes at 24 months, and 13 hours and 17 minutes at 5 years.

Parent and child speech were quantified through a combination of automatic annotation by LENA software and manual (human) annotation. The LENA software produces an automatic count of child vocalizations (child vocalization count, CVC), words produced by nearby adults (adult word count, AWC), and adult-child conversational turns (conversational turn count, CTC). Recent studies have sought to assess and validate LENA’s classification performance (e.g., Cristia, Bulgarelli & Bergelson, 2020; Lehet, Arjmandi, Houston & Dilley, 2020; Wang, Williams, Dilley & Houston, 2020). According to one meta-analysis, LENA achieves a mean recall and precision of 0.59 and 0.68,

respectively, for recognizing adult words and a mean recall of 0.77 for recognizing child vocalizations (Cristia et al., 2020). LENA's CTC measure looks for adult and child speech in close temporal proximity – but, critically, without differentiating between child-directed and overheard speech. This means that an unknown proportion of LENA's CTCs are identified in error, such as when a parent is talking on the phone and the infant is babbling to herself nearby (i.e., accidental contiguity). The frequency of accidental contiguity has recently been shown to be high (Ferjan Ramírez et al., 2021), leading us to limit our analysis of turn-taking to manually identified conversational turns. However, we do rely on LENA's automatic identified AWC and CVC (see Table 1 for definitions). AWCs are used to approximate the amount of adult speech heard by the infant, and CVCs are used as a measure of child volubility in infancy. Note that AWC is an estimate of ALL speech occurring near the child wearing the recorder, including all adult speech directed to the child and all adult speech overheard by the child. CVC is an estimate of ANY articulations that originate from the vocal tract of the child, except for fixed signals (screams, cries), sounds related to respiration (breaths) or digestion (burps; see Xu, Richards & Gilkerson, 2014; Gilkerson et al., 2017 for a more detailed description of

**Table 1.** Language variable names, types, and definitions. AWC = Adult Word Count; CTC = Conversational Turn Count; CVC = Child Vocalization Count; MLU-m: Mean Length of Utterance in morphemes; LexDiv = Lexical Diversity; LENA = LENA estimate; Manual = manually coded.

Variable name	Variable type	Agess measured	Variable definition
CVC	Automatic	6, 10, 14, 18, 24 mo	Number of vocalizations containing speech-related activity produced by the child wearing the recorder, estimated automatically by LENA and averaged over two recording days. Child vocalizations can be of any length, as long as they are surrounded by 300+ ms of non-speech.
AWC	Automatic	6, 10, 14, 18, 24 mo	Total number of adult words heard by the child during the recording, estimated automatically by LENA and averaged over two recording days.
%Parentese	Manual	6, 10, 14, 18, 24 mo	Percent of segments where mother, father, or other adult spoke directly to the infant, parentese speech style was used (high pitch, larger pitch range), and one or more than one adult voice was recorded during the interval.
CTC	Manual	6, 10, 14, 18, 24 mo; 5 years	Total number of adult utterances directed to child, followed within 5 seconds by child utterances directed to adult, or vice versa; counted in discrete pairs (child-to-parent=1 turn, parent-to-child-to-parent=1 turn, child-to-parent-to-child-to-parent=2 turns; See Ferjan Ramírez et al., 2021)
LexDiv	Manual	5 years	Mean number of unique lexical types across 100 C-units per participant.
MLU	Manual	5 years	Mean number of morphemes per C-unit across 100 C-units per participant.



these metrics). Early communication skills are a strong predictor of later language ability, so CVC was considered in order to explore whether children's own volume of speech-related vocalizations between 6 and 24 months may predict later language outcomes. However, note that CVC is a purely quantitative measure that does NOT differentiate between different kinds of language-related vocalizations that infants are known to produce between 6 and 24 months (i.e., a vocalization such as "ba" and a full multi-word utterance such as "I want a cookie" would both be counted as one CV). A higher CVC value thus indicates a higher volume of child speech, but not necessarily more complex speech.

The LENA audio files were further processed using the Advanced Data Extractor Tool (ADEX) for the purposes of manual annotation. For identification of parentese (6-24 months) and conversational turns (6-24 months and 5 years), the same procedures were followed as in previously published studies (Ferjan Ramírez et al., 2018, 2020, 2021; Ramírez-Esparza et al., 2014, 2016, 2017). In brief, ADEX was used to identify intervals with the language activity of interest (high AWC), in order to avoid coding when there is no social or linguistic activity (for example, during naps). Each participant's two daily recordings were segmented into 30-second intervals. For each of the two recording days, 50 intervals with the highest adult word count that were at least 3 minutes apart were selected, yielding a total of 100 30-second coding intervals per participant. Ten research assistants listened to each 30-second interval and determined the presence or absence of parentese speaking style, and counted the number of conversational turns (CTs), using the same audio files, training, and reliability assessment as described by Ramírez-Esparza and colleagues (2014, 2016, 2017) and Ferjan Ramírez and colleagues (2018, 2020, 2021). Note that child-directed speech consists of parentese and standard speech. The focus of the present study is on the long-term effects of parentese, which is distinguished from standard child-directed speech by its acoustic features. To identify parentese and distinguish it from standard child-directed speech, the same criteria were adopted as described previously by Ramírez-Esparza and colleagues (2014, 2016, 2017) and Ferjan Ramírez and colleagues (2018, 2020, 2021): Ramírez-Esparza et al. (2014) independently verified that the intervals defined as parentese or standard speech contained the acoustic differences characteristic of these two speech styles (i.e., higher pitch and larger pitch range for parentese). In these analyses, 60 occurrences of the word 'you' were analyzed. The 60 occurrences of 'you' represented 30 pairs (30 produced as parentese and 30 as standard speech) produced by the same adult addressing the same infant. Mean pitch and pitch range were significantly higher for parentese than standard speech ( $p < 0.001$ ); see Table 1 for variable definitions).

For tabulating the number of CTs within each 30-s segment, the same procedures were followed as in Ferjan Ramírez et al. (2021). In brief, as with the LENA algorithm, CTs were counted in discrete pairs, and pauses of 5s or more constituted the end of a conversation. Critically, and unlike the LENA algorithm, cases of accidental contiguity were not counted as CTs. The total number of CTs was then counted across all 100 intervals for each participant. After training, all coders were tested independently with a training file from Ramírez-Esparza and colleagues (2014); and a training file from the present dataset, used to evaluate inter-coder reliability. The reliability analysis produced an average intra-class correlation of 95% for parentese and 97% for CTs. This indicates effective training and reliable coding.

Language complexity at 5 years was assessed by measuring children's mean length of utterance (MLU) in morphemes, and lexical diversity (LexDiv) by manually annotating the recordings collected at 5 years of age. The following procedures were used: for each

child, both 5-year daylong audio recordings were segmented into one hour-long segments. Using ADEX, the hours from both days were then arranged from highest to lowest in CVC. Recall that CVC represents an estimated number of meaningful child speech utterances of any length, and can include babbling, individual words, or sentences. The hour with the highest CVC was used for analyses of MLU and LexDiv in order to be able to consistently obtain a sample of 100 consecutive utterances for analyses from each child.

Nine research assistants, all native speakers of English, transcribed the dataset. They were first trained on the protocol, transcribed two training files, and received feedback on their transcription from their supervisor, the third author of the manuscript (K. Sheth). Within the highest CVC hour selected for transcription, the utterances were segmented into communication units (C-units; Miller, Andriacchi & Nockerts, 2019), which include one main clause with all subordinate clauses and cannot be further divided without losing its meaning. For example, [*He went to the store because he was out of milk*] would be considered one C-unit, but [*He climbed up on the branches*] [*but they weren't branches*] would be considered two C-units were chosen as a consistent way for segmenting utterances across all participants. Following the standard Systematic Analysis of Language Transcripts (SALT) conventions (Miller et al., 2019), the first 100 C-units in the highest CVC hour were transcribed for each participant. Within these 100 C-units, transcribers also completed a morpheme transcription, marking the bound morphemes as specified in the standard SALT Conventions (Miller et al., 2019). For example, to indicate regular plural inflections, the bound morpheme /s was used (EX: frog/s, tree/s). To indicate third person singular verb inflections, the bound morpheme /3s was used (EX: he look/3s, she jump/3s). Each 100-utterance transcription was then checked by a second transcriber and marked for any disagreements or errors. The transcription supervisor (K. Sheth) then resolved any disagreements. Utterances with excessive noise that were impossible to reliably transcribe were removed from analyses. In two participants, the entire highest CVC hour was too noisy to allow for reliable transcription (i.e., lots of overlapping voices from other children). In both cases, the second highest CVC hour was used for analyses instead.

For each participant, the 100 transcribed C-units were then further analyzed to derive the MLU in morphemes and LexDiv. MLU was calculated using the morpheme transcription and utilized formulas in Microsoft Excel to get the average number of morphemes per C-unit across the 100 C-units for each participant. LexDiv represents the average number of unique lexical types across all 100 C-units per participant, and was calculated utilizing a Python (Version 3.8) (Van Rossum & Drake, 2009) script with Natural Language ToolKit (NLTK) (Bird, Loper & Klein, 2009). For example, *he goes* counts as two unique words. A subsequent occurrence of *he* or *goes* in the same participant would not be counted as additional unique lexical items, but words such as *go*, *went*, *him*, or *his* would count as additional unique lexical items. What this means is that our measure of unique lexical types refers to the words and not the lemmas.

Because the LENA recordings varied in duration, projected 12-h values were used for all LENA automatic measures (AWC, CVC). The 12-h projections are generated automatically for recordings at least 10h in length, and represent the interpolated values for AWC and CVC at the 12-h mark for the day's recording. The values for "manual" variables are based on 100 30-s segments for parentese and CTC, and on 100 consecutive utterances for MLU and LexDiv, annotated by humans as described above.

## Results

### Descriptive statistics

The descriptive statistics for all raw data can be found in Table 2. Across all infancy timepoints (6, 10, 14, 18, and 24 months), infants heard an estimated average of 16,424

**Table 2.** Variables and their distributions. Abbreviations: SD, Standard Deviation; CI, Confidence Interval; SES, Socioeconomic status; mo, months; CVC, Child Vocalization Count; AWC, Adult Word Count; CTC, Conversational Turn Count; yr, year; LexDiv, Lexical Diversity; MLU, Mean Length of Utterance. Parentese is entered as the proportion of coded intervals.

Variable	Mean	SD	Range	95% CI
SES	49.41	10.45	30–66	46.23–52.59
Parentese 6mo	41.32	17.26	13–86	36.07–46.56
Parentese 10mo	48.66	19.01	12–93	42.88–54.44
Parentese 14mo	53.07	19.19	21–92	47.23–58.90
Parentese 18mo	58.36	20.60	18–89	52.10–64.63
Parentese 24mo	69.93	15.19	33–94	65.31–74.55
GLOBAL Parentese	271.34	77.93	133–404	247.65–295.03
CVC 6mo	1195.52	338.46	689–2052	1092.62–1298.42
CVC 10mo	1333.90	355.09	708–2341	1225.94–1441.85
CVC 14mo	1320.30	444.64	450–2181	1185.11–1455.48
CVC 18mo	2040.40	745.11	663–5062	1813.86–2266.93
CVC 24mo	2847.02	873.44	1154–5120	2581.47–3112.57
GLOBAL CVC	8737.14	1780.29	5476–12885	8195.88–9278.39
AWC 6mo	16893.57	6688.80	4573–28641	14859.99–18927.15
AWC 10mo	16272.36	7250.50	3775–34225	14068.01–18476.72
AWC 14mo	14978.45	5303.64	6809–27285	13366.00–16590.91
AWC 18mo	16667.65	5508.00	5070–31288	14993.06–18342.23
AWC 24mo	17310.11	5879.86	6760–26668	15522.47–19097.76
GLOBAL AWC	82122.15	24661.64	36547–125414	74624.32–89619.98
CTC 6mo	21.68	13.31	6–74	17.64–25.73
CTC 10mo	47.61	25.32	7–129	39.91–55.31
CTC 14mo	79.73	41.87	17–197	67.00–92.46
CTC 18mo	127.02	60.56	31–303	108.61–145.44
CTC 24mo	211.52	62.12	111–341	192.64–230.41
GLOBAL CTC	487.57	151.91	257–955	441.38–533.75
5yr CTC	168.52	62.66	44–316	149.47–187.57
5yr LexDiv	175.30	24.24	120–252	167.93–182.66
5yr MLU	4.75	0.79	3.41–6.98	4.51–4.99

words per day, and produced an average of 1,747 vocalizations per day. An average of 54% of the coded intervals contained parentese speaking style.

We created a correlation matrix between parentese, CTC, and AWC measured at 6, 10, 14, 18, and 24 months, and the three 5-year outcome measures: CTC, LexDiv and MLU (Table 3). We use  $\alpha = 0.05$  as denoting statistical significance, and  $\alpha = 0.1$  as denoting marginal significance. With the sample size of 44 and a 0.8 power, both of these  $\alpha$  values allow us to reliably capture medium to large effect sizes (Serdar, Cihan, Yücel & Serdar, 2021).

Parental parentese at 6 months, 10 months, 14 months, and 18 months was significantly correlated with 5-year CTC. At 24 months, the correlation between parental parentese and 5-year CTC was marginally significant. Further, parental parentese at 18 months and 24 months was significantly correlated with 5-year LexDiv. At 6 months, the correlation between parental parentese and 5-year LexDiv was marginally significant. Finally, parental parentese at 18 months was significantly correlated with 5-year MLU. At

**Table 3.** Correlations between parentese, CTC, and AWC measured at 6, 10, 14, 18, and 24 months, and the three 5-year measures: 5yr CTC, 5yr LexDiv, and 5yr MLU. N=44 \*\*  $p < 0.05$ , \*  $0.05 < p < 0.1$  (marginal significance). Abbreviations: mo, months; AWC, Adult Word Count; CTC, Conversational Turn Count; yr, year; LexDiv, Lexical Diversity; MLU, Mean Length of Utterance.

	Parentese		CTC		AWC	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
6 mo						
5yr CTC	0.38	0.012**	0.10	0.506	0.22	0.153
5yr LexDiv	0.28	0.071*	0.07	0.664	0.07	0.636
5yr MLU	0.26	0.089*	-0.01	0.936	0.11	0.495
10 mo						
5yr CTC	0.38	0.012**	0.26	0.087*	-0.03	0.855
5yr LexDiv	0.24	0.115	0.06	0.713	-0.20	0.194
5yr MLU	0.24	0.111	0.03	0.872	-0.13	0.393
14 mo						
5yr CTC	0.34	0.024**	0.06	0.713	0.12	0.457
5yr LexDiv	0.21	0.165	-0.06	0.680	-0.09	0.571
5yr MLU	0.25	0.096*	0.3	0.835	-0.11	0.499
18 mo						
5yr CTC	0.53	0.000**	0.33	0.029**	0.23	0.141
5yr LexDiv	0.32	0.033**	0.06	0.721	0.10	0.513
5yr MLU	0.36	0.017**	0.06	0.720	0.05	0.764
24 mo						
5yr CTC	0.26	0.084*	0.34	0.023**	0.17	0.281
5yr LexDiv	0.35	0.020**	0.11	0.485	0.04	0.797
5yr MLU	0.20	0.189	0.17	0.278	0.09	0.585

6 months and 14 months, the correlation between parental parentese and 5-year LexDiv was marginally significant.

Because parental use of parentese was significantly or marginally significantly correlated with all three 5-year measures at one or more of the timepoints (Table 3), and one of our main questions pertained to the CONSISTENCY of parental parentese use, we calculated the “global” infancy parentese score, summing the parentese values across all 5 ages (6, 10, 14, 18, and 24 months, Table 2), which serves here as our proxy for “parentese consistency” between 6 and 24 months. To examine the association between parentese consistency between 6-24 months of age and 5-year language measures, we ran unadjusted linear regression models, including the “global parentese” variable, and the three 5-year measures (CTC, LexDiv, MLU). Next, adjusted models were run, which included child volubility across all ages in infancy (“global CVC”, Table 2) as a potential predictor of the 5-year measures. This step was done to examine whether children’s own volubility of speech related vocalizations in infancy may predict later language outcomes.

CTC at 18 and 24 months was significantly correlated with 5-year CTC. The correlation between 10 month CTC and 5-year CTC was marginally significantly (Table 3). There were no significant correlations between CTC in infancy and MLU or LexDiv at 5 years. To consider the hypothesis that CONSISTENT use of turn-taking in infancy would predict parent-child turn taking at Kindergarten entry, we considered the association between CTC summed across all 5 ages in infancy (“global CTC”; Table 2) and 5-year CTC, first in an unadjusted model, and then in adjusted models. CTC was not included in the models predicting LexDiv and MLU due to a lack of correlation between CTC and these two variables.

AWC at any one of the timepoints measured in infancy was not significantly or marginally significantly correlated with any one of the 5-year measures (Table 3), and was therefore not included in any one of the models. SES was also not significantly correlated with any one of the 5-year measures (all  $ps > .26$ ), and was not included in the models.

### Predicting 5-year CTC

Across all participating children, the mean number of conversational turns in 100 30-s segments at age 5 years was 168.52 (SD = 62.66, Table 2). The correlation between global parentese and 5-year CTC was significant ( $r = 0.45$ ,  $p = 0.002$ ), as was the correlation between global CTC and 5-year CTC ( $r = 0.34$ ,  $p = 0.024$ ). To evaluate whether these correlations could be accounted for by the child’s own language characteristics, we added simultaneously collected, potentially related measures of child language volubility (global CVC). Global CVC was not significantly correlated with global parentese ( $p = 0.55$ ), but was significantly correlated with global CTC, ( $r = 0.40$ ,  $p = 0.008$ ).

In Table 4, we display regression metrics for 6 models to predict turn-taking at age 5 years: Model 1: global parentese only; Model 2: global CTC only; Model 3: global CVC only; Model 4: global parentese plus global CTC; Model 5: global parentese plus global CVC; Model 6: global parentese plus global CTC plus global CVC. Note that we are comparing models with different numbers of independent variables. In this scenario, R-squared will increase every time an independent variable is added to the model (i.e., it never declines). By contrast, adjusted R-squared increases only when independent variable is significant and affects the dependent variable. Therefore, R-squared is used

**Table 4.** Multiple linear regression models used to predict 5-year CTCs from infancy measures of global parentese, global CTC, and global CVC. N=44. Abbreviations: STD Coef, Standardized Coefficient; UNSTD Coef, Unstandardized Coefficient; Adj, Adjusted; CTC, Conversational Turn Count; CVC, Child Vocalization Count.

	STD Coef	UNSTD Coef	t test value	p	Model R <sup>2</sup>	Model Adj R <sup>2</sup>
Model 1						
Global Parentese	.451	.363	3.275	.002	.203	.184
Model 2						
Global CTC	.340	.140	2.343	.024	.116	.095
Model 3						
Global CVC	.043	-.001	-.276	.784	.002	-.022
Model 4						
Global Parentese	.419	.337	2.140	.038	.204	.166
Global CTC	.045	.019	.230	.819		
Model 5						
Global Parentese	.459	.369	3.292	.002	.210	.172
Global CVC	-.085	-.003	-.607	.547		
Model 6						
Global Parentese	.372	.299	1.814	.077	.217	.158
Global CTC	.128	.053	.577	.567		
Global CVC	-.127	-.004	-.802	.427		

to judge which model is “best” (Schroeder, Sjoquist & Stephan, 2016). The best model in Table 4 is Model 1 (only global parentese), which demonstrates that 18.4% of the variance in 5-year CTC can be explained by parental use of parentese between 6 and 24 months of age (Table 4, Model 1). Adding CTC and/or CVC to the model did not improve the predictive power of global parentese (Table 4, Models 2-6).

### *Predicting 5-year Lexical Diversity*

Across all participating children, the mean number of different lexical items produced in 100 consecutive utterances was 175.3 (SD = 24.24; Table 2). Global parentese correlated significantly with 5-year LexDiv ( $r = 0.326$ ,  $p = 0.031$ ). However, results from multiple linear regression analyses predicting 5-year LexDiv scores from global parentese when controlling for CVC revealed that its addition to the model significantly improved the predictive power of global parentese. In Table 5, we display regression metrics for 3 models: Model 1: global parentese only; Model 2: global CVC only; Model 3: global parentese and global CVC. The best model is Model 3, which combines global parentese with CVC, and demonstrates that 17.2% of the variance in Lexical Diversity at age 5 years can be explained by a combination of parental parentese and child volubility between 6 and 24 months of age (Table 6, Model 3). Note, however, that the global

**Table 5.** Multiple linear regression models used to predict 5-year Lexical Diversity from infancy measures of global parentese and global CVC. N=44. Abbreviations: STD Coef, Standardized Coefficient; UNSTD Coef, Unstandardized Coefficient; Adj, Adjusted; CVC, Child Vocalization Count.

	STD Coef	UNSTD Coef	t test value	p	Model R <sup>2</sup>	Model Adj R <sup>2</sup>
Model 1						
Global Parentese	.326	.101	2.234	.031	.106	.085
Model 2						
Global CVC	-.291	-.004	-1.971	.055	.085	.063
Model 3						
Global Parentese	.356	.111	2.552	.015	.210	.172
Global CVC	-.324	-.004	-2.322	.025		

**Table 6.** Multiple linear regression models used to predict 5-year MLU from infancy measures of global parentese and global CVC. N=44. Abbreviations: STD Coef, Standardized Coefficient; UNSTD Coef, Unstandardized Coefficient; Adj, Adjusted; CVC, Child Vocalization Count.

	STD Coef	UNSTD Coef	t test value	p	Model R <sup>2</sup>	Model Adj R <sup>2</sup>
Model 1						
Global Parentese	.314	.003	2.14	.038	.098	.077
Model 2						
Global CVC	-.363	.000	-2.523	.015	.132	.111
Model 3						
Global Parentese	.350	.004	2.581	.006	.253	.217
Global CVC	-.395	.000	-2.914	.014		

CVC's standard coefficient is negative, meaning that children whose overall speech volubility between 6 and 24 months was LOWER had higher lexical diversity at age 5 years (see Discussion).

### Predicting 5-year MLU

Across all participating children, the mean MLU in morphemes was 4.75 (SD = 0.79; Table 2). Global parentese correlated significantly with 5-year MLU ( $r = 0.314, p = 0.038$ ). Results from multiple linear regression analyses predicting 5-year MLU scores from global parentese when controlling for CVC revealed that its addition to the model significantly improved the predictive power of global parentese. In Table 6, we display regression metrics for 3 models: Model 1: global parentese only; Model 2: global CVC only; Model 3: global parentese and global CVC. The best model is Model 3, which combines global parentese with CVC, and demonstrates that 21.7% of the variance in MLU at age 5 years can be explained by a combination of parental parentese and child volubility between 6 and 24 months of age (Table 6, Model 3). Note that the global CVC's

standard coefficient is negative, meaning that children whose overall speech volubility between 6 and 24 months was LOWER had higher MLU at age 5 years

## Discussion

The present study tested the hypothesis that consistent use of parentese and/or conversational turns in infancy predicts children's conversation and language outcomes at the age of 5 years, just prior to Kindergarten entry. Supporting our hypothesis, our results show that, at the age of 5 years, children whose parents consistently used high amounts of parentese in infancy demonstrated higher lexical diversity, produced longer sentences, and engaged in higher rates of turn-taking compared to children whose parents used lower amounts of parentese in infancy. Also confirming our hypotheses, the sheer quantity of adult speech in infancy (i.e., AWC) was not associated with children's conversation or language complexity outcomes at the age of 5 years. Together, these findings support the notion that language input in infancy is the foundation of later language skills (Hirsh-Pasek et al., 2015; Rodriguez & Tamis-LeMonda, 2011; Rowe, 2012), and that associations between qualitative aspects language input and later child language skills may be stronger than those between the sheer quantity of adult speech and child language skills (Hirsh-Pasek et al., 2015; Zimmerman, Gilkerson, Richards, Christakis, Xu, Gray & Yapanel, 2009). We extend previous findings by proposing that parentese may be a key component of high-quality language input in infancy, especially when delivered at the age at which infants are known to be sensitive to the acoustic properties of the speech signal (Kuhl, 2004), but may not yet be producing large quantities of speech themselves.

Given the present results, it is important to consider HOW and WHY parentese may be helpful for language learning in infancy and beyond. With its exaggerated acoustics, unique syntactic structure, accompanying social behaviors, and exaggerated facial movements (Werker, Pegg & McLeod, 1994), parentese conveys a positive emotion that makes the speaker sound "happy" (Singh et al., 2002). This combination attracts and holds infants' attention to the speaker and what they are saying, giving infants ample time to babble in response (Ferjan Ramírez et al., 2020). Through such exchanges, which may at first be sparse, and primarily driven by parental parentese, infants adjust their vocalizations when they respond to parental talk, and increase the complexity of their language (Bornstein et al., 1999; Braarud & Stormark, 2008; Goldstein & Schwade, 2008; Smith & Trainor, 2008). Parents, in turn, provide contingent feedback that is continuously adjusted to their child's linguistic needs, thereby creating a positive feedback loop that further promotes language growth through back and forth exchanges in toddlerhood and through the preschool years (Warlaumont et al., 2014). The quantity of parent-child conversational turns between 4-6 years of age, a variable corresponding closely to our 5-year CTC, has previously been linked to children's cognitive performance, as well as the functional and structural measures of language-related brain networks (Romeo et al., 2018a, 2018b, 2021). Our findings demonstrate that parental use of parentese between 6 and 24 months accounted for 18.4% of variance in 5-year CTC, suggesting the important role of parentese in infancy in establishing a strong communication foundation prior to Kindergarten entry.

The present study also demonstrates a link between parentese in infancy, particularly between 18 and 24 months, and lexical diversity and sentence length at age 5 years. Parentese may facilitate lexical acquisition by exaggerating acoustical properties such as



linguistic focus. For example, mothers of one-year old infants have been shown to highlight target words by placing an exaggerated pitch peak at sentence ends (Aslin, Woodward, LaMendola & Bever, 1996; Fernald & Mazzie, 1991). Parentese may also be helpful for identifying words because of a preponderance of prosodically isolated single words (Brent & Siskind, 2001). For example, Brent and Siskind found that these words cross the spectrum of grammatical categories, and are a significant predictor of vocabulary acquisition. Another possible role for phonological properties of parentese is in classifying words by grammatical category, which may be related to the presently-identified link between parentese in infancy and MLU at age 5 years. For example, Shi, Morgan, and Allopena (1998) found that phonological and acoustic properties of parentese differentiate content and function words, and perceptual experiments have shown that even newborns can differentiate these word categories (Shi, Werker & Morgan, 1999). Furthermore, some researchers have proposed that some acoustic properties that are exaggerated in parentese are associated with syntactic boundaries (Kemler Nelson, Hirsh-Pasek, Jusczyk & Cassidy, 1989). In support of this notion, laboratory studies have demonstrated that infants use these cues to group words into syntactically-relevant sequences (Mandel, Jusczyk & Kemler Nelson, 1994; Mandel, Kemler Nelson & Jusczyk, 1996; Nazzi, Kemler Nelson, Jusczyk & Jusczyk, 2000).

Importantly, for both lexical diversity and utterance length, our results indicate that the predictive power of parentese was improved when the child's own volubility was added to the models. Specifically, parentese and infant language volubility between 6 and 24 months of age together accounted for 17.2% and 21.7% of lexical diversity and MLU at 5 years respectively. Of note, infant speech volubility contributed in the *NEGATIVE* direction (i.e., children whose overall volume of speech between 6 and 24 months was lower had higher lexical diversity and MLU at age 5 years). One potential interpretation of this finding is that associations between parentese and later language complexity may be particularly strong in children who are less "chatty" (i.e., children who are not (yet) producing large quantities of speech themselves). It may be that parentese acts as a mechanism that "drives" infants into later conversational exchanges, and especially so if their own volume of talk is low; future studies with larger numbers of infants who vary in their levels of "chattiness" will be necessary to further unpack this relation. Particularly interesting would be studies using methodologies that capture non-verbal interactions between parents and children, in addition to their verbal exchanges. For example, studies have found that parental contingent comments during *PASSIVE* joint engagements (when infants do not produce speech themselves, but are looking at the same object as the parent) were positively associated with children's later language outcomes (Rollins, 2003; Trautman & Rollins, 2006). Similarly, infants are known to activate their motor brain areas in distinct ways to speech compared to non-speech sounds several months prior to generating intelligible speech (Kuhl, Ramírez, Bosseler, Lin & Imada, 2014), suggesting that they are creating internal speech motor models of native language in response to parental speech, even when they may not yet be talking themselves. Furthermore, it is well documented that infants in the age range studied here (6-24 months) often use gesture to reveal knowledge that they cannot yet express in speech (Goldin-Meadow, 2009). In fact, gestures in infancy are known to selectively predict children's lexical skills at age 3.5 years, even with early child speech controlled (Rowe & Goldin-Meadow, 2009a, 2009b). Furthermore, previous research has found that child gesture *TYPES*, rather than child gesture volume (frequency), predicts later spoken language vocabulary size (Rowe, Özçalışkan & Goldin-Meadow, 2008). Along similar lines, it may be the case that children's later language skills are predicted by vocalization *TYPES* in infancy, which

were not measured in the present study (recall that CVC is a purely quantitative measure). It is also possible that maternal and paternal parentese (analyzed in the present study together as “parental” parentese) are in fact associated with infant vocalizations in distinct ways (Shapiro et al., 2021). Finally, it is important to acknowledge that naturalistic daylong recordings do not control for contextual or conversational style differences, both of which have been shown to play a role in shaping how language input benefits child language growth (Crain-Thoreson, Dahlin & Powell, 2001; Yoder & Kaiser, 1989). For example, conversational styles during toy play, dressing, feeding, and book reading are known to differ in their rates of behavior directives vs conversation eliciting utterances (Hoff-Ginsberg, 1991), and different types of parental utterances have been shown to predict different emergent language skills (Crain-Thoreson et al., 2001; Reese, 1995). Future studies will thus have to unpack the relation between parental parentese and the complexities of children’s linguistic vocalizations between 6 and 24 months, across different contexts, and separating the contributions of various caregivers (i.e., mothers vs fathers; see Shapiro et al., 2021).

Contrary to our hypotheses, in the present dataset, consistent turn taking between 6 and 24 months of age was NOT identified as a significant predictor of any of the 5-year outcomes. This is somewhat surprising, considering that one previous study has linked turn-taking between 18 and 24 months to language outcomes 10 years later (Gilkerson et al., 2018). There are multiple possible interpretations for this apparent discrepancy. First, it is important to note that parentese between 6 and 14 months has previously been proposed to enhance turn-taking at 18 months (Ferjan Ramírez et al., 2020). Considering this finding, it is possible that the predictive power of turn-taking at 18-24 months reported by Gilkerson and colleagues could, in fact, be attributed to parental use of parentese in the first year of life, which was not measured in that study. That is, parentese may act as a mechanism that initially “drives” infants to engage in conversational exchanges, which have been demonstrated to change in quantity and quality somewhere around 18-24 months of age (Gilkerson et al., 2018). For example, prior to 18 months of age, children rarely produce combinatorial speech. Then, around 18 months of age, the first word combinations are typically produced, and children’s vocabularies increase rapidly. While the existence of a “word spurt” is debatable (Dapretto & Bjork, 2000; but see Ganger & Brent, 2004), researchers agree that the rate of word learning and landscape changes in child-language use are observed in the second half of the second year of life (Fromkin, Rodman & Hyams, 2013). Correspondingly, normative data collected in children’s naturalistic environments suggest that the rate of turn-taking between parents and children increases rapidly between 18 and 24 months of age (Gilkerson & Richards, 2009), perhaps to the level where turn-taking finally becomes a reliable predictor of later language outcomes. However, our present data, in combination with previous findings (Ferjan Ramírez et al., 2020), suggest that these enhancements in turn-taking may be partly related to earlier use of parental parentese.

It is also important to point out that conversational turns were measured differently in the study by Gilkerson and colleagues compared to the present study. Specifically, Gilkerson et al. (2018) used LENA’s automatic estimate of turn-taking at the daylong level, while in the present study, conversational turns were annotated manually in 100 30-second segments per participant in order to exclude cases of accidental contiguity, which are abundant in the age range studied here (Ferjan Ramírez et al., 2021). However, while manual annotation is more accurate, the downside is that it is limited to only a portion of the daylong recording, and as such may not be sensitive enough to capture significant relations with later language outcomes. Enhancements in conversational turns

in infancy have previously been linked to larger vocabularies in infancy (Ferjan Ramírez et al., 2020). Furthermore, turn-taking in infancy has recently been proposed to shape the circuitry of language-related brain networks at age 26 months (Huber, Corrigan, Yarnykh, Ferjan Ramírez & Kuhl, 2023), suggesting that the physiology facilitating the back-and-forth interactions is indeed being set up in infancy. Future studies using larger datasets will have to look further into conversational turns and their associations with behavioral measures of language and cognition at Kindergarten entry and beyond. However, the results of the present study, along with previous research (Ferjan Ramírez et al., 2020; Kuhl, 2004), suggests that parentese may be the key to igniting “conversation” in infancy, toddlerhood, and beyond.

Finally, it is important to acknowledge that, while the consistency of high-quality parental input across the first two years of life may be an important contributing factor to Kindergarten language skills, the precise nature of high quality parental language behaviors may change over the course of development. For example, Rowe, (2012) has previously demonstrated that using a diverse and sophisticated vocabulary may be particularly beneficial for toddlers, while using decontextualized language may be particularly beneficial for preschoolers. Likewise, the present dataset suggests that parental use of parentese may have differential effects within the infancy period studied here (6 to 24 months). The links between the exact timing at which parental behaviors have particularly strong associations with specific child language outcomes will have to be further investigated with carefully designed interventions that enroll a sufficiently high number of participants in a single study spanning an age range from infancy to Kindergarten. Unfortunately, this was not possible here, due to the exploratory nature of the present study (i.e., data collection at 5 years was not pre-planned), and due to its correlational design. Future pre-planned and pre-registered studies with higher numbers of participants will also allow for more stringent  $\alpha$  levels. In the present study, we use  $\alpha = 0.05$  as denoting statistical significance, and  $\alpha = 0.1$  as denoting “marginal significance”, allowing us to reliably capture medium to large effect sizes. It is important to acknowledge that, in most fields,  $\alpha = 0.05$  has been used as the gold standard cutoff denoting statistical significance (Miller & Ulrich, 2019), although researchers agree that this cutoff is arbitrary, that there are good reasons to believe that no single  $\alpha$  level is optimal for all research, and that there are certainly contexts, such as exploratory or preliminary studies, where higher  $\alpha$  levels can be appropriate (Michaels, 2017; Serdar et al., 2021). We chose  $\alpha = 0.1$  to denote marginal significance because of the exploratory nature of our study (see Serdar et al., 2021). While this inevitably leads to an increased likelihood of a false positive (i.e., detecting an effect when it is not there in the full population), it also decreases the likelihood of a false negative (i.e., not detecting an effect when it is there in the full population), which we consider important, given the scarcity of previous research on the long-term effects of parentese.

In addition to the exploratory nature and use of a correlational design in the present study, there are other limitations that should be considered. The current sample was originally part of a longitudinal, randomized control study designed to establish the efficacy of a parent coaching intervention for enriching parental language input (Ferjan Ramírez et al., 2018, 2020). The goal of the Intervention was to examine the effect of changing parents’ behavior, while holding other factors constant across participants. Therefore, the sample was intentionally homogenous and included children raised by predominantly White, English-speaking mothers and fathers, within middle to upper-middle SES households in Washington state. While this helped to assure that the Coaching and Control groups were closely matched through the first 24 months of

age, it is important to acknowledge that this demographic may exhibit different patterns of language input compared to families who are not represented in the current study. Intriguingly, recent cross-cultural studies suggest that non-Western children attain developmental milestones on a similar timeline as Western children, even in cultures where parentese use and turn-taking between parents and children may be rare (see Casillas, Brown & Levinson, 2020, 2021). It has also been argued that parentese and turn-taking between caregivers and children are simply a teaching model preferred by middle- and upper-SES families (Avineri, Johnson, Brice-Heath, McCarty, Ochs, Kremer-Sadlik, Blum, Zentella, Rosa, Flores, Alim & Paris, 2015; Sperry, Sperry & Miller, 2018). However, back-and-forth conversational interactions between caregivers and children have been shown to predict language outcomes WITHIN socio-economic groups, including lower income samples (see Golinkoff et al., 2019; Hirsh-Pasek, et al., 2015; Masek, Paterson, Golinkoff, Bakeman, Adamson, Owen, Pace & Hirsh-Pasek, 2021). Additionally, rates of interactive speech have been shown to vary with rates of child-initiated speech, child vocabulary size, and processing speed (Shneidman & Goldin-Meadow, 2012; Weisleder & Fernald, 2013) in addition to child-initiated communicative behavior (Salomo & Liszkowski, 2013) within and across cultures, arguing that differences in qualitative aspects of caregiver input still affect the sophistication of child language. Nevertheless, future research should assess the use of parentese and turn-taking and their predictive power in more diverse populations, such as non-White families, multilingual and non-English-speaking households, single-parent families, and families with same-sex parents.

In summary, the present study provides evidence that parents' consistent use of the speech style that is elicited naturally at home when parents speak to their children – parentese – can predict long-term language outcomes in children when they reach the age of 5 years, while the sheer quantity of language input cannot. Our findings suggest that consistent parental use of parentese speech produces highly robust patterns of association with future language across a large age range, from 6 months of age to 24 months of age. This buttresses the idea that language skills are malleable, and that parentese speech in the natural social context in which it is delivered may present an ideal catalyst for language learning, a fact that needs to be incorporated into theoretical accounts of language acquisition, and one with implications for parents and society.

**Acknowledgements.** This study was supported by the Overdeck Family Foundation and the University of Washington's Language Acquisition and Multilingualism Endowment. The authors thank Lili Correa, Julia Mizrahi, Denise Padden and Neva Corrigan for valuable assistance throughout study design, data collection, and analyses.

**Competing interests.** The authors declare none

## References

- Aslin, R. N., Woodward, J. Z., LaMendola, N. P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 117–134). Mahwah, NJ: Erlbaum.
- Avineri, N., Johnson, E., Brice-Heath, S., McCarty, T., Ochs, E., Kremer-Sadlik, T., Blum, S., Zentella, A. C., Rosa, J., Flores, N., Alim, H. S., & Paris, D. (2015). Invited forum: Bridging the “language gap”. *Journal of linguistic anthropology*, 25(1), 66–86.
- Bird, S., Loper, E., & Klein, E. (2009). *Natural Language Processing with Python*. O'Reilly Media Inc.

- Bornstein, M. H., Tamis-LeMonda, C. S., & Haynes, O. M.** (1999). First words in the second year: Continuity, stability, and models of concurrent and predictive correspondence in vocabulary and verbal responsiveness across age and context. *Infant Behavior & Development*, *22*(1), 65–85. [https://doi.org/10.1016/S0163-6383\(99\)80006-X](https://doi.org/10.1016/S0163-6383(99)80006-X)
- Braarud, H. C., & Stormark, K. M.** (2008). Prosodic modification and vocal adjustments in mothers' speech during face-to-face interaction with their two to four-month-old infants: A double video study. *Social Development*, *17*, 1074–1084. <https://doi.org/10.1111/j.1467-9507.2007.00455>
- Brent, M. R., & Siskind, J. M.** (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*(2), B33–B44.
- Bruner, J.** (1981). The social context of language acquisition. *Language & Communication*, *1*, 155–178. [https://doi.org/10.1016/0271-5309\(81\)90010-0](https://doi.org/10.1016/0271-5309(81)90010-0)
- Burnham, D., Kitamura, C., & Vollmer-Conna, U.** (2002). What's new pussycat? On talking to babies and animals. *Science*, *296*, 1435.
- Byers-Heinlein, K., Tsui, A. S. M., Bergmann, C., Black, A. K., Brown, A., Carbajal, M. J., Durrant, S., Fennell, C. T., Fiévet, A. C., Frank, M. C., Gampe, A., Gervain, J., Gonzalez-Gomez, N., Hamlin, J. K., Havron, N., Hernik, M., Kerr, S., Killam, H., Klassen, K., Kosie, J. E., Kovács, Á. M., Lew-Williams, C., Liu, L., Mani, N., Marino, C., Mastroberardino, M., Mateu, V., Noble, C., Orena, A. J., Polka, L., Potter, C. E., Schreiner, M., Singh, L., Soderstrom, M., Sundara, M., Waddell, C., Werker, J. F., & Wermeinger, S.** (2021). A multi-lab study of bilingual infants: Exploring the preference for infant-directed speech. *Advances in Methods and Practices in Psychological Science*, *4*(1), <https://doi.org/10.1177/2515245920974622>
- Casillas, M., Brown, P., & Levinson, S. C.** (2020). Early language experience in a Tzeltal Mayan village. *Child Development*, *91*(5), 1819–1835.
- Casillas, M., Brown, P., & Levinson, S. C.** (2021). Early language experience in a Papuan community. *Journal of Child Language*, *48*(4), 792–814.
- Cole, P. M., Armstrong, L. M., & Pemberton, C. K.** (2010). The role of language in the development of emotion regulation. In S. D. Calkins & M. A. Bell (Eds.), *Child development at the intersection of emotion and cognition*. (pp. 59–77). American Psychological Association. <https://doi.org/10.1037/12059-004>
- Conboy, B. T., Brooks, R., Meltzoff, A. N., & Kuhl, P. K.** (2015). Social interaction in infants' learning of second-language phonetics: An exploration of brain-behavior relations. *Developmental Neuropsychology*, *40*(4), 216–229. <https://doi.org/10.1080/87565641.2015.1014487>
- Cooper, R. P., & Aslin, R. N.** (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, *61*(5), 1584–1595. <https://doi.org/10.2307/1130766>
- Crain-Thoreson, C., Dahlin, M. P., & Powell, T. A.** (2001). Parent-child interaction in three conversational contexts: variations in style and strategy. *New directions for child and adolescent development*, *(92)*, 23–98. <https://doi.org/10.1002/cd.13>
- Cristia, A., Bulgarelli, F., & Bergelson, E.** (2020). Accuracy of the Language Environment Analysis system segmentation and metrics: A systematic review. *Journal of Speech, Language, and Hearing Research*, *63*(4), 1093–1105. [https://doi.org/10.1044/2020\\_JSLHR-19-00017](https://doi.org/10.1044/2020_JSLHR-19-00017)
- Dapretto, M., & Bjork, E. L.** (2000). The development of word retrieval abilities in the second year and its relation to early vocabulary growth. *Child Development*, *71*(3), 635–648. <https://doi.org/10.1111/1467-8624.00172>
- Durham, R. E., Farkas, G., Hammer, C. S., Tomblin, J. B., & Catts, H. W.** (2007). Kindergarten oral language skill: A key variable in the intergenerational transmission of socioeconomic status. *Research in Social Stratification and Mobility*, *25*, 294–305. <https://doi.org/10.1016/j.rssm.2007.03.001>
- Ferguson, C. A.** (1964). Baby Talk in Six Languages. *American Anthropologist*, *66*, 103–114.
- Ferjan Ramirez, N., Hippe, D. S., Correa, L., Andert, J., & Baralt, M.** (2022). Habla conmigo, daddy! Fathers' language input in North American bilingual Latinx families. *Infancy*, *00*, 1–23. <https://doi.org/10.1111/inf.12450>
- Ferjan Ramirez, N., Hippe, D. S., & Kuhl, P. K.** (2021). Comparing automatic and manual measures of parent-infant conversational turns: A word of caution. *Child Development*, *92*(2), 672–681. <https://doi.org/10.1111/cdev.13495>
- Ferjan Ramirez, N., Lytle, S., Fish, M., & Kuhl, P. K.** (2018). Parent coaching at 6 and 10 months improves language outcomes at 14 months: A randomized controlled trial. *Developmental Science*, *22*(3). <https://doi.org/10.1111/desc.12762>

- Ferjan Ramírez, N., Lytle, S., & Kuhl, P. K. (2020). Parent coaching increases conversational turns and advances infant language development. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(7), 3484–3491. <https://doi.org/10.1073/pnas.1921653117>
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, *8*, 181–195. [https://doi.org/10.1016/S0163-6383\(85\)80005-9](https://doi.org/10.1016/S0163-6383(85)80005-9)
- Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior & Development*, *10*(3), 279–293. [https://doi.org/10.1016/0163-6383\(87\)90017-8](https://doi.org/10.1016/0163-6383(87)90017-8)
- Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology*, *27*(2), 209–221.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, *20*(1), 104–113. <https://doi.org/10.1037/0012-1649.20.1.104>
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*, 477–501. <https://doi.org/10.1017/S0305000900010679>
- Fromkin, V., Rodman, R., & Hyams, N. (2013). *An Introduction to Language*. 10th ed. Boston, MA: Cengage Learning
- Ganger, J., & Brent, M. R. (2004). Reexamining the vocabulary spurt. *Developmental Psychology*, *40*(4), 621–632. <https://doi.org/10.1037/0012-1649.40.4.621>
- Garnica, O. (1977). On some prosodic and paralinguistic features of speech to young children. In C. E. Snow, & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition*, (pp. 271–285). Cambridge University Press.
- Gergely, A., Faragó, T., Galambos, Á., & Topál, J. (2017). Differential effects of speech situations on mothers' and fathers' infant-directed and dog-directed speech: An acoustic analysis. *Scientific Reports*, *7*, 13739. <https://doi.org/10.1038/s41598-017-13883-2>
- Gilkerson, J., & Richards, J. A. (2009). *The power of talk: Impact of adult talk, conversational turns, and TV during the critical 0-4 years of child development*. In *The power of talk* (2nd ed.). LENA Foundation.
- Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Oller, K. D., Hansen, J., & Paul, T. D. (2017). Mapping the Early Language Environment Using All-Day Recordings and Automated Analysis. *American journal of speech-language pathology*, *26*(2), 248–265. [https://doi.org/10.1044/2016\\_AJSLP-15-0169](https://doi.org/10.1044/2016_AJSLP-15-0169)
- Gilkerson, J., Richards, J. A., Warren, S. F., Oller, D. K., Russo, R., & Vohr, B. (2018). Language experience in the second year of life and language outcomes in late childhood. *Pediatrics*, *142*(4). <https://doi.org/10.1542/peds.2017-4276>
- Goldin-Meadow, S. (2009). How gesture promotes learning throughout childhood. *Child development perspectives*, *3*(2), 106–111. <https://doi.org/10.1111/j.1750-8606.2009.00088.x>
- Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, *19*(5), 515–523. <https://doi.org/10.1111/j.1467-9280.2008.02117.x>
- Golinkoff, R. M., Hoff, E., Rowe, M. L., & Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2019). Language matters: Denying the existence of the 30-million-word gap has serious consequences. *Child Development*, *90*(3), 985–992. <https://doi.org/10.1111/cdev.13128>
- 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>
- 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*(1), 14–20. <https://doi.org/10.1037/0012-1649.24.1.14>
- Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.
- Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., Yust, & Suma, K. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science*, *26*, 7, 1071–1083. <https://doi.org/10.1177/0956797615581493>
- Hirsh-Pasek, K., & Treiman, R. (1982). Doggerel: Motherese in a new context. *Journal of Child Language*, *9*, 229–237.
- Hoff-Ginsberg, E. (1991). Mother-child conversation in different social classes and communicative settings. *Child development*, *62*(4), 782–796. <https://doi.org/10.1111/j.1467-8624.1991.tb01569.x>

- Huber, E., Corrigan, N., Yarnykh, V., Ferjan Ramírez, N., & Kuhl, P. K. K. (2023). Language experience during infancy predicts white matter myelination at age two years. *Journal of Neuroscience, in press*.
- Jacobson, J. L., Boersma, D. C., Fields, R. B., Olson, K. L., & David, C. (1983). Paralinguistic features of adult speech to infants and small children. *Child Development, 54*(2), 436–442. <https://doi.org/10.2307/1129704>
- 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.
- 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. <https://doi.org/10.1111/j.1467-7687.2007.00572>.
- 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. <https://doi.org/10.1126/science.277.5326.684>
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B, 363*, 979–1000.
- Kuhl, P. K., Ramírez, R. R., Bosseler, A., Lin, J. F., & Imada, T. (2014). Infants’ brain responses to speech suggest analysis by synthesis. *Proceedings of the National Academy of Sciences of the United States of America, 111*(31), 11238–11245. <https://doi.org/10.1073/pnas.1410963111>
- Kuhl, P. K., Tsao, F., & Liu, H. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America, 100* (15), 9096–9101. <https://doi.org/10.1073/pnas.1532872100>
- Landry, S. H., Smith, K. E., & Swank, P. R. (2003). The importance of parenting during early childhood for school-age development. *Developmental Neuropsychology, 24*(2-3), 559–591. <https://doi.org/10.1080/87565641.2003.9651911>
- Lehet, M., Arjmandi, M. K., Houston, D., & Dilley, L. (2020). Circumspection in using automated measures: Talker gender and addressee affect error rates for adult speech detection in the Language ENvironment Analysis (LENA) system. *Behavior Research Methods, 53*, 113–138. <https://doi.org/10.3758/s13428-020-01419-y>
- LENA. (2015). LENA™ Pro [Computer software]. Boulder, CO
- Levinson, S. C. (2016). Turn-taking in human communication—Origins and implications for language processing. *Trends in Cognitive Sciences, 20*(1), 6–14. <https://doi.org/10.1016/j.tics.2015.10.010>
- Ma, W., Golinkoff, R. M., Houston, D., & Hirsh-Pasek, K. (2011). Word learning in infant- and adult-directed speech. *Language Learning and Development. The Official Journal of the Society for Language Development, 7*(3), 185–201. <https://doi.org/10.1080/15475441.2011.579839>
- Mandel, D. R., Jusczyk, P. W., & Kemler Nelson, D. G. (1994). Does sentential prosody help infants to organize and remember speech information? *Cognition, 53*, 155–180.
- Mandel, D. R., Kemler Nelson, D. G., & Jusczyk, P. W. (1996). Infants remember the order of words in a spoken sentence. *Cognitive Development, 11*, 181–196.
- The ManyBabies Consortium. (2020). Quantifying sources of variability in infancy research using the infant-directed-speech preference. *Advances in Methods and Practices in Psychological Science, 3*(1), 24–52. <https://doi.org/10.1177/2515245919900809>
- Masek, L. R., Paterson, S. J., Golinkoff, R. M., Bakeman, R., Adamson, L. B., Owen, M. T., Pace, A., & Hirsh-Pasek, K. (2021). Beyond talk: Contributions of quantity and quality of communication to language success across socioeconomic strata. *Infancy, 26*(1), 123–147. <https://doi.org/10.1111/inf.12378>
- McCarthy, D. (1954). Language development in children. In L. Carmichael (Eds.), *Manual of child psychology* (pp. 492–630). John Wiley & Sons Inc.
- Merz, E. C., Maskus, E. A., Melvin, S. A., He, X., & Noble, K. G. (2020). Socioeconomic disparities in language input are associated with children’s language-related brain structure and reading skills. *Child Development, 91*(3), 846–860. <https://doi.org/10.1111/cdev.13239>
- Michaels, R. (2017). Confidence in courts: A delicate balance. *Science, 357*(6353):764–764. [10.1126/science.aao3967](https://doi.org/10.1126/science.aao3967)
- Miller, J., & Ulrich, R. (2019). The quest for an optimal alpha. *PLoS ONE, 14*(1), Article e0208631. <https://doi.org/10.1371/journal.pone.0208631>

- Miller, J. F., Andriacchi, K., & Nockerts, A. (2019). *Assessing language production using SALT software: A clinician's guide to language sample analysis* (3rd ed.). Madison, WI: SALT Software, LLC.
- Mitchell, R. W. (2001). Americans' talk with dogs: Similarities and differences with talk to infants. *Research on Language and Social Interaction*, *34*(2), 183–210.
- Narayan, C. R., & McDermott, L. C. (2016). Speech rate and pitch characteristics of infant-directed speech: Longitudinal and cross-linguistic observations. *Journal of the Acoustical Society of America*, *139*(3), 1272–1281. <https://doi.org/10.1121/1.4944634>
- Nazzi, T., Kemler Nelson, D. G., Jusczyk, P. W., & Jusczyk, A. M. (2000). Six-month-olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. *Infancy*, *1*, 123–147.
- Newport, E. L., Gleitman, H., & Gleitman, L. R. (1977). Mother, I'd rather do it myself: some effects and non-effects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: language input and acquisition* (pp. 109–149). New York: C.U.P
- Pace, A., Burchinal, P., Alper, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2019). Measuring success: Within and cross-domain predictors of academic and social trajectories in elementary school. *Early Childhood Research Quarterly*, *46*, 112–125. <https://doi.org/10.1016/j.ecresq.2018.04.001>
- Phillips, J. R. (1973). Syntax and vocabulary of mothers' speech to young children: Age and sex comparisons. *Child Development*, *44*, 182–185.
- Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2014). Look who's talking: Speech style and social context in Language input to infants are linked to concurrent and future speech development. *Developmental Science*, *17*, 880–891. <https://doi.org/10.1111/j.1467-7687.2005.00445>
- Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2016). The impact of early social interactions on later language development in Spanish–English bilingual infants. *Child Development*, *88*, 1216–1234. <https://doi.org/10.1111/cdev.12648>
- Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2017). Look who's talking now! Parentese speech, social context, and language development across time. *Frontiers in Psychology*, *8*, 1173. <https://doi.org/10.3389/fpsyg.2017.01008>
- Reese, E. (1995). Predicting children's literacy from mother-child conversations. *Cognitive Development*, *10* (3), 381–405. [https://doi.org/10.1016/0885-2014\(95\)90003-9](https://doi.org/10.1016/0885-2014(95)90003-9)
- Reilly, J. S., & Bellugi, U. (1996). Competition on the face: Affect and language in ASL motherese. *Journal of Child Language*, *23*(1), 219–239. <https://doi.org/10.1017/S0305000900010163>
- Rodríguez, E. T., & Tamis-LeMonda, C. S. (2011). Trajectories of the home learning environment across the first 5 years: Associations with children's vocabulary and literacy skills at prekindergarten. *Child Development*, *82*(4), 1058–1075. <https://doi.org/10.1111/j.1467-8624.2011.01614.x>
- Rollins, P. R. (2003). Caregivers' contingent comments to 9-month-old infants: Relationships with later language. *Applied psycholinguistics*, *24*(2), 221–234. <https://doi.org/10.1017/S0142716403000110>
- Romeo, R. R., Leonard, J. A., Grotzinger, H. M., Robinson, S. T., Takada, M. E., Mackey, A. P., Scherer, E., Rowe, M. L., West, M. R., & Gabrieli, J. D. E. (2021). Neuroplasticity associated with changes in conversational turn-taking following a family-based intervention. *Developmental Cognitive Neuroscience* *49*. 1878–9293 <https://doi.org/10.1016/j.dcn.2021.100967>
- Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. E. (2018a). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science*, *29*(5), 700–710. <https://doi.org/10.1177/0956797617742725>
- Romeo, R. R., Segaran, J., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Yendiki, A., Rowe, M. L., & Gabrieli, J. D. E. (2018b). Language exposure relates to structural neural connectivity in childhood. *Journal of Neuroscience*, *38*(36), 7870–7877. <https://doi.org/10.1523/JNEUROSCI.0484-18.2018>
- 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>
- Rowe, M. L., & Goldin-Meadow, S. (2009a). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science* (New York, N.Y.), *323*(5916), 951–953. <https://doi.org/10.1126/science.1167025>
- Rowe, M. L., & Goldin-Meadow, S. (2009b). Early gesture selectively predicts later language learning. *Developmental science*, *12*(1), 182–187. <https://doi.org/10.1111/j.1467-7687.2008.00764.x>
- Rowe, M. L., Ozcaliskan, S., & Goldin-Meadow, S. (2008). Learning words by hand: Gesture's role in predicting vocabulary development. *First language*, *28*(2), 182–199. <https://doi.org/10.1177/0142723707088310>



- Rowe, M. L., & Snow, C. E. (2020). Analyzing input quality along three dimensions: interactive, linguistic, and conceptual. *Journal of Child Language*, *47*(1), 5–21. <https://doi.org/10.1017/S0305000919000655>
- Rowe, M. L., & Weisleder, A. (2020). Language Development in Context. *Annual Review of Developmental Psychology*, *2*(1), 22.
- Salomo, D., & Liszkowski, U. (2013). Sociocultural settings influence the emergence of prelinguistic deictic gestures. *Child development*, *84*(4), 1296–1307.
- Schroeder, L. D., Sjoquist, D. L., & Stephan, P. E. (2016). *Understanding regression analysis: An introductory guide* (Vol. 57). Sage Publications.
- Serdar, C. C., Cihan, M., Yücel, D., & Serdar, M. A. (2021). Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. *Biochemia medica*, *31* (1), 010502. <https://doi.org/10.11613/BM.2021.010502>
- Shapiro, N. T., Hippe, D. S., & Ferjan Ramírez, N. (2021). How Chatty Are Daddies? An Exploratory Study of Infants' Language Environments. *Journal of speech, language, and hearing research: JSLHR*, *64*(8), 3242–3252. [https://doi.org/10.1044/2021\\_JSLHR-20-00727](https://doi.org/10.1044/2021_JSLHR-20-00727)
- Shi, R., Morgan, J. L., & Allopena, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective. *Journal of Child Language*, *25*, 169–201.
- Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, *72*, B11–B21.
- Shneidman, L. A., Arroyo, M. E., Levine, S. C., & Goldin-Meadow, S. (2013). What counts as effective input for word learning? *Journal of Child Language*, *40*(3), 672–686. <https://doi.org/10.1017/S0305000912000141>
- Shneidman, L. A., & Goldin-Meadow, S. (2012). Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science*, *15*(5), 659–673.
- Singh, L., Morgan, J. L., & Best, C. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy*, *3*, 365–394.
- Singh, L., Nestor, S., Parikh, C., & Yull, A. (2009). Influences of infant-directed speech on early word recognition. *Infancy*, *14*(6), 654–666. <https://doi.org/10.1080/15250000903263973>
- Smith, N. A., & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy*, *13*, 410–420. <https://doi.org/10.1080/15250000802188719>
- Snow, C. E. (1977a). Mothers' speech research: From input to interaction. In C. E. Snow, & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition*, (pp. 31–49). Cambridge University Press.
- Snow, C. E. (1977b). The development of conversation between mothers and babies. *Journal of Child Language*, *4*(1), 1–22. <https://doi.org/10.1017/S0305000900000453>
- Snow, C. E., van Eeden, R., & Muysken, P. (1981). The interactional origins of foreigner talk: Municipal employees and foreign workers. *International Journal of the Sociology of Language*, *28*, 81–91.
- Sperry, D. E., Sperry, L. L., & Miller, P. J. (2018). Reexamining the Verbal Environments of Children from Different Socioeconomic Backgrounds. *Child Development*, *90*(4), 1303–1318
- Stern, D. N., Spieker, S., Barnett, R. K., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language*, *10*, 1–15. <https://doi.org/10.1017/S03050009000005092>
- Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Current Directions in Psychological Science*, *23*, 121–126. <https://doi.org/10.1177/0963721414522813>
- Tamis-LeMonda, C. S., Luo, R., McFadden, K. E., Bandel, E. T., & Vallotton, C. (2019). Early home 5th grade academic skills. *Applied Developmental Science*, *23*(2), 153–169.
- Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*, 53–71. [https://doi.org/10.1207/s15327078in0701\\_5](https://doi.org/10.1207/s15327078in0701_5)
- Trautman, C., & Rollins, P. (2006). Child-centered behaviors of caregivers with 12-month-old infants: Associations with passive joint engagement and later language. *Applied psycholinguistics*, *27*(3), 447–463. doi:10.1017/S0142716406060358
- Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication*, *49*(1), 2–7. <https://doi.org/10.1016/j.specom.2006.10.003>
- Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.
- Vygotsky, L. S. (1979). Interaction between learning and development. In M. Gauvain, & M. Cole (Eds.), *Readings on the development of children*, (pp. 34–40). Scientific American Books.

- Wang, Y., Williams, R., Dilley, L., & Houston, D. M. (2020). A meta-analysis of the predictability of LENATM automated measures for child language development. *Developmental Review*, *57*, 100921. <https://doi.org/10.1016/j.dr.2020.100921>
- Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, K. D. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, *25*(7), 1314–1324. <https://doi.org/10.1177/0956797614531023>
- 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>
- Werker, J. F., Pegg, J. E., & McLeod, P. J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior & Development*, *17*(3), 323–333. [https://doi.org/10.1016/0163-6383\(94\)90012-4](https://doi.org/10.1016/0163-6383(94)90012-4)
- Xu, D., Richards, J. A., & Gilkerson, J. (2014). Automated analysis of child phonetic production using naturalistic recordings. *Journal of speech, language, and hearing research : JSLHR*, *57*(5), 1638–1650. [https://doi.org/10.1044/2014\\_JSLHR-S-13-0037](https://doi.org/10.1044/2014_JSLHR-S-13-0037)
- Yoder, P. J., & Kaiser, A. P. (1989). Alternative explanations for the relationship between maternal verbal interaction style and child language development. *Journal of child language*, *16*(1), 141–160. <https://doi.org/10.1017/s0305000900013489>
- 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*(1), 342–349. <https://doi.org/10.1542/peds.2008-2267>

---

Cite this article: Ferjan Ramírez, N., Weiss, Y., Sheth, K.K., & Kuhl, P.K. (2024). Parentese in infancy predicts 5-year language complexity and conversational turns. *Journal of Child Language* *51*, 359–384, <https://doi.org/10.1017/S0305000923000077>