Early Childhood Research Quarterly 64 (2023) 199–215

Contents lists available at ScienceDirect

Early Childhood Research Quarterly

journal homepage: www.elsevier.com/locate/ecresq

Variations in the home language environment and early language development in a peri-urban community in China

Tianli Feng[a], Xinwu Zhang[b,*], Lulu Zhou[b], Yue Zhang[c], Lucy Pappas[b], Sarah-Eve Dill[b], Scott Rozelle[b], Yue Ma[b]

[a] School of Management and Economics, University of Electronic Science and Technology of China, Chengdu 610054, China
[b] Stanford Center on China's Economy and Institutions, Encina Hall, 616 Jane Stanford Way, Stanford University, Stanford, California 94305, United States
[c] National Center for Women and Children's Health, Chinese Center for Disease Control and Prevention, Beijing 100081, China

ARTICLE INFO

Keywords:
Home language environment
Language development
Peri-urban China
Language environment analysis (LENA)
Low-socioeconomic status (SES)

ABSTRACT

Parental investment in the home language environment during the earliest years is a critical predictor of early language development. Because most studies investigating the home language environment and child language development have been conducted in Western, high-income, and developed settings, less is known about such environments in low- or middle-income settings. This study was conducted in a peri-urban area in Southwestern China in a sample of 81 rural migrant and urbanized farmer families with children aged 18-24 months. The home language environment was measured using Language Environment Analysis (LENA) recorders and software, while early language development was measured using the MacArthur-Bates Communicative Developmental Inventories expressive vocabulary scale. Findings reveal large and substantial variation in the sample's home language environments and a strong association between the home language environment and child language development. Certain demographic characteristics, such as household resources, maternal employment, and gender, are associated with the home language environment. These findings highlight the needs for interventions specifically targeting the home language environment to improve early language development of young children and for more research on early childhood development in peri-urban China.

1. Introduction

The first three years of life is one of the most critical periods for early childhood development (ECD). During this time, environmental variables within the home correlate significantly with child development, and may lead to either positive or negative developmental trajectories (Gottfried, 2013). Linguistic and social experiences in early childhood have been found to impact growth in cognitive functioning, language acquisition, and later academic skills (Romeo et al., 2018). According to the investment model—one of the many theoretical frameworks describing relations between various environmental variables and ECD outcomes—better caregiver investment (typically predicated on the availability of economic resources) during the early years of a child's life leads to the largest returns in skills attainment and later-life outcomes (Coddington et al., 2014; Heckman et al., 2013; Conger et al., 2010; Cunha et al., 2010; ; Heckman, 2006, 2008).

Specifically, a body of ECD literature has highlighted the home language environment as a critical area of early investment. In this paper, home language environment is used as an umbrella term to describe the language-related interactions and processes (also referred to as language inputs) that children are exposed to at home. Research has shown that the quality of the home language environment is a strong predictor of a child's language development outcomes (Gilkerson et al., 2018; Ramírez-Esparza et al., 2014; Weisleder & Fernald, 2013). For example, studies conducted in Western contexts have documented how children with caregivers who speak to those children more frequently tend to have larger vocabularies and process information faster than children with less developmentally-engaging home language environments (Hurtado et al., 2014; Gilkerson & Richards, 2009; Hart & Risley, 1995;). Moreover, because early language development has been identified as a determinant of academic outcomes, what a caregiver invests linguistically into the home language environment can significantly influence their child's overall developmental outcome (Fernald et al., 2013; Lee & Burkam, 2002).

Early investment into the home language environment is especially important for children from low socio-economic backgrounds, a population that is at high risk for experiencing significant disparities in lan-

* Corresponding authors.
E-mail addresses: tianlie@uestc.edu.cn (T. Feng), zhangxinwuceee@126.com (X. Zhang), luluzhou19@gmail.com (L. Zhou), zhangyue0416@163.com (Y. Zhang), lmpappas@stanford.edu (L. Pappas), sedill@stanford.edu (S.-E. Dill), rozelle@stanford.edu (S. Rozelle), yma3@stanford.edu (Y. Ma).

https://doi.org/10.1016/j.ecresq.2023.03.005
Received 8 November 2021; Received in revised form 15 March 2023; Accepted 24 March 2023

guage and literacy skills (Mendive et al., 2020; Merz et al., 2020). There have been consistent and robust findings that early language disparities contribute to cumulative disadvantages over time (; Leung et al., 2020; Hart and Risley, 1995). In addition, evidence has shown that, among preschool-aged children who experience untreated speech and language delay, 40% to 60% of the children may experience persistent language difficulties (Law et al., 1998). On top of that, an estimated 250 million young children living in low- and middle-income countries (LMICs) are already vulnerable to reduced cognition and developmental delay (e.g., in literacy-numeracy, learning/cognition, physical development, and socioemotional development) (Bai et al., 2019; Black et al., 2017; McCoy et al., 2016). Seeing as language skills developed in the first years of life provide a foundation for later skill development that impacts factors such as school readiness and academic achievement, targeted language development interventions may be a key strategy for improving ECD and child outcomes in at-risk settings (Fernald et al., 2013; Meisenberg & Woodley, 2013; Wake et al., 2012).

1.1. Describing the home language environment: language quality and quantity

It is well-established that children's early language skills are strongly correlated with their exposure to language inputs in the home context. Variability in children's language skills has been documented as dependent on quantitative and qualitative differences in their language-learning experiences, such as their exposure to conversational speech with adults and the content of that speech (Goodrich et al., 2021; Scheele et al., 2010). According to research by Hart and Risley (1995), the number of adult words spoken to children from birth to age 3 strongly correlates with their language ability by preschool age. Children who are exposed to larger quantities of adult speech in the early stages of language acquisition have been found to experience a reduction in processing time for encoding, which aids in their vocabulary development (Ellis, 2002). Yet simply overhearing adult speech in the home environment does not necessarily facilitate language learning in young children; indeed, social interaction is at the core of children's language development, because language is embedded in a child's daily interactions with their caregivers (; Hirsh-Pasek et al., 2015; Bruner, 1985). Empirical studies have revealed that young children learn language most effectively from adult speech directed at them (i.e., child-directed speech) and by participating in extended conversations through multiple speaker turns (i.e., conversational turns with adults) (Leech & Rowe, 2021; Leung et al., 2020).

Additionally, not all home language environments are the same; this is reflected in the wide variability of child language development outcomes across socioeconomic-status (SES) levels. Studies in the United States have consistently found that low-SES families tend to have less-engaging home language environments (e.g., fewer adult words spoken and fewer adult-child conversations had) compared to high-SES families (Ramírez et al., 2020; Ganek & Eriks-Brophy, 2018; Hirsh-Pasek et al., 2015; Noble et al., 2015; Hurtado et al., 2014; ; ; Ramírez-Esparza et al., 2014;Weisleder & Fernald, 2013; Rowe, 2012; Rowe & Goldin-Meadow, 2009; Hoff, 2003). Moreover, disparities arise in both the quality and quantity of caregiver-child interactions. Higher-SES caregivers tend to have more advanced and stimulating verbal interactions with their children more frequently than lower-SES caregivers, and higher-SES caregivers tend to speak far more words to their children than low-SES caregivers (Buckingham et al., 2014; Hart & Risley, 1995). Partially due to their less-engaging language environments, then, children from low-SES backgrounds have been found to have weaker cognitive and linguistic performance than their higher-SES peers in early childhood (Ramey & Ramey, 2004). For example, a study of 18- to 24-month-old children in Fernald et al. (2013) found that the language processing efficiency of children from lower-SES families trailed their higher-SES peers of the same age by an average of six months.

1.2. Measuring the home language environment: LENA

Although researchers have long recognized the importance of the home language environment for early language development, dedicated studies are generally rare due to the difficulties inherent in measuring the home language environment. Observer bias is a common difficulty in recording home language environments, as audio recordings are intrusive and disrupt the natural environment. However, recent developments in research technology have allowed researchers to conduct unobtrusive, naturalistic studies on the home language environment. Thanks to the Language Environment Analysis (LENA) system—a small, wearable recorder and complementary language processing software—researchers can quantify a child's home language environment without disrupting it (Wang et al., 2017). The LENA recorders are worn by children and record all language in their vicinity for 16 hours. With these day-long recordings, the LENA software can distinguish between adult words, adult-child conversations, child vocalizations, electronic noise, silence, and more, providing researchers with a comprehensive, quantitative depiction of home language environments.

Using LENA technology, numerous studies in Western settings have demonstrated the importance of the home language environment on language development outcomes. Studies employing LENA have shown that adult speech, adult-child interactions, and child vocalizations are significantly correlated to early language outcomes, such that more adult words spoken to children, more conversations had with children, and more vocalizations made by the child are positively correlated to stronger early language development (e.g., Lopez et al., 2020; d'Apice et al., 2019; Uccelli et al., 2019; Gilkerson et al., 2018; Romeo et al., 2018, 2021).

1.3. Influences on the home language environment: Using LENA

Confirming what other language environment studies have found, preliminary evidence using LENA devices shows that households of different SES have different home language environments, with higher SES households reporting stronger child development outcomes. Furthermore, studies suggest that even within low-SES households, there is large variation in home language environments. A LENA study of 29 Spanish-learning infants from low-SES families in the United States found significant variability in the amount of overheard speech and child-directed speech (Weisleder & Fernald, 2013). The study also found that more diverse, child-directed speech was associated with stronger language processing abilities within the sample, even though sample families had similarly disadvantaged demographic characteristics (i.e., low income, low educational attainment, limited access to quality health care). Beyond the study in low-income areas of the United States, however, there are few LENA studies that have attempted to measure variations, or the sources of variations, in the home language environment within non-Western, low-SES populations. To our knowledge, only a small number of studies have examined variations in the home language environment within low-SES groups in low- and middle-income settings, where often large shares of the population have lower-SES backgrounds (Casillas et al., 2020, 2021; Pae et al., 2016; Zhang et al., 2015). Currently, few studies using LENA technology have analyzed the home language environment in non-Western samples, including one study from rural Senegal (Weber et al., 2017), two studies from rural Shaanxi Province and urban Shanghai, China, respectively (Ma et al., 2021; Zhang et al., 2015), and one from urban Seoul, South Korea (Pae et al., 2016). Although the latter two studies were carried out in East Asian countries, both targeted high-SES populations in urban megacities. Furthermore, even fewer existing LENA studies have examined the home language environment of low-SES communities in an Asian country, and these studies found that children living in rural communities heard less adult words and experienced less conversational turns than children living in urban, higher-SES communities (Ma et al., 2021, 2023).

In addition to the inclusion of SES as a possible associated factor with the home language environment, the literature on the home language environment and early language development has identified several child and family demographic characteristics that may act as significant factors. Previous studies (using LENA and other methods) have found significant differences in language development and home language environment measures between children of different ages (Gilkerson et al., 2018; Doyle et al., 2012) and between boys and girls (Bornstein, 2002). Other studies have identified mixed findings on the links between maternal age and child cognitive development, which often overlaps with language development. Some studies have found a positive relationship between maternal age and child cognition, while others have suggested an inverse U-shape relationship such that the lowest and highest ages are negatively associated with cognitive development (Duncan et al., 2018). Research has found positive correlations between parental education levels and child language development, as well as correlations between parental education levels and linguistic interactions between adults and children (Dahl & Lochner, 2005; Tamis-LeMonda et al., 2004). The evidence on associations between maternal employment and early development is mixed, with research pointing toward an overall neutral effect of maternal employment in the first year of a child's life on early development (Brooks-Gunn et al., 2010).

1.4. The present study

To address the lack of evidence from low-SES, non-Western populations in the existing literature, this study examines the home language environment and early language development in a low-SES and previously understudied population: children from a peri-urban community in Southwestern China. In China, low-income, peri-urban spaces, known as "villages-within-cities" (*chengzhongcun*), are effectively rural communities within an urban setting (Buckingham & Chan, 2018; Li & Xiong, 2018). These communities have residents from diverse backgrounds, including urban and rural migrants, as well as urbanized (i.e., former) farmers. Our sampling focused on rural migrant and urbanized farmer families because they tend to have lower SES than urban families. To be clear, in the rest of this study, "urbanized farmers" (*shidi nongmin*) refers to local residents of the community who were formerly rural farmers but lost their farmland in the urbanization process that has physically transformed their environment (Heurlin, 2019). Because of this urbanization of (former) rural communities, many urbanized farmers of our sample have obtained new legal urban status, also known as having an urban *hukou*, as defined by the Chinese national household registration system.[1] "Rural migrants," on the other hand, have out-migrated from their rural hometowns to urban areas for work and are still legally considered to be rural residents, thus holding rural *hukou* (Buckingham & Chan, 2018; Li, 2010). In most cases, urbanized farmers own their own homes in these peri-urban communities, while most rural migrants rent their housing. However, despite differences in their legal status in China's national registration system, both groups share rural backgrounds and, on average, have relatively low levels of SES (education and income/asset levels) compared to their fully urban local peers (Heurlin, 2019; Tong et al., 2019; He & Wang, 2016; ; Li & Placier, 2015; Zhan, 2011; Li, 2010).

To date, little is known about the home language environment or early language development among peri-urban communities in China. One study of ECD in rural communities across China found that among rural migrants, 39% of children under age 3 have language delays, which is more than twice that of a healthy population (Wang et al., 2019). The same study found that levels of parental investment in interactive parenting practices were also low among rural migrants: only 21% of migrant caregivers (mothers and fathers) had read a book to their child and only 29% had told a story to their child in the previous two days (Wang et al., 2019). However, the aforementioned study is the only study to date on language development among China's rural migrants, and yet it did not examine other peri-urban populations such as urbanized farmers. Additionally, this study did not focus on the home language environment specifically, nor did it examine variations in language inputs and/or child language development.

In the current study, we aim to describe the home language environment and early language development of young children aged 18-24 months from rural migrant and urbanized farmer families in a peri-urban community in Southwestern China. To do so, we pursue four research questions. First, using LENA technology, what is the quality and the heterogeneity of the home language environment in terms of adult words, adult-child conversations, and child vocalizations? Second, what is the distribution of child language development and the prevalence of language delays among our sample? Third, how are the characteristics of families with different home language environments similar and different? Fourth, how do language development outcomes of children in different home language environments compare to one another?

2. Materials and methods

2.1. Ethics statement

This study received ethical approval from the Stanford University Institutional Review Board (Protocol ID 57196). Trained members of the field survey team received informed oral consent from all caregivers of sample children. Caregivers were aware that their audio recording data would be collected and used for the purposes of this study.

2.2. Sample selection

Data for the study were collected in a peri-urban district located in Southwestern China. The district is considered part of the second tier of the provincial capital municipality, which is less developed than the central first tier.[2] Unlike the 100% urbanization rates of the provincial capital's five first-tier districts, the urbanization rate of this peri-urban district was 73% in 2020, with the majority of the district's urban development occurring between 2010 and 2020 (Wang, 2020). Many residents in the district still have rural roots, such as the rural migrants who relocated from rural regions (both within and outside of the province) and the urbanized farmers who held rural *hukou* (household registration status) before the urbanization of the district (Zhu, 2017). About a third of the district's population still has rural *hukou* status, including both migrant and local residents (Wang, 2020).

Given this split in residency status of the population, there are also differences in the SES characteristics (e.g., income and educational attainment levels) of the groups. On average, the SES of residents tends to be lower than that of residents in the central urban districts of the provincial capital. For example, compared to the RMB 49193 (USD 7028) per capita income of residents in the capital's central urban districts in 2019, the average per capita disposable income for rural residents in the peri-urban district was considerably lower at RMB 30405 (USD 4344) (Wang, 2020). Moreover, unlike the sampled district, the five first-tier districts of the provincial capital are fully urbanized; nearly all their residents have urban *hukou* status (District B Statistical Report, 2020; District C Statistical Report, 2020; District D Statistical Report, 2020, District E Statistical Report, 2020; District F Statistical

[1] China's household registration system (hukou), which started in the late 1950s, is a legal system that rigidly classifies individuals as either rural or urban residents and determines their social benefits accordingly. A hukou identifies individuals by their "places of origin" in China. An urban hukou has more privileges than a rural hukou (S. Li, 2010).

[2] In China, the tier system is a hierarchical classification of cities and districts. First tiers tend to have greater levels of economic development, political influence, and population size than second, third, and fourth tiers (Bland & Hernández, 2016).

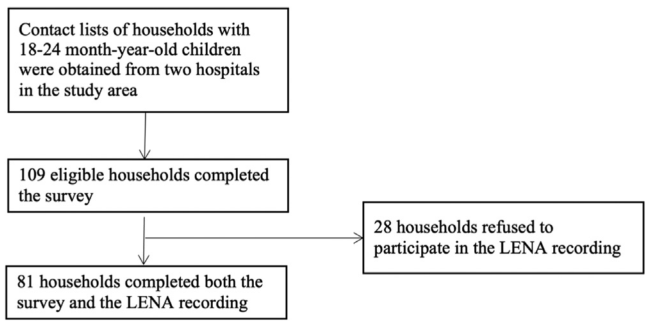

Fig. 1. The flowchart of sample selection of families with children between 18-24 months.

Report, 2020). In other words, although the peri-urban district is a rapidly developing, urbanized region, a large share of the population still has rural roots, as well as educational levels and incomes that are lower than those of urban residents who live in other first-tier districts.

Within the district, we used a three-step sampling strategy to select the study sample (Figure 1). First, we narrowed our sampling frame by randomly selecting two neighborhoods in the district. Next, with the help of the main public hospital in each neighborhood, we obtained lists of children in the target age range for this study (18-24 months). Finally, using this list, we randomly screened families on the list via phone calls to confirm whether they met the eligibility criteria to be considered rural migrants or urbanized farmers as defined by the Chinese government. A family was considered to be rural migrant: a.) if the *hukou* of either parent was outside of the provincial capital municipality and of rural status; and b.) if that parent had been living the peri-urban district for at least 6 months (National Working Committee on Children and Women (NWCCW) et al., 2018). On the other hand, a family was considered to be an urbanized farmer: a.) if the *hukou* of either parent was within the peri-urban district; b.) if all or part of the family's farmland had been reclaimed by the local government; and c.) if any remaining cultivated land that was contracted by the family was, on average, fewer than 200 square meters of cultivated land per adult (Hu, 2020).

Based on these criteria, a total of 109 households were enrolled in the survey. Within this group, 81 households agreed to participate in collection of home language environment data using LENA technology. Therefore, our final analytical sample included 81 households that participated in both the survey and the assessment of the home language environment. To reduce bias and improve the robustness of results, we conducted an analysis of sample attrition. Appendix Table A1 shows that there are no statistically significant differences between the demographic characteristics of households that participated in data collection of the home language environment (81 households) and those that did not (28 households).

2.3. Data collection

Data were collected in July 2020 by trained survey enumerators following a standardized data collection protocol. For each family, the child's primary caregiver was first identified as the individual most often responsible for the child's daily care (typically the child's mother or grandmother). The primary caregiver was initially administered a quantitative survey (described below) that had two main blocks of data: measures of child language development and a set of demographic characteristics of the child and family. After the survey was completed, the study team collected data on the home language environment using LENA audio technology.

2.3.1. The home language environment: LENA

Data on the home language environment were collected using LENA audio recorders and software. LENA is a recording and analysis system that quantitatively measures a child's home language environment (Ford et al., 2008). In other words, LENA measures linguistic input in a child's home environment. To perform the in-home measurement, each child's caregiver was given a fully charged LENA recorder, a specialized LENA shirt, and a LENA charger. A fully charged LENA recorder can record 16 continuous hours of audio data capturing a family's everyday home language environment. In our study, caregivers were instructed to record two 16-hour days that were representative of their child's typical at-home experience. Following LENA validation protocols (Zhang et al., 2015), the recorder was placed in the chest pocket of the specialized shirt that the child wore throughout the day. Caregivers were instructed to remove the LENA recorder and LENA-specialized shirt only when their child bathed or slept at night. Caregivers were also instructed to charge the LENA recorder between the first and second days of recording. After two full days of recording were completed, enumerators retrieved the LENA recorder and interviewed the caregiver to ensure compliance with the LENA recording protocol. According to post-recording questioning of the caregivers and analysis of the LENA recordings, all 81 participating families followed these protocols successfully.

After LENA recordings were collected from the family, they were analyzed using the LENA software to produce three standardized measures of the home language environment: adult word count (AWC), conversational turn count (CTC), and child vocalization count (CVC). AWC measures the number of adult words spoken near the child, excluding words from electronic devices such as a television. CTC refers to the number of conversational turns (alternation from one speaker to another) between an adult and the child. Finally, CVC measures the number of pre-speech or speech productions made by the child. All three measures of the home language environment have been reliably used by previous studies in the U.S., China, and other countries (Gilkerson et al., 2018; Canault et al., 2016; Pae et al., 2016; Zhang et al., 2015; Weisleder & Fernald, 2013). In our sample, families spoke Standard Mandarin and/or the provincial dialect of Mandarin. The provincial dialect comes from the Mandarin dialect and shares the same syllable structure as Standard Mandarin, making it widely distinguishable and easily interpreted (Zhang, 2007). Moreover, this provincial dialect of Mandarin is well-established as a Mandarin-based dialect, and more than that, bears a very close linguistic similarity to Mandarin (Tang & van Heuven, 2009).

Counts of the home language environment measures (AWC, CTC, and CVC) are reported as average counts across 12 hours of audio data. In our collection, we had households record 16 hours of audio data. The objective of restricting analysis to 12-hour segments was to improve comparability across families and capture a time window containing the most household interaction. Thus, we instructed families not to start recording until the child was awake, and not to stop recording until the child went to bed. To normalize 16 hours of recording into 12 hours of audio data, we applied Chebyshev polynomials transformation to normalize the distribution (Mason & Handscomb, 2002) and then used LASSO regression models to select the final Chebyshev polynomials model to transform the data (Neto, 2021). We then predicted residuals using the final Chebyshev polynomials model, estimating residualized count variables from the transformed data and rescaling residualized count variables back to the original count metric. Thus, outcomes (AWC, CTC, and CVC) were totals from first usable 12 hours of recordings. See the details of our LENA Data Management Protocol in Appendix File A1. This approach is consistent with previous studies that have used the LENA system to study the quality of the language environment of children (Cunha et al., 2020; Gilkerson et al., 2018).

Past literature has proven the versatility and reliability of LENA recordings across diverse linguistic contexts including American English, Spanish, French, Korean, Dutch, and Vietnamese when compared with trained human transcribers (Busch et al., 2018; ; Ganek & Eriks-Brophy, 2018; Gilkerson et al., 2018; Pae et al., 2016; Canault et al., 2016; Xu et al., 2009). LENA technology has also been used and validated for Mandarin-speaking populations in China (Ma et al., 2021, 2023; Zhang et al., 2015). The reliability of LENA for use in Mandarin Chinese was tested in Gilkerson et al. (2015), which found that correlations across LENA and human annotations were strong for AWC (r = 0.73), but not for CTC (r = 0.22).

2.3.2. Child language development

To measure child language development, we used the Mandarin version of the MacArthur-Bates Communicative Development Inventories (MCDI) (Hao et al., 2008). We utilized the parent-reported expressive vocabulary assessment of the MCDI for children between 16 and 30 months (our participating children were 18-24 months old). Using a list of 113 words, enumerators asked the child's primary caregiver whether their child could say each word in Mandarin or the provincial dialect; each word the child could say counted for one point. By comparing the parent-reported results with empirically determined cutoff scores established using the MCDI manual (Tardif et al., 2008), the status of a child's expressive language developmental progress (rate of delay) was determined. In contrast to LENA, the MCDI measures the outcomes of child language development, and indicates the level of delay in expressive vocabulary development. Rates of delay were established for each age group using group cutoff scores from Tardif et al. (2008) (any child under the 10th percentile of language development in their one-month age group was considered delayed) and then combined to produce a full sample cutoff for delay. The MCDI has been adapted to and validated in Mandarin Chinese (Tardif et al., 2008; Fenson et al., 2007) and proven reliable in studying early language development in China (Ma et al., 2021; Zhang et al., 2015). Specifically, three methods of reliability were used to assess the reliability of the MCDI, including internal consistency (how likely each of the items on the scale are measuring the same ability), test-retest reliability (how likely one is to receive a similar ability ranking when the same test is used at different times), and comparison statistics for a "fake" short form based on analyzing the same subset of items in the larger norming study sample (Tardif, 2008). First, internal consistency was calculated by computing Cronbach's alpha for each of the short form subscales, and Cronbach's alpha was more than 0.8, indicating that internal consistency was very good. Second, simple bivariate correlations between the various sections of each of the instruments were high (0.67-0.90) for short forms. Third, simple bivariate correlations (0.80-0.92) in "fake" version of the short form which included the data from only the items that appeared on the final short form were highly consistent with the short form.

2.3.3. Demographic information

The research team also collected information about the demographic characteristics of sample children and families. Child characteristics included the child's age (reported in months), gender, and whether the child was born prematurely. Family characteristics consisted of both the educational level and *hukou* status of each of the sample child's parents, the age of the mother, whether the mother was the primary caregiver, whether the father lived at home for at least six months during the previous 12 months prior to the survey, the number of adults in the household, a household asset index, and whether the family was legally characterized as urbanized farmers or rural migrants. We established a household asset index for participating families using polychoric principal components analysis based on whether the family owned or had access to running water, a flush toilet, a water heater, a washing machine, a computer, Internet, a refrigerator, an air conditioner, a motorbike/motorcycle, and a car/truck (Kolenikov & Angeles, 2009).

2.4. Statistical analysis

Our statistical analysis consists of four parts. First, to describe the quality and the heterogeneity of the home language environment, we graphically present the AWC, CTC, and CVC scores for the participating families in rank order. Second, we graphically present the distribution of child MCDI scores to examine the nature of child language development and prevalence of language delays. For both the LENA and MCDI distributions, we describe the means, standard deviations, and ranges of each of the key variables. Third, we analyze demographic characteristics associated with more and less robust home language environments by conducting t-tests comparing the demographic characteristics of children and households in the top and bottom terciles of AWC, CTC, and CVC. We also conduct Chi-squared tests to compare demographic characteristics (binary variables) between top terciles of AWC, CTC and CVC and bottom terciles of AWC, CTC and CVC to improve the robustness of results. Additionally, we have conducted multivariate simple linear regression analysis and multivariate multiple linear regression analysis to examine the potential relation between demographic characteristics and the language environment to ensure that the results of t-tests and Chi-squared tests are robust (Appendix Table A2 and A3). Finally, we conduct additional t-tests comparing the MCDI scores of children in the top and bottom terciles of AWC, CTC, and CVC to analyze the correlations between the home language environment and child language development. We used STATA 16.1 to perform all statistical analyses. P-values at or below 0.05 are considered statistically significant.

Table 1
Descriptive Statistics of Child and Household Characteristics (N = 81)

Variables	Mean/N (SD/Percent)
Child characteristics	
Age (months)	21.10 (1.60)
Gender (1 = boy)	46 (56.79%)
Prematurity (1 = yes)	8 (9.88%)
Household characteristics	
Age of mother (years)	29.11 (4.69)
Maternal education (1 = completed high school or above)	62 (76.54%)
Mother has job (1 = yes)	51 (62.96%)
Mother is primary caregiver (1 = yes)	46 (56.79%)
Mother has urban *hukou* (1 = yes)	25 (30.86%)
Father has urban *hukou* (1 = yes)	22 (27.16%)
Paternal education (1 = completed high school or above)	64 (79.01%)
Father lived at home during most of last year (1 = yes)	64 (79.01%)
Number of adults in the household	3.05 (1.01)
Asset index (PCA score)	0.00 (1.28)
Urbanized farmer (1 = yes)	47 (58.02%)

Note. *Hukou* refers to legal residency status as determined by the Chinese government and can be either urban or rural; "PCA" means principal component analysis; "SD" means standard deviation.

3. Results

3.1. Descriptive statistics of children and families

Table 1 presents the descriptive statistics of the 81 sample children and their families. The average age of all the children in the study was 21 months old. More than half (57%) of children in the study were male, and 10% were born prematurely. For mothers in the study, their average age was 29 years, and about 77% of mothers had completed high school or above. In 57% of the families, the mother was the child's primary caregiver. Almost two-thirds (63%) of the sample mothers had a job at the time of the survey. Approximately 69% of mothers and 73% of fathers had rural *hukou* (legal residency status). About 79% of fathers had completed high school, and in 79% of families the father had lived at home for at least six months of the past year. The average number of adults living in each household was three. Finally, urbanized farmers accounted for 58% of the sample, and rural migrants accounted for the remaining 42%.

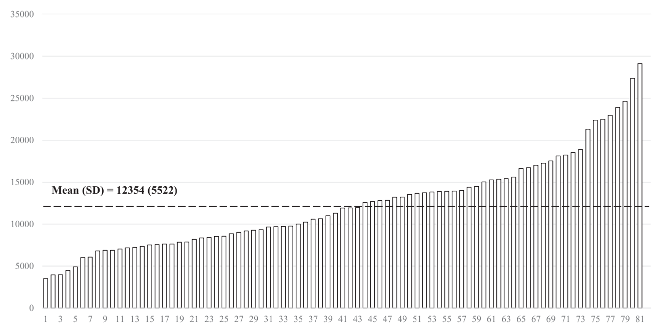

Fig. 2. The distribution of Adult Word Count (AWC) (N = 81). Source: authors' survey. Note: "SD" means standard deviation.

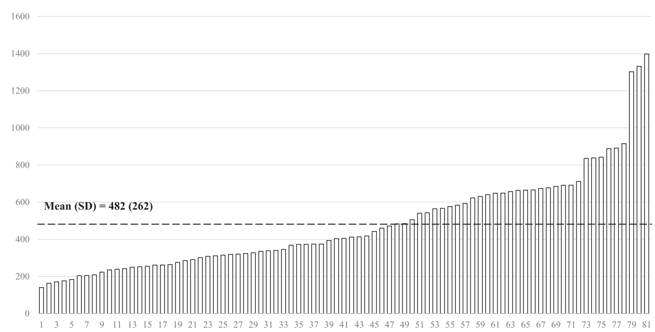

Fig. 3. The distribution of Conversational Turn Count (CTC) (N = 81). Source: authors' survey. Note: "SD" means standard deviation.

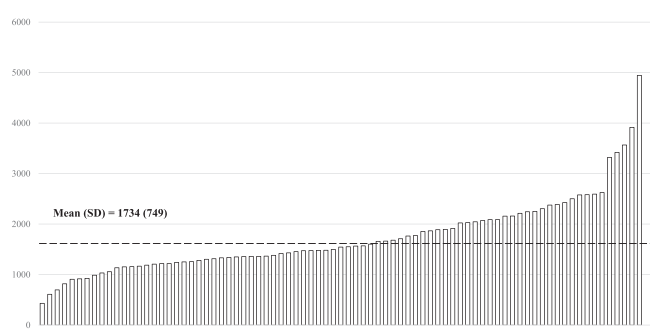

Fig. 4. The distribution of Child Vocalization Count (CVC) (N = 81). Source: authors' survey. Note: "SD" means standard deviation.

3.2. The home language environment

Figures 2, 3, and 4 show the distributions of the three measures among the sample families of the home language environment—AWC, CTC, and CVC, respectively. As reported in Figure 2, the mean AWC of the sample was 12354 words (SD = 5522), and the difference in scores between the highest-ranking (AWC = 29118) and lowest-ranking (AWC = 3515) participants was 8.3-fold. For CTC, shown in Figure 3, the mean score was 482 (SD = 262) with a 9.9-fold difference between the highest score (CTC = 1397) and the lowest score (CTC = 141). Finally, Figure 4 shows that the mean score for CVC was 1734 (SD = 749), and the difference in CVC between the highest (CVC = 4945) and lowest (CVC = 431) scores was 11.5-fold.

3.3. Child language development

Figure 5 shows the distribution of MCDI scores among the children in the sample. The mean MCDI score was 37, with a standard deviation of 27. About 25% of children were below the CDI cutoff for proficient language development (Tardif et al., 2008). The results also show a large

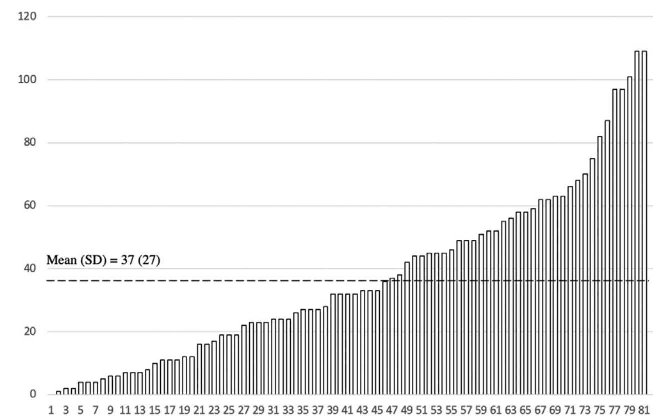

Fig. 5. The distribution of MacArthur-Bates Communicative Development Inventories (MCDI) scores (N = 81).

variation in MCDI scores among the sample children, with a 6.8-fold ($p < 0.001$) difference between the first-quartile (MCDI = 10) and third-quartile (MCDI = 68) participants.

3.4. What is associated with variation in the home language environment?

Tables 2, 3, and 4 present the results of our descriptive t-tests and Chi-squared tests comparing demographic characteristics of sample families in the top and bottom terciles of AWC, CTC, and CVC, respectively. Table 2 compares households in the top and bottom terciles of AWC. The results show that whether the mother had a job was significantly associated with higher AWC. Children of mothers who had jobs were significantly more likely to have AWC scores in the top tercile than those whose mothers did not have jobs ($p < 0.05$); specifically, 78% of mothers with jobs were in the top tercile of AWC. In other words, mothers who were employed at the time of the survey were more likely to produce a greater number of adult words than mothers who were unemployed. Compared to employed mothers in the bottom tercile of AWC, there were 30% more employed mothers in the top tercile.

The results of our descriptive t-test and Chi-squared tests comparing the demographic characteristics of families in the top and bottom terciles of CTC are shown in Table 3. Child's gender and household asset index showed large and significant variation among CTC scores. For example, male children made up 70% of the scores in the bottom tercile and 30% of scores in the top tercile of CTC. In other words, female children were more likely to be in the top tercile of CTC than their male counterparts ($p < 0.01$). Similarly, the difference in the asset index was large and significant ($p < 0.01$) among CTC scores, as children in the top tercile of CTC had higher levels of household asset values than those in the bottom tercile.

Table 4 compares the characteristics of sample families in the top and bottom terciles of CVC. Similar to the results from CTC in Table 3, the differences in gender and household asset level were large and significant between the top and bottom terciles of CTC scores. Female children and children from families with higher asset index scores were significantly more likely to be in the top tercile of CVC ($p < 0.01$ and $p < 0.05$, respectively). Furthermore, difference in age was large and significant between the top and bottom terciles of CVC scores ($p < 0.01$). The average age of children in the top tercile was 21.7 months, while the average age of children in the bottom tercile was 20.5 months.

Appendix Table A2 illustrates correlations between household characteristics and the home language environment based on a simple linear regression model. Results show that child's age was statistically significantly correlated with CTC and CVC ($p < 0.05$ and $p < 0.01$, respectively). The table also shows that child's gender was statistically significantly correlated with CTC and CVC ($p < 0.01$). Additionally, we found a positive and significant association between maternal employment and

Table 2
Differences in Demographic Characteristics Between Top and Bottom Terciles (⅓) of Adult Word Count (AWC)

Variables	Bottom ⅓ of AWC	Top ⅓ of AWC	T-test	Chi-square test	
				Pearson chi-square	P-value
	(1)	(2)	(3) = (2) – (1)	(4)	(5)
Child characteristics					
Age (months)	21.121	21.235	0.114		
	(1.542)	(1.475)	[0.439]		
Gender (1 = boy)	19	12		3.711	0.054
	(70.370%)	(44.444%)			
Prematurity (1 = yes)	3	4		0.164	0.685
	(11.111%)	(14.815%)			
Household characteristics					
Age of mother (years)	28.667	29.037	0.370		
	(4.261)	(5.110)	[1.287]		
Maternal education (1 = completed high school or above)	22	20		0.429	0.513
	(81.481%)	(74.074%)			
Mother has a job (1 = yes)	13	21		5.082	0.024
	(48.148%)	(77.778%)			
Mother is primary caregiver (1 = yes)	16	12		1.187	0.276
	(59.259%)	(44.444%)			
Mother has urban *hukou* (1 = yes)	6	9		0.831	0.362
	(22.222%)	(33.333%)			
Father has urban *hukou* (1 = yes)	7	6		0.101	0.750
	(25.926%)	(22.222%)			
Paternal education (1 = completed high school or above)	22	20		0.429	0.513
	(81.481%)	(74.074%)			
Father lived at home for at least 6 months in the past 12 months (1 = yes)	20	23		1.028	0.311
	(74.074%)	(85.185%)			
Number of adults in the household	2.778	3.074	0.296		
	(1.050)	(0.874)	[0.272]		
Asset index (PCA score)	-0.432	0.111	0.543		
	(1.888)	(0.869)	[0.341]		
Urbanized farmers (1 = yes)	12	16		1.187	0.276
	(44.444%)	(59.259%)			

Note. Data source is author's survey. Standard deviations/percent in parentheses, standard errors in brackets; "PCA" means principal component analysis.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3
Differences in Demographic Characteristics Between Top and Bottom Terciles (⅓) of Conversational Turn Count (CTC)

Variables	Bottom ⅓ of CTC	Top ⅓ of CTC	T-test	Chi-square test	
				Pearson chi-square	P-value
	(1)	(2)	(3) = (2) – (1)	(4)	(5)
Child characteristics					
Age (months)	20.759	21.579	0.820*		
	(1.511)	(1.522)	[0.429]		
Gender (1 = boy)	19	8		8.963	0.003
	(70.370%)	(29.63%)			
Prematurity (1 = yes)	1	4		1.984	0.159
	(3.704%)	(14.815%)			
Household characteristics					
Age of mother (years)	29.148	29.296	0.148		
	(4.538)	(4.697)	[1.291]		
Maternal education (1 = completed high school or above)	20	22		0.429	0.513
	(74.074%)	(81.481%)			
Mother has a job (1 = yes)	15	20		2.030	0.154
	(55.556%)	(74.074%)			
Mother is primary caregiver (1 = yes)	14	13		0.074	0.785
	(51.852%)	(48.148%)			
Mother has urban *hukou* (1 = yes)	6	10		1.421	0.233
	(22.222%)	(37.037%)			
Father has urban *hukou* (1 = yes)	10	8		0.333	0.564
	(37.037%)	(29.630%)			
Paternal education (1 = completed high school or above)	21	23		0.491	0.484
	(77.778%)	(85.185%)			
Father lived at home for at least 6 months in the past 12 months (1 = yes)	23	25		0.750	0.386
	(85.185%)	(92.593%)			
Number of adults in the household	3.074	3.444	0.370		
	(1.035)	(0.892)	[0.263]		
Asset index (PCA score)	-0.486	0.455	0.941***		
	(1.886)	(0.384)	[0.336]		
Urbanized farmers (1 = yes)	15	18		0.701	0.402
	(55.556%)	(66.667%)			

Note. Data source is author's survey. Standard deviations/percent in parentheses, standard errors in brackets; "PCA" means principal component analysis.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4
Differences in Demographic Characteristics Between Top and Bottom Terciles (⅓) of Child Vocalization Count (CVC)

Variables	Bottom ⅓ of CVC	Top ⅓ of CVC	T-test	Chi-square test	
				Pearson chi-square	P-value
	(1)	(2)	(3) = (2) - (1)	(4)	(5)
Child characteristics					
Age (months)	20.532	21.741	1.209***		
	(1.495)	(1.433)	[0.418]		
Gender (1 = boy)	21	6		16.667	0.000
	(77.778%)	(22.222%)			
Prematurity (1 = yes)	1	4		1.984	0.159
	(3.704%)	(14.815%)			
Household characteristics					
Age of mother (years)	29.556	28.370	-1.185		
	(5.056)	(3.127)	[1.283]		
Maternal education (1 = completed high school or above)	19	23		1.714	0.190
	(70.370%)	(85.185%)			
Mother has a job (1 = yes)	16	18		0.318	0.573
	(59.259%)	(66.667%)			
Mother is primary caregiver (1 = yes)	16	15		0.076	0.783
	(59.259%)	(55.556%)			
Mother has urban *hukou* (1 = yes)	6	11		2.146	0.143
	(22.222%)	(40.741%)			
Father has urban *hukou* (1 = yes)	8	8		0.000	1.000
	(29.630%)	(29.630%)			
Paternal education (1 = completed high school or above)	20	23		1.028	0.311
	(74.074%)	(85.185%)			
Father lived at home for at least 6 months in the past 12 months (1 = yes)	23	25		0.750	0.386
	(85.185%)	(92.593%)			
Number of adults in the household	3.000	3.333	0.333		
	(1.038)	(0.961)	[0.272]		
Asset index (PCA score)	-0.312	0.470	0.782**		
	(1.734)	(0.368)	[0.340]		
Urbanized farmers (1 = yes)	15	17		0.307	0.580
	(55.556%)	(62.963%)			

Note. Data source is author's survey. Standard deviations/percent in parentheses, standard errors in brackets; "PCA" means principal component analysis.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Appendix Table A1
Differences in Demographic Characteristics Between LENA Consent and Reject Households

Variables	Participated in LENA recording	Did not participate in LENA recording	Difference
	(1)	(2)	(3) = (2) - (1)
Child characteristics			
Age (months)	21.103	21.304	0.200
	(1.596)	(1.420)	[0.341]
Gender (1 = boy)	46	13	-0.104
	(56.790%)	(46.429%)	[0.110]
Prematurity (1 = yes)	8	2	-0.027
	(9.877%)	(7.143%)	[0.064]
Household characteristics			
Age of mother (years)	29.111	29.893	0.782
	(4.685)	(3.871)	[0.985]
Maternal education (1 = completed high school or above)	62	23	0.056
	(76.543%)	(82.143%)	[0.092]
Mother has a job (1 = yes)	51	17	-0.022
	(62.963%)	(60.714%)	[0.107]
Mother is primary caregiver (1 = yes)	46	18	0.075
	(56.790%)	(64.286%)	[0.109]
Mother has urban *hukou* (1 = yes)	25	6	-0.094
	(30.864%)	(21.429%)	[0.099]
Father has urban *hukou* (1 = yes)	22	8	0.014
	(27.160%)	(28.571%)	[0.099]
Paternal education (1 = completed high school or above)	64	23	0.031
	(79.012%)	(82.143%)	[0.089]
Father lived at home for at least 6 months in the past 12 months (1 = yes)	64	21	-0.040
	(79.012%)	(75.000%)	[0.092]
Number of adults in the household	3.049	2.821	-0.228
	(1.011)	(0.723)	[0.208]
Asset index (PCA score)	0.014	-0.040	-0.053
	(1.236)	(1.524)	[0.288]
Urbanized farmers (1 = yes)	47	11	-0.187*
	(58.025%)	(39.286%)	[0.109]

Note. Data source is author's survey. *hukou* refers to legal residency status as determined by the Chinese government and can be either urban or rural; "PCA" means principal component analysis; standard deviations/percent in parentheses, standard errors in brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Appendix Table A2
Associations Between Home Language Environment Measurements and Demographic Characteristics (Simple Linear Regression Model)

Variables	AWC (1)	CTC (2)	CVC (3)
Child characteristics			
Age (months)	178.604	36.529**	142.032***
	[388.710]	[17.994]	[50.313]
Gender (1 = boy)	-1,988.58	-162.892***	-543.933***
	[1226.231]	[56.197]	[157.591]
Prematurity (1 = yes)	729.328	14.644	61.009
	[2067.990]	[98.130]	[280.613]
Household characteristics			
Age of mother (years)	119.356	1.456	-11.674
	[131.933]	[6.287]	[17.938]
Maternal education (1 = completed high school or above)	-661.111	91.408	359.693*
	[1455.316]	[68.334]	[193.451]
Mother has a job (1 = yes)	2,572.966**	112.267*	194.55
	[1245.447]	[59.305]	[172.033]
Mother is primary caregiver (1 = yes)	-2,413.389*	-93.582	-113.049
	[1216.541]	[58.164]	[168.577]
Mother has urban *hukou* (1 = yes)	1,310.73	81.501	112.543
	[1328.536]	[62.721]	[180.850]
Father has urban *hukou* (1 = yes)	-320.052	-19.873	-8.861
	[1387.755]	[65.793]	[188.278]
Paternal education (1 = completed high school or above)	-0.111	92.041	286.767
	[1516.289]	[71.155]	[203.104]
Father lived at home for at least 6 months in the past 12 months (1 = yes)	94.683	1.137	120.379
	[1516.251]	[71.904]	[205.204]
Number of adults in the household	750.176	51.634*	114.614
	[608.603]	[28.552]	[82.329]
Asset index (PCA score)	630.527	49.838**	128.648**
	[480.557]	[22.343]	[64.273]
Urbanized farmers (1 = yes)	1,139.60	47.835	0.248
	[1244.559]	[59.086]	[169.690]
Observations	81	81	81

Note. "AWC" means adult word count; "CTC" means conversational turns count; "CVC" means child vocalization count; *hukou* refers to legal residency status as determined by the Chinese government and can be either urban or rural; "PCA" means principal component analysis; standard errors in brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

AWC ($p < 0.05$), as well as an association between family asset and CTC, CVC ($p < 0.05$, both). Appendix Table A3 illustrates correlations between household characteristics and the home language environment based on a multiple linear regression model. Results show that child's age was not statistically significantly correlated with AWC or CTC but was statistically significantly correlated with CVC at $p < 0.05$. The table also shows that child gender was statistically significantly correlated with CTC and CVC at $p < 0.05$ and $p < 0.01$, respectively. Additionally, we also found positive associations between maternal job status and family asset index with the home language environment, but they were all statistically insignificant.

3.5. The home language environment and early language development

Figures 6, 7, and 8 show the results of the *t*-tests comparing child language development between the top and bottom terciles of AWC, CTC, and CVC, respectively. Children in the top tercile of AWC had significantly higher MCDI scores than children in the bottom tercile of AWC (Figure 6). Specifically, children in the top tercile of AWC had an average MCDI score of 51.4, whereas children in the bottom tercile of AWC had an average score of 27.4, representing a difference of 24 ($p = 0.001$). Figure 7 shows the difference in MCDI scores between top and bottom terciles of CTC. The difference was 29.6, as children in the top tercile of CTC had an average MCDI score of 58.2 and those in the bottom tercile of CTC had an average MCDI score of 28.6. Finally, for CVC (Figure 8), the participants in the top tercile had an average MCDI score of 57.8, compared to an average MCDI score of 22.6 of participants in the bottom tercile; thus, the difference between top-tercile and bottom-tercile MCDI scores is 35.2. In other words, the differences between the top and bottom terciles for all three measures were statistically significant ($p = 0.003$ for AWC, and $p < 0.001$ for both CTC and CVC). Furthermore, holding control variables constant, we found a strong association between all three measures of the home language environment and MCDI. Appendix Table A4 illustrates that AWC and CTC were both statistically significantly and positively associated with child language development outcomes at $p < 0.01$, and CVC was statistically significant and positively associated with child language outcomes at $p < 0.05$. Over-

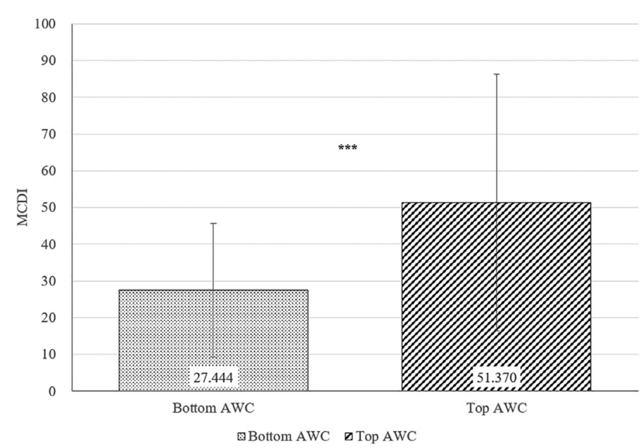

Fig. 6. Difference in MacArthur-Bates Communicative Development Inventories (MCDI) scores between top and bottom terciles of Adult Word Count (AWC) ($p = 0.003$).

Appendix Table A3
Associations Between Home Language Environment Measurements and Demographic Characteristics (Multiple Linear Regression Model)

Variables	AWC (1)	CTC (2)	CVC (3)
Child characteristics			
Age (months)	-10.356	25.717	119.982**
	[404.114]	[17.996]	[51.505]
Gender (1 = boy)	-2,557.852*	-162.463**	-491.224***
	[1,379.040]	[61.412]	[175.762]
Prematurity (1 = yes)	-818.512	-6.042	129.522
	[2,325.239]	[103.548]	[296.357]
Household characteristics			
Age of mother (years)	194.102	6.270	2.749
	[152.369]	[6.785]	[19.420]
Maternal education (1 = completed high school or above)	-675.410	81.658	337.829
	[1,916.529]	[85.347]	[244.266]
Mother has a job (1 = yes)	1,326.358	71.884	208.589
	[1,543.574]	[68.739]	[196.732]
Mother is primary caregiver (1 = yes)	-2,305.563	-72.454	-22.650
	[1,549.828]	[69.017]	[197.529]
Mother has urban *hukou* (1 = yes)	1,321.222	40.007	-31.140
	[1,776.566]	[79.115]	[226.427]
Father has urban *hukou* (1 = yes)	-2,603.255	-104.701	-26.164
	[1,734.245]	[77.230]	[221.033]
Paternal education (1 = completed high school or above)	-1,105.230	-44.781	-116.532
	[2,042.131]	[90.941]	[260.274]
Father lived at home for at least 6 months in the past 12 months (1 = yes)	721.554	-34.592	-90.407
	[1,766.438]	[78.664]	[225.136]
Number of adults in the household	-6.132	23.375	54.360
	[720.224]	[32.073]	[91.794]
Asset index (PCA score)	643.349	34.024	66.977
	[550.298]	[24.506]	[70.137]
Urbanized farmers (1 = yes)	978.578	34.725	-45.751
	[1,518.455]	[67.620]	[193.530]
Observations	81	81	81
R-squared	0.179	0.276	0.275

Note. "AWC" means adult word count; "CTC" means conversational turns count; "CVC" means child vocalization count; *hukou* refers to legal residency status as determined by the Chinese government and can be either urban or rural; "PCA" means principal component analysis; standard errors in brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

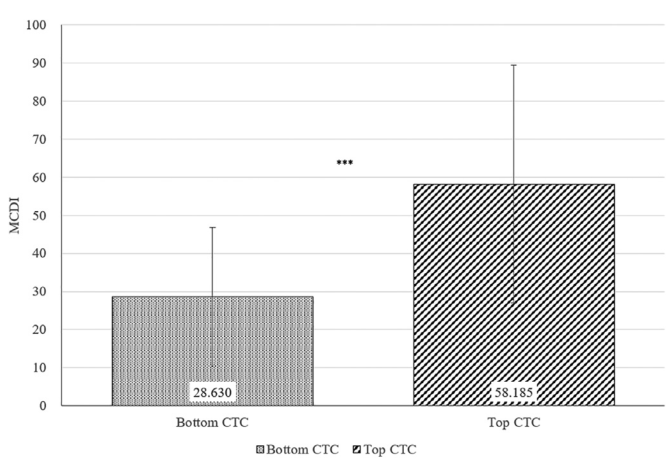

Fig. 7. Difference in MacArthur-Bates Communicative Development Inventories scores between top and bottom terciles of Conversational Turn Count (CTC) ($p < 0.001$).

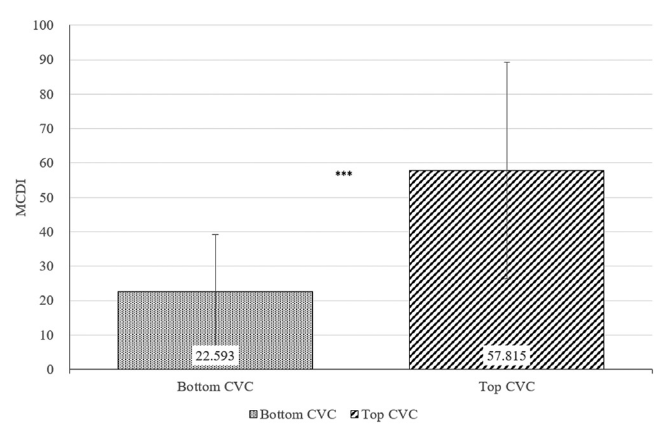

Fig. 8. Difference in MacArthur-Bates Communicative Development Inventories (MCDI) scores between top and bottom terciles of Child Vocalization Count (CVC) ($p < 0.001$).

all, these findings show that the home language environment and child language development have a strong positive association.

4. Discussion

Findings from an expansive literature underscore the importance of children's early cognitive, socioemotional, and linguistic development for later-life success (Romeo et al., 2018; Heckman & Mosso, 2014; Conger et al., 2010; Cunha et al., 2010). In particular, delays in speech and language skills, if left untreated, have a high potential of persisting in a large percentage of affected children, impacting future outcomes in educational attainment and employment (Law et al., 1998). Because low-resourced children living in LMIC settings already face an especially high risk of developmental delay in the absence of interventions, and given the scarcity of research on early language development outside of

Appendix Table A4
Correlations between home language environment measurements, demographic characteristics, and child language development (holding control variables constant)

Variables	MCDI		
	(1)	(2)	(3)
Home language environment measurements			
AWC (Adult Word Count)	0.002***		
	[0.000]		
CTC (Conversational Turn Count)		0.043***	
		[0.011]	
CVC (Child Vocalization Count)			0.008**
			[0.004]
Child characteristics			
Age (months)	5.747***	4.619***	4.752***
	[1.569]	[1.564]	[1.743]
Gender (1 = boy)	-19.740***	-17.278***	-20.286***
	[5.492]	[5.529]	[6.046]
Prematurity (1 = yes)	-4.026	-5.220	-6.535
	[9.037]	[8.865]	[9.654]
Household characteristics			
Age of mother (years)	-0.493	-0.418	-0.170
	[0.599]	[0.585]	[0.632]
Maternal education (1 = completed high school or above)	6.619	1.897	2.669
	[7.448]	[7.357]	[8.060]
Mother has a job (1 = yes)	-0.906	-1.649	-0.247
	[6.027]	[5.933]	[6.454]
Mother is primary caregiver (1 = yes)	2.024	1.052	-1.889
	[6.118]	[5.958]	[6.426]
Mother has urban *hukou* (1 = yes)	-2.124	-1.502	0.477
	[6.927]	[6.786]	[7.366]
Father has urban *hukou* (1 = yes)	11.152	11.042	6.739
	[6.848]	[6.703]	[7.191]
Paternal education (1 = completed high school or above)	5.568	5.536	4.553
	[7.947]	[7.800]	[8.479]
Father lived at home for at least 6 months in the past 12 months (1 = yes)	-2.853	-0.079	-0.835
	[6.867]	[6.744]	[7.332]
Number of adults in the household	0.486	-0.533	0.032
	[2.796]	[2.757]	[2.994]
Asset index (PCA score)	-0.492	-0.816	0.106
	[2.159]	[2.128]	[2.297]
Urbanized farmers (1 = yes)	-6.050	-5.808	-3.938
	[5.914]	[5.801]	[6.298]
Observations	81	81	81
R-squared	0.503	0.520	0.433

Notes. MCDI = MacArthur-Bates Communicative Development Inventories; MCDI is a measure of child language development; *hukou* refers to legal residency status as determined by the Chinese government and can be either urban or rural. Standard errors in brackets; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Western or high-income contexts, investigating the quality of the home language environment for an underrepresented population was the aim of the current study (Black et al., 2017; Knudsen et al., 2006).

This observational study of 81 children aged 18-24 months from rural migrant and urbanized farmer families in China reveals the first-ever findings on variations in the home language environment and early child language development of low-SES, peri-urban families in a non-Western, developing setting. Overall, we find that the home language environment of our study's peri-urban sample had less linguistic interaction between adults and children than the home language environment of families living in both poor, rural China and in more highly resourced, urban China. In addition, sample families had large and substantial variation across all measures of the home language environment (i.e., some families provided a relatively engaging home language environment, while some families did not), despite having similarly disadvantaged demographic characteristics. Finally, our results show a strong positive association between home language environment measures and child language skills.

4.1. Variations in the home language environment of peri-urban families

The results of the study demonstrate that, despite having relatively similar demographic characteristics, such as low levels of parental educational attainment and living in disadvantaged, low-income neighborhoods, there was substantial variation across all measures of the home language environment among peri-urban families. Between the highest and lowest ranking child for each home language environment measure, the sample had a 8.3-fold difference in the number of adult words spoken, a 9.9-fold difference in the number of adult-child conversations had, and a 11.5-fold difference in the number of child vocalizations made. We note that the substantial variation found in this study is consistent with the findings of other studies examining the home language environment of low-SES communities (Ma et al., 2021; Weisleder & Fernald, 2013; Hurtado et al., 2008; Pan et al., 2005); however, this study's variations were greater than those found in low-SES samples from rural communities in China. For example, variation in the home language environments of rural Shaanxi households in Ma et al. (2021) was 5-fold between the highest and lowest ranking child in both AWC and CVC measures, and 6.6-fold in the CTC measure—all smaller than the differences identified in our results. Yet the findings of Weisleder & Fernald (2013), a study that analyzed the home language environments of 29 low-SES Spanish-learning children in the U.S., had a 15-fold difference between the highest and lowest ranking child in terms of AWC—a variation somewhat greater than that of this study sample. In other words, substantially heterogeneous home language environment results within similarly disadvantaged, low-SES group are seemingly common across diverse linguistic, cultural, and geographic contexts. A potential reason for this may be found in Heath (1983), which suggests that because caregivers from different socio-cultural groups hold different beliefs about their roles in child language development, their children's home language environments may have marked differences outside of SES-level differences.

Evidence from this study also indicates that the quality of the home language environment of peri-urban families was even lower than that of families in underdeveloped, low-SES rural China. In a sample of 38 children aged 20-27 months from rural Shaanxi Province in northwestern China, Ma et al. (2021) reported that the sample children heard, on average, 13428 adult words, engaged in 559 conversational turns, and generated 2140 vocalizations throughout 12-hour recordings. In comparison, this sample's peri-urban families showed weaker results: on average, sample children heard 12354 adult words, engaged in 482 conversational turns, and produced 1734 vocalizations during the same time duration. Based on this comparison, we believe that the home language environments of rural migrant and urbanized farmer families are substantially less linguistically engaging (fewer adult words spoken and fewer adult-child conversations had) than those of families from nationally designated poverty counties of rural China.

One potential reason explaining this rather surprising finding may be that, despite living in more urban environments, rural migrants and urbanized farmers face unique challenges in their living conditions that may directly and indirectly influence the way they interact with their children, challenges that caregivers from rural areas may not typically encounter. For example, rural migrants in China may face social exclusion after moving from rural to urbanized regions; struggle with securing permanent residency status; experience family separation; and suffer more from mental health issues and limited access to health care (Chen et al., 2016; Zhan, 2011; Hesketh et al., 2008; Wong & Song, 2008). Indeed, given the institutional barriers to urban social welfare benefits in China, the marginalization of rural migrants typically lasts for extended periods of time and has been found to contribute to widening urban income inequality (Chen et al., 2016). Urbanized farmers may also suffer from more severe health problems and lack access to adequate health care, both of which have been shown to negatively affect mental health and social stability (Wang et al., 2019)(). Thus, while peri-urban regions potentially offer more work opportunities and economic mobility, rural caregivers who move to such areas may be preoccupied by a variety of external challenges that risk reducing their investment in the home language environment of their children.

Furthermore, this study found overall low levels of language development (on average) alongside high levels of variation in language development (measured by MCDI) among children in peri-urban families. Approximately 25% of this study's sample children were below the MCDI cutoff for proficient language development (Tardif et al., 2008). While this number may seem low in comparison to other studies finding rates of language delay in rural China as high as 46% (Emmers et al., 2020), the sample had high heterogeneity, with a 6.8-fold difference between participants scoring in the first and third quartiles: On average, the first quartile participant had a MCDI score of 10, while the third quartile participant had a score of 68. As stated before, our observed variation is also substantially larger than the variation seen in the poor rural China sample from Ma et al. (2021), which reported an approximately 2-fold difference between the first (MCDI = 28) and third quartiles (MCDI = 55). Thus, this is further evidence for the surprising heterogeneity of children's early language development within a relatively demographically homogenous disadvantaged population. An implication of this finding is that there exists considerable variability in caregiver language engagement with their children outside of factors related to socioeconomic class. Although not assessed in the current study, examples of potential factors could be caregiver beliefs or caregiver emotional well-being (Conger et al., 1984). Nevertheless, low-SES groups, especially those living in peri-urban China, remain at a heightened risk of language development delays that can negatively influence children's long-term educational and life outcomes.

4.2. Influence of maternal employment on language development in peri-urban families

Our findings also indicate that certain demographic characteristics are associated with more interactive home language environments. In other words, specific child and family characteristics encourage more adult speech, adult-child conversations, and child vocalizations; identifying these characteristics may thus help researchers to take steps toward improving individual home language environments for peri-urban populations.

Our results show that a child whose mother had a job was more likely to hear more adult speech (higher AWC) than a child whose mother was unemployed. Although the number of adult-child conversations was not significantly correlated to maternal job status (only adult words), there was a substantial difference between the upper and bottom terciles of CTC (74% of employed mothers vs. 56%). This finding is somewhat unique to our sample, as past research in Western and high-SES samples has not focused specifically on maternal employment or found much evidence of its significance.

One plausible reason for our study findings linking maternal employment and language outcomes is that employed mothers may have higher levels of educational attainment than those who are unemployed; since a mother's education tends to be positively associated with stronger parenting practices (Yue et al., 2017), employed and more-educated mothers may speak to and interact more often with their children. Coddington et al. (2014) also identified positive correlations between maternal education and child vocabulary outcomes. Reinforcing the link between maternal employment and maternal education level, we note that we find significant associations between maternal employment and maternal educational attainment (college level and above) for our sample at the 10% significance level.

Another possible reason behind the significant association between maternal employment and AWC may be that employed mothers are more likely to speak more with other adults in the home environment. Given that AWC measures the sheer volume of words spoken in range of the child wearing the LENA recorder, mothers who are employed may have more news to share with other members of the household after a day of work. Given that employment can lead to larger social circles, greater social engagement, and more social support among female caregivers (Shen et al., 2022; Drummond et al., 2017), work activity outside of the household could be leading to a greater quantity of language input for children in the home environment.

4.3. Influence of gender on language development in peri-urban families

Our findings also suggest significant early language differences between boys and girls in relation to their home language environments. One potential reason for this may be that female children are more likely to be exposed to a more linguistically diverse home language environment than male children. In our study, female sample children on average heard more speech from adults, had more conversations with adults, and made more vocalizations than their male peers. Although some research identified insignificant differences between boys and girls in their early language experiences (Eriksson & Berglund, 1999), our finding is consistent with the trend in the international literature suggesting that young girls and boys display significant differences in their early language development and in their home language experiences. In a review of over 82,000 administrations of the MCDI, results showed that girls aged 18–30 months produced significantly more words than did boys in 25 out of 29 languages (Frank et al., 2021). Results from Western samples also suggest that language milestones, such as a child's first word, may appear earlier in girls than in boys (Bornstein et al., 2004; Galsworthy et al., 2000; Maccoby and Jacklin, 1978). In addition, young girls tend to produce more words and longer sentences, be more verbal, and demonstrate more facility with grammatical rules in comparison to boys of the same age (Simonsen et al., 2014; Eriksson et al., 2012).

On the one hand, biological differences may be an underlying factor in the differences between boys' and girls' early language use (Bornstein et al., 2004; Galsworthy et al., 2000). On the other hand, many studies suggest that the home language environments of boys and girls simply look different, with evidence suggesting that caregivers—specifically maternal caregivers, who have been documented to speak more to female children versus male children at birth—converse with girl children more often than they do with boy children, and engage in more symbolic and linguistically structured play with girls (Bleses et al., 2018; Johnson et al., 2014; Leaper et al., 1998). Compounding that, caregivers often perceive girls to be more mature than boys, even at a young age, creating feedback loops for engaging in more conversations with girls (Bornstein et al., 2004; Lovas, 2011). We note, however, that studies such as Huttenlocher et al. (1991) present conflicting findings, detecting no significant difference in the number of words spoken by maternal caregivers to girl children compared to boy children; thus, they attributed gender differences in early language acquisition to early capacity differences.

As a potential additional reason for this study's finding on gender differences, we consider China's history of son-preference. In traditional Chinese culture, girls were mainly prescribed to housekeeping and childcare gender roles, while economic and social mobility and responsibility was placed on boys; this influenced how the genders were socialized. However, as China's urbanization increased and more families began living in urban areas, the culture changed. After the introduction of the one-child policy, urban China in the 1990s saw a major shift in culture surrounding child preference, as girls in cities received more parental investment in their education, which was associated with greater economic and social value placed on daughters. In other words, the playing field leveled between sons and daughters to some degree. Given our sample's placement in peri-urban areas, perhaps the influence of urban Chinese culture and transforming gender roles has exerted some influence on how caregivers perceive and linguistically engage with daughters, partially explaining why the girl children in our sample appear to hear more adult words, have more adult-child conversations, and speak more than boys. Coupled with biological factors that influence earlier speech in girls than boys, we believe that changes in traditional gender role perception may have spurred on more engaging home language environments for young girls in peri-urban China to a certain degree.

4.4. Influence of SES on language development in peri-urban families

Confirming past research from both high- and low-SES samples, as well as Western and non-Western samples, our data suggest that the home language environment is significantly associated with the resources available to families—that is, a family's asset index. Overall, our finding suggests that low-resourced families have lower quality home language environments, leaving children from such households at the greatest risk of less linguistic stimulation and worse ECD outcomes. Based on our findings, compared to the amount of adult speech, the number of adult-child conversations and the number of child vocalizations may be more directly connected to a family's socioeconomic resources. Within our sample, children who came from families with higher asset index scores (meaning households that had access to more household items such as running water, internet access, personal transportation vehicles, etc.) had more conversations with adults and made more vocalizations. We find no statistically significant associations between the number of adult words and household resources. In the literature on the home language environment, it is widely accepted that children from lower-SES backgrounds hear fewer adult words and less complex language than their higher-SES peers (Rowe, 2008; Hart & Risley, 1995). In fact, compared to an even younger sample (5-30 months) from high-SES families in urban China, we find that children from our low-SES peri-urban sample only hear about half as many adult words as their higher-SES peers, and have only half as many conversations with adults (Zhang et al., 2015). When comparing caregiver-child verbal interactions between mothers from different SES backgrounds, Hoff (2003) found that mothers from higher-SES families produced more topic-continuing replies to their children than mothers from lower-SES families. In other words, mothers from higher-SES families engaged in more conversational speech with their children than mothers from lower-SES families.

If we apply the theory of the investment model (that families with more resources are able to invest more time, money, and energy in their child's early development, e.g., Shin & McCoy, 2022; Conger et al., 2010; Bradley & Corwyn, 2002) to our findings, we can start to understand what might be underlying the associations between household resources and the home language environment. Given that households with higher asset scores have more conversational speech and child vocalizations, this relation could be influenced by more resourced houses having the time and energy to have more conversations with children. On the other hand, while the number of adult words is not significantly associated to household resources, this may indicate that household resources are more closely related to caregivers' capacity to interact with children (through conversations), rather than related to their capacity of saying more adult words in the home. While the investment model does help us better understand what underlying relations may be at play in the home language environment, more research is needed to draw further conclusions, especially in underrepresented samples such as our own.

Similar to previous studies (Ma et al., 2023; Zhang et al., 2023; Weisleder & Fernald, 2013; Greenwood et al., 2011; Hart & Risley, 1995;), our research found strong positive associations between the home language environment and child language development. The differences in MCDI scores across all three home language environment measures between the top-tercile and bottom-tercile participants are all statistically significant ($p < 0.01$). For AWC, CTC, and CVC, children in the top tercile scored higher than children in the bottom tercile, with differences of 24, 30, and 35 MCDI points, respectively. Demonstrating the importance of caregiver-child interaction and child-directed speech in early childhood development, this study indicates that CTC and CVC have a stronger connection with child language development than AWC. However, overall, all home language environment measures and child language development measures are strongly associated.

4.5. Strengths and limitations

We acknowledge several strengths of this study. First, this paper is one of the first few studies to quantitatively measure the home language environment using LENA technology in a non-western, developing setting. Second, and importantly, this study is the first to quantitatively examine the home language environment and child language development of rural migrants and urbanized farmers from a peri-urban community, one of the fastest growing sub-populations in China. Although our peri-urban sample shows great heterogeneity in both home language environment and child language development measures, the sample children's average rate of developmental delays is more severe than that of children in both rural and urban regions of China (Ma et al., 2021; Zhang et al., 2015). As China is one of the world's most unequal countries, with huge socioeconomic gaps between rural and urban demographics, this study further reveals how children from low-SES communities are disadvantaged beginning in early childhood (Xie & Zhou, 2014).

This research has several limitations. First, using LENA recorders, we recorded only two days of audio for each sample family, compared to the weekly or biweekly recording of other studies (Suskind et al., 2016; Zhang et al., 2015). However, before recording, we asked all caregivers whether the next two days would be representative of their daily home life. If not, the caregivers were instructed to wait until they had two days of stable, daily life so that the LENA recorders could capture sample children's most typical at-home experiences. Additionally, our meticulous four-step process standardized the 16-hour recordings into 12 hours of audio data, and sample families' results of AWC, CTC, and CVC were reported as the average of two days of recording. Second, the small sample size makes comparability to larger studies more difficult. However, based on our robustness checks, we feel confident in the power and validity of our data and results. Future studies of similar sample sizes could incorporate exploratory moderation analysis. Third, the results in this study are correlational, not causal, and thus must be interpreted without assuming causality. Although we identify associations between the home language environment and language abilities, as well as demographic characteristics and the language development/home language environment measures, future studies should apply experimental or longitudinal research designs to identify the factors that causally influence early language development in peri-urban China. Fourth, our relatively small sample size restricts the generalizability and implications of our results in the context of peri-urban China. Given that this is one of the first studies on the home language environment of peri-urban Southwestern China, it should serve as a starting point for future research to be conducted in larger samples that are more representative. Last, despite the widespread use of LENA in assessing home language environments, the LENA measurements we use in this study (AWC and CTC) are not entirely accurate, especially CTC (Cristia et al., 2020, 2021). Interpretation of our data on CTC should be conducted with these lower accuracy rates in mind, and future studies should continue validation efforts of LENA measures.

4.6. Contributions

Ultimately, this study shows that there is a significant, pressing need to enhance policies and programs to support ECD for families in low-SES communities in China. Children's early language delays can have significant and long-term negative impacts on their future performance, as early childhood development plays a critical role in one's future human capital development. Using LENA technology to further study how the home language environment, early language development, and early childhood development are related in more low-SES communities in other regions, practitioners and policymakers can develop more targeted interventions to improve children's language development. Specifically, policymakers should invest more in parenting trainings—particularly community-based approaches—that teach caregivers how to engage in interactive, conversational-driven, and stimulating parenting practices, because adult-child conversational turns and child vocalizations are significantly associated with child language skills. Strengthening parental knowledge of early childhood language development and parental interaction with children will enable children to develop stronger long-term linguistic and cognitive capacities, which can ultimately contribute to poverty reduction in low-SES communities.

Conclusion

Evidence from this study shows that within a community of similarly disadvantaged, low-SES families in peri-urban China, a strong association between the home language environment and child language development exists. Additionally, there are large and substantial variations in the quality of the home language environment in peri-urban families. These findings suggest that there is a great need for interventions targeting improvement of the home language environment for better ECD outcomes, as well as a need for more research on rural migrant and urbanized farmer communities in peri-urban China.

Declaration of Competing Interest

The authors declare that they have no competing interests.

CRediT authorship contribution statement

Tianli Feng: Conceptualization, Resources, Writing – review & editing, Supervision, Project administration. **Xinwu Zhang:** Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Lulu Zhou:** Methodology, Investigation, Data curation, Writing – original draft, Visualization. **Yue Zhang:** Conceptualization, Methodology, Formal analysis, Resources, Data curation, Supervision, Project administration. **Lucy Pappas:** Formal analysis, Writing – review & editing, Visualization, Supervision. **Sarah-Eve Dill:** Resources, Writing – review & editing, Supervision, Project administration. **Scott Rozelle:** Conceptualization, Resources, Writing – review & editing, Supervision, Project administration. **Yue Ma:** Conceptualization, Methodology, Formal analysis, Resources, Data curation, Supervision, Project administration.

Data availability

Data will be made available on request.

Acknowledgments

We would like to thank the local collaborators and survey enumerators for their assistance with contacting local families and assisting with the survey and LENA data collection, and the children and families who participated in this study.

Funding

This study was supported by private gifts from individual donors. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix File A1. LENA data management protocol

Unfortunately, not all parents always follow the recording protocol as instructed, even when our enumerators confirm with households that all recordings have been correctly produced. We adopted the following criteria to determine whether the data we received from the parents was valid or invalid. We divided each recording session data into five-minute segments. In our dataset, there are 28,798 such segments. A segment is defined to be valid if it satisfies four conditions, as listed below.

1 The segment must be complete, meaning that the recording lasted precisely 300 seconds. About 1.35% of our recording segments were not complete, and we dropped them from our final recording dataset.
2 The segment does not have either of two recording errors:

a. The first recording error arises when the audio file does not have enough child speech.
b. The second when the audio file does not have enough overall speech.

(In our study, if the recordings had either of the above errors, we asked the caregivers to record again. Thus, the final recording segments did not have these two recording errors.)

1 The recording segment took place between 6:20 AM and 11:55 PM. The objective of imposing this restriction is to improve comparability across families as children differ when they go to bed and wake up. Additionally, most children wake up after 6:20 AM, and go to bed before 11:55 PM. We instructed families not to start the recording session until the child was awake and removed and turned off the device when they went to bed.
2 We required that the recording session lasted at least nine hours. Four families did not provide a valid file with at least ten hours of recording. The average recording duration was 14.72 hours, with a standard deviation of 1.80 hours.

In addition, to adjust for skewing commonly seen with count data and variations in the recording start time among families (Cunha et al., 2020), our study used following process to standardize the 16-hour recordings into 12-hours of audio data.

Step 1: We subtracted the start time of each five-minute audio segment from 6:20 AM to generate the number of segments between start time of each five-minute audio and 6:20 AM (for example, the number of segments between 6:35 AM and 6:20 AM is 3 segments).

Step 2: We produced Chebyshev polynomials on the segments between the start time of each five-minute audio segment and 6:20 AM to minimize the Runge phenomenon and provide the best consistent approximation of polynomials in continuous functions (Mason & Handscomb, 2002).

Step 3: We used least absolute shrinkage and selection operator (adaptive LASSO) regression models to select the final Chebyshev polynomials model used in transforming the data (Neto, 2020).

Step 4: We conducted Ordinary Least Square to predict residuals controlling for the segments between the start times of each five-minute audio and 6:20 AM to adjust for variations in the recording start time among families, by using the original value of each five-minute segment's AWC, CTC and CVC as the dependent variable and the final Chebyshev polynomials as the independent variable.

Step 5: We added residuals to constant term to replace original value of five-minutes' AWC, CTC and CVC (no matter missing or not).

Step 6: The fitted AWC, CTC, and CVC were totals from the first 144 valid five-minute segments (usable 12 hours) of each one-day recording.

References

Mason JC, Handscomb DC. Chebyshev polynomials. Chapman and Hall/CRC; 2002.

Neto D. Adaptive LASSO for selecting Fourier coefficients in a functional smooth time-varying cointegrating regression: An application to the Feldstein–Horioka puzzle. Math Comput Simul. 2021;179: 253–264. doi: 10.1016/j.matcom.2020.08.011

References

Chengdushi. (2020). Chenghuaqu guomin jingji he Shehui Fazhan Tongji Gongbao. [2019 People's economic and social development statistical report of Chenghua District, Chengdu municipality]. *Chengdu Bureau of Statistics*. http://www.cdstats.chengdu.gov.cn/htm/detail_181401.html.

Chengdushi. (2020). Jinjiangqu Guomin Jingji he Shehui Fazhan Tongji Gongbao. [2019 People's economic and social development statistical report of Jinjiang District, Chengdu municipality]. *Chengdu Bureau of Statistics*. http://www.cdstats.chengdu.gov.cn/htm/detail_183530.html.

Chengdushi. (2020). Jinniuqu Guomin Jingji he Shehui Fazhan Tongji Gongbao. [2019 People's economic and social development statistical report of Jinniu District, Chengdu municipality]. *Chengdu Bureau of Statistics*. http://www.cdstats.chengdu.gov.cn/htm/detail_181354.html.

Chengdushi. (2020). Qingyangqu Guomin Jingji he Shehui Fazhan Tongji Gongbao. [2019 People's Economic and social development statistical report of Qingyang District, Chengdu municipality]. *Chengdu Bureau of Statistics*. http://www.cdstats.chengdu.gov.cn/htm/detail_183530.html.

Chengdushi. (2020). Wuhouqu guomin jingji he shehui fazhan tongji gongbao [2019 People's economic and social development statistical report of Wuhou District, Chengdu municipality]. *Chengdu Bureau of Statistics*. http://www.cdstats.chengdu.gov.cn/htm/detail_182598.html.

Bai, Y., Zhang, S., Wang, L., Dang, R., Abbey, C., & Rozelle, S. (2019). Past successes and future challenges in rural China's human capital. *Journal of Contemporary China, 28*(120), 883–898. 10.1080/10670564.2019.1594102.

Black, M. M., Walker, S. P., Fernald, L. C. H., Andersen, C. T., DiGirolamo, A. M., Lu, C., McCoy, D. C., Fink, G., Shawar, Y. R., Shiffman, J., Devercelli, A. E., Wodon, Q. T., Vargas-Barón, E., & Grantham-McGregor, S.Lancet Early Childhood Development Series Steering Committee. (2017). Early childhood development coming of age: Science through the life course. Lancet *(London, England), 389*(10064), 77–90. 10.1016/S0140-6736(16)31389-7.

Bland, D., & Hernández, M. (2016). *Urban legend: China's tiered city system explained*. South China Morning Post https://multimedia.scmp.com/2016/cities/.

Bleses, D., Jensen, P., Højen, A., & Dale, P. S. (2018). An educator-administered measure of language development in young children. *Infant Behavior and Development, 52*, 104–113. 10.1016/j.infbeh.2018.06.002.

Bornstein, M. H. (2002). *Handbook of parenting: Children and parenting. In Parenting infants* (pp. 3–43). Lawrence Erlbaum Associates Publishers. https://psycnet.apa.org/record/2002-02629-001.

Bornstein, M. H., Hahn, C.-S., & Haynes, O. M. (2004). Specific and general language performance across early childhood: Stability and gender considerations. *First Language, 24*(3), 267–304.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53*(1), 371–399. 10.1146/annurev.psych.53.100901.135233.

Brooks-Gunn, J., Han, W.-J., & Waldfogel, J. (2010). First-year maternal employment and child development in the first seven years. *Monographs of the Society for Research in Child Development, 75*(2), 7–9. 10.1111/j.1540-5834.2010.00562.x.

Bruner, J. (1985). Child's talk: Learning to use language. *Child Language Teaching and Therapy, 1*(1), 111–114. 10.1177/026565908500100113.

Buckingham, J., Beaman, R., & Wheldall, K. (2014). Why poor children are more likely to become poor readers: The early years. *Educational Review, 66*(4), 428–446. 10.1080/00131911.2013.795129.

Buckingham, W., & Chan, K. W. (2018). One city, two systems: Chengzhongcun in China's urban system. *Journal of Contemporary China, 27*(112), 584–595. 10.1080/10670564.2018.1433576.

Busch, T., Sangen, A., Vanpoucke, F., & van Wieringen, A. (2018). Correlation and agreement between language environment analysis (lena™) and manual transcription for Dutch natural language recordings. *Behavior Research Methods, 50*(5), 1921–1932. 10.3758/s13428-017-0960-0.

Canault, M., Le Normand, M.-T., Foudil, S., Loundon, N., & Thai-Van, H. (2016). Reliability of the language environment analysis system (LENA™) in European French. *Behavior Research Methods, 48*(3), 1109–1124. 10.3758/s13428-015-0634-8.

Casillas, M., Brown, P., & Levinson, S. C. (2020). Early language experience in a Tseltal Mayan village. *Child Development, 91*(5), 1819–1835. 10.1111/cdev.13349.

Casillas, M., Brown, P., & Levinson, S. C. (2021). Early language experience in a Papuan community. *Journal of Child Language, 48*(4), 792–814. 10.1017/S0305000920000549.

Chen, D., Wang, Y., Ren, F., & Du, Q. (2016). Spatio-temporal differentiation of urban-rural equalized development at the county level in Chengdu. *Sustainability, 8*(5) Article 5. 10.3390/su8050422.

Coddington, C. H., Mistry, R. S., & Bailey, A. L. (2014). Socioeconomic status and receptive vocabulary development: Replication of the parental investment model with Chilean preschoolers and their families. *Early Childhood Research Quarterly, 29*(4), 538–549. 10.1016/j.ecresq.2014.06.004.

Conger, R. D., Conger, K. J., & Martin, M. J. (2010). Socioeconomic status, family processes, and individual development. *Journal of Marriage and Family, 72*(3), 685–704. 10.1111/j.1741-3737.2010.00725.x.

Conger, R. D., McCarty, J. A., Yang, R. K., Lahey, B. B., & Kropp, J. P. (1984). Perception of child, child-rearing values, and emotional distress as mediating links between environmental stressors and observed maternal behavior. *Child Development, 55*(6), 2234–2247. 10.2307/1129795.

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. 10.1044/2020_JSLHR-19-00017.

Cristia, A., Lavechin, M., Scaff, C., Soderstrom, M., Rowland, C., Räsänen, O., Bunce, J., & Bergelson, E. (2021). A thorough evaluation of the language environment analysis (LENA) system. *Behavior Research Methods, 53*(2), 467–486. 10.3758/s13428-020-01393-5.

Cunha, F., Gerdes, M., & Nihtianova, S. (2020). Language environment and maternal expectations: An evaluation of the LENA start program. *Economics of Education*, 68.

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica, 78*(3), 3 Article. 10.3982/ECTA6551.

d'Apice, K., Latham, R. M., & von Stumm, S. (2019). A naturalistic home observational approach to children's language, cognition, and behavior. *Developmental Psychology, 55*(7), 1414–1427. https://doi.org/10.1037/dev0000733

Dahl, G., & Lochner, L. (2005). *The impact of family income on child achievement*. Mass., USA: National Bureau of Economic Research Cambridge.

Doyle, L. W., Davis, P. G., Schmidt, B., & Anderson, P. J. (2012). Cognitive outcome at 24 months is more predictive than at 18 months for IQ at 8-9 years in extremely low birth weight children. *Early Human Development, 88*(2), 95–98. 10.1016/j.earlhumdev.2011.07.013.

Drummond, S., O'Driscoll, M. P., Brough, P., Kalliath, T., Siu, O.-L., Timms, C., Riley, D., Sit, C., & Lo, D (2017). The relationship of social support with well-being outcomes via work–family conflict: Moderating effects of gender, dependants and nationality. *Human Relations, 70*(5), 544–565. 10.1177/0018726716662696.

Duncan, G. J., Lee, K. T. H., Rosales-Rueda, M., & Kalil, A. (2018). Maternal age and child development. *Demography, 55*(6), 2229–2255. 10.1007/s13524-018-0730-3.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*(2), 143–188. 10.1017/S0272263102002024.

Emmers, D., Jiang, Q., Xue, H., Zhang, Y., Zhang, Y., Zhao, Y., Liu, B., Dill, S.-E., Qian, Y., Warrinnier, N., Johnstone, H., Cai, J., Wang, X., Wang, L., Luo, R., Li, G., Xu, J., Liu, M., Huang, Y., & Rozelle, S. (2020). Early childhood development and parental training interventions in rural China: A systematic review and meta-analysis. *BMJ Global Health, 6*(8), 13. 10.1136/bmjgh-2021-005578.

Eriksson, M., & Berglund, E. (1999). Swedish early communicative development inventories: Words and gestures. *First Language, 19*(55), 55–90. 10.1177/014272379901905503.

Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg, S., Marjanovič-Umek, L., Gayraud, F., Kovacevic, M., & Gallego, C. (2012). Differences between girls and boys in emerging language skills: Evidence from 10 language communities. *British Journal of Developmental Psychology, 30*(2), 326–343. 10.1111/j.2044-835X.2011.02042.x.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates communicative development inventories* (2nd ed.). Paul H. Brookes.

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science, 16*(2), 234–248. 10.1111/desc.12019.

Ford, M., Baer, C., Xu, D., Yapanel, U., & Gray, S. (2008). The LENA(TM) language environment analysis system: Audio specifications of the DLP-0121 (LENA(TM) Technical Report LTR-03-2). *LENA Foundation*. http://www.lena.org/wp-content/uploads/2016/07/LTR-03-2_Audio_Specifications.pdf.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2021). *Variability and consistency in early language learning: The wordbank project*. MIT Press https://langcog.github.io/wordbank-book/.

Galsworthy, M. J., Dionne, G., Dale, P. S., & Plomin, R. (2000). Sex differences in early verbal and non-verbal cognitive development. *Developmental Science, 3*(2), 206–215. 10.1111/1467-7687.00114.

Ganek, H., & Eriks-Brophy, A. (2018). Language environment analysis (LENA) system investigation of day long recordings in children: A literature review. *Journal of Communication Disorders, 72*, 77–85. 10.1016/j.jcomdis.2017.12.005.

Ganek, H. V., & Eriks-Brophy, A. (2018). A concise protocol for the validation of language environment analysis (LENA) conversational turn counts in Vietnamese. *Communication Disorders Quarterly, 39*(2), 371–380. 10.1177/1525740117705094.

Gilkerson, J., & Richards, J. A. (2009). *Impact of adult talk, conversational turns, and TV during the critical 0-4 years of child development*. LENA Foundation LENA Foundation Technical Report 2nd Edition; p. 36 http://www.lena.org/wp-content/uploads/2016/07/LTR-01-2_PowerOfTalk.pdf .

Gilkerson, J., Richards, J. A., Warren, S. F., Oller, D. K., Russo, R., & Vohr, B. (2018). Language experience in the second year of life predicts language outcomes in late childhood. *Pediatrics, 142*(4). 10.1542/peds.2017-4276.

Gilkerson, J., Zhang, Y., Xu, D., Richards, J. A., Xu, X., Jiang, F., Harnsberger, J., & Topping, K. (2015). Evaluating language environment analysis system performance for Chinese: A pilot study in Shanghai. *Journal of Speech, Language, and Hearing Research, 58*(2), 445–452. 10.1044/2015_JSLHR-L-14-0014.

Goodrich, J. M., Lonigan, C. J., Phillips, B. M., Farver, J. M., & Wilson, K. D. (2021). Influences of the home language and literacy environment on Spanish and English vocabulary growth among dual language learners. Early Childhood Research Quarterly*, 57*, 27–39. 10.1016/j.ecresq.2021.05.002.

Gottfried, A. W. (2013). *Home environment and early cognitive development: Longitudinal research*. Academic Press.

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*(2), 83–92. 10.1177/1525740110367826.

Hao, M., Shu, H., Xing, A., & Li, P. (2008). Early vocabulary inventory for Mandarin Chinese. *Behavior Research Methods, 40*(3), 728–733. 10.3758/brm.40.3.728.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.

He, S., & Wang, K. (2016). China's new generation migrant workers' urban experience and well-being. In D. Wang, & S. He (Eds.), *Mobility, sociability and well-being of urban living* (pp. 67–91). Berlin Heidelberg: Springer. 10.1007/978-3-662-48184-4_4.

Heath, S. B. (1983). *Ways with words: Language, life and work in communities and classrooms*. Cambridge University Press.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782), 1900–1902. https://doi.org/10.bbgfv3

Heckman, J. J. (2008). The case for investing in disadvantaged young children. *CESifo DICE Report, 06*(2), 3–8.

Heckman, J. J., & Mosso, S. (2014). The economics of human development and social mobility (Working Paper No. 19925). *National Bureau of Economic Research*. 10.3386/w19925.

Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review, 103*(6), 2052–2086. 10.1257/aer.103.6.2052.

Hesketh, T., Jun, Y. X., Lu, L., & Mei, W. H. (2008). Health status and access to health care of migrant workers in China. *Public Health Reports, 123*(2), 189–197. 10.1177/003335490812300211.

Heurlin, C. (2019). Unemployment among land-losing farmers in China: Evidence from the 2010 census. *Journal of Contemporary China, 28*(117), 434–452. 10.1080/10670564.2018.1542223.

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., Yust, P. K. S., & Suma, K. (2015a). The contribution of early communication quality to low-income children's language success. *Psychological Science, 26*(7), 1071–1083. 10.1177/0956797615581493.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*(5), 1368–1378. 10.1111/1467-8624.00612.

Hu, X. (2020). Ruhe caisuan shidi nongmin? [How to define land-lost farmers?]. *FindLaw.cn*. https://china.findlaw.cn/ask/question_jx_56514.html.

Hurtado, N., Grüter, T., Marchman, V. A., & Fernald, A. (2014). Relative language exposure, processing efficiency and vocabulary in Spanish–English bilingual toddlers. *Bilingualism: Language and Cognition, 17*(1), 189–202. 10.1017/S136672891300014X.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science, 11*(6), F31–F39. 10.1111/j.1467-7687.2008.00768.x.

Johnson, K., Caskey, M., Rand, K., Tucker, R., & Vohr, B. (2014). Gender differences in adult-infant communication in the first months of life. *Pediatrics, 134*(6), e1603–e1610. 10.1542/peds.2013-4289.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences of the United States of America, 103*(27), 10155–10162. 10.1073/pnas.0600888103.

Kolenikov, S., & Angeles, G. (2009). Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer? *Review of Income and Wealth, 55*(1), 128–165. 10.1111/j.1475-4991.2008.00309.x.

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (1998). Screening for speech and language delay: A systematic review of the literature. Health Technology Assessment *(Winchester, England), 2*(9), 1–184.

Leaper, C., Anderson, K. J., & Sanders, P. (1998). Moderators of gender effects on parents' talk to their children: A meta-analysis. *Developmental Psychology, 34*(1), 3.

Lee, V. E., & Burkam, D. T. (2002). Inequality at the starting gate: Social background differences in achievement as children begin school. *Economic Policy Institute, 1660 L Street, N*.

Leech, K. A., & Rowe, M. L. (2021). An intervention to increase conversational turns between parents and young children. *Journal of Child Language, 48*(2), 399–412. 10.1017/S0305000920000252.

Leung, C. Y. Y., Hernandez, M. W., & Suskind, D. L. (2020). Enriching home language environment among families from low-SES backgrounds: A randomized controlled trial of a home visiting curriculum. *Early Childhood Research Quarterly, 50*, 24–35. 10.1016/j.ecresq.2018.12.005.

Li, M., & Xiong, Y. (2018). Demolition of Chengzhongcun and social mobility of Migrant youth: A case study in Beijing. *Eurasian Geography and Economics, 59*(2), 204–223. 10.1080/15387216.2018.1503966.

Li, N., & Placier, P. (2015). Migrant and non-migrant families in Chengdu, China: Segregated lives, segregated schools. *Social Sciences, 4*(2) Article 2. 10.3390/socsci4020339.

Li, S. (2010). The economic situation of rural migrant workers in China. *China Perspectives* 2010(2010/4), Article 4. 10.4000/chinaperspectives.5332.

Lopez, L. D., Walle, E. A., Pretzer, G. M., & Warlaumont, A. S. (2020). Adult responses to infant prelinguistic vocalizations are associated with infant vocabulary: A home observation study. *PLOS ONE, 15*(11), Article e0242232. https://doi.org/10.1371/journal.pone.0242232.

Lovas, G. S. (2011). Gender and patterns of language development in mother-toddler and father-toddler dyads. *First Language, 31*(1), 83–108. 10.1177/0142723709359241.

Ma, Y., Jonsson, L., Feng, T., Weisberg, T., Shao, T., Yao, Z., Zhang, D., Dill, S.-E., Guo, Y., Zhang, Y., Friesen, D., & Rozelle, S. (2021). Variations in the home language environment and early language development in rural China. *International Journal of Environmental Research and Public Health, 18*(5) Article 5. 10.3390/ijerph18052671.

Ma, Y., Pappas, L., Zhang, X., Feng, T., Su, W. B., Wang, Q., Zeng, Y., Dill, S.-E., & Rozelle, S. (2023). Family-level factors of early childhood development: Evidence from rural China. *Infant Behavior and Development, 70*, Article 101787. 10.1016/j.infbeh.2022.101787.

Maccoby, E. E., & Jacklin, C. N. (1978). *The psychology of sex differences*: 2. Stanford University Press.

Mason, J. C., & Handscomb, D. C. (2002). *Chebyshev polynomials*. Chapman and Hall/CRC. 10.1201/9781420036114.

McCoy, D. C., Peet, E. D., Ezzati, M., Danaei, G., Black, M. M., Sudfeld, C. R., Fawzi, W., & Fink, G. (2016). Early childhood developmental status in low- and middle-income countries: National, regional, and global prevalence estimates using predictive modeling. *PLOS Medicine, 13*(6), Article e1002034. 10.1371/journal.pmed.1002034.

Meisenberg, G., & Woodley, M. A. (2013). Are cognitive differences between countries diminishing? Evidence from TIMSS and PISA. *Intelligence, 41*(6), 808–816. 10.1016/j.intell.2013.03.009.

Mendive, S., Mascareño Lara, M., Aldoney, D., Pérez, J. C., & Pezoa, J. P. (2020). Home language and literacy environments and early literacy trajectories of low-socioeconomic status chilean children. *Child Development, 91*(6), 2042–2062. 10.1111/cdev.13382.

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. 10.1111/cdev.13239.

National Working Committee on Children and Women (NWCCW), National Bureau of Statistics (NBS), & United Nations Children's Fund (UNICEF). (2018). *Children in China: An Atlas of Social Indicator*. UNICEF https://www.unicef.cn/sites/unicef.org.china/files/2019-04/Atlas%202018%20final%20ENG.pdf.

Neto, D. (2021). Adaptive LASSO for selecting Fourier coefficients in a functional smooth time-varying cointegrating regression: An application to the Feldstein–Horioka puzzle. *Mathematics and Computers in Simulation, 179*(C), 253–264.

Noble, K. G., Engelhardt, L. E., Brito, N. H., Mack, L. J., Nail, E. J., Angal, J., Barr, R., Fifer, W. P., & Elliott, A. J.in collaboration with the PASS Network. (2015). Socioeconomic disparities in neurocognitive development in the first two years of life: SES cognitive disparities during infancy. *Developmental Psychobiology, 57*(5), 535–551. 10.1002/dev.21303.

Pae, S., Yoon, H., Seol, A., Gilkerson, J., Richards, J. A., Ma, L., & Topping, K. (2016). Effects of feedback on parent–child language with infants and toddlers in Korea. *First Language, 36*(6), 549–569. 10.1177/0142723716649273.

Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development, 76*(4), 763–782. 10.1111/j.1467-8624.2005.00876.x.

Ramey, C. T., & Ramey, S. L. (2004). Early learning and school readiness: Can early intervention make a difference? *Merrill-Palmer Quarterly, 50*(4), 471–491. 10.1353/mpq.2004.0034.

Ramírez, N. F., Lytle, S. R., & Kuhl, P. K. (2020). Parent coaching increases conversational turns and advances infant language development. *Proceedings of the National Academy of Sciences, 117*(7), 3484–3491. 10.1073/pnas.1921653117.
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*(6), 880–891. 10.1111/desc.12172.
Romeo, R. R., Leonard, J. A., Grotzinger, H. M., Robinson, S. T., Takada, M. E., Mackey, A. P., … Gabrieli, J. D. E. (2021). Neuroplasticity associated with changes in conversational turn-taking following a family-based intervention. *Developmental Cognitive Neuroscience, 49*, 100967. 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. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science, 29*(5), 700–710. 10.1177/0956797617742725.
Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language, 35*(1), 185–205. 10.1017/s0305000907008343.
Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development, 83*(5), 1762–1774. 10.1111/j.1467-8624.2012.01805.x.
Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. Science *(New York, N.Y.), 323*(5916), 951–953. 10.1126/science.1167025.
Scheele, A. F., Leseman, P. P. M., & Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics, 31*(1), 117–140. 10.1017/S0142716409990191.
Shen, Z., Fang, X., & Zheng, X. (2022). The impact of women's off-farm employment on depressive symptoms: Evidence from rural China. *Social Science & Medicine, 311*, Article 115309. 10.1016/j.socscimed.2022.115309.
Shin, S. Y., & McCoy, D. C. (2022). Pathways between family SES, parent characteristics, early experiences, and child language outcomes in South Korea: A combined analysis of the family stress model and the family investment model. *Developmental Psychology, 58*(11), 2081. 10.1037/dev0001408.
Simonsen, H. G., Kristoffersen, K. E., Bleses, D., Wehberg, S., & Jørgensen, R. N. (2014). The Norwegian communicative development inventories: Reliability, main developmental trends and gender differences. *First Language, 34*(1), 3–23. 10.1177/0142723713510997.
Suskind, D. L., Leffel, K. R., Graf, E., Hernandez, M. W., Gunderson, E. A., Sapolich, S. G., Suskind, E., Leininger, L., Goldin-Meadow, S., & Levine, S. C. (2016). A parent-directed language intervention for children of low socioeconomic status: A randomized controlled pilot study. *Journal of Child Language, 43*(2), 366–406. 10.1017/S0305000915000033.
Tamis-LeMonda, C. S., Shannon, J. D., Cabrera, N. J., & Lamb, M. E. (2004). Fathers and mothers at play with their 2- and 3-year-olds: Contributions to language and cognitive development. *Child Development, 75*(6), 1806–1820. 10.1111/j.1467-8624.2004.00818.x.
Tang, C., & van Heuven, V. J. (2009). Mutual intelligibility of Chinese dialects experimentally tested. *Lingua, 119*(5), 709–732. 10.1016/j.lingua.2008.10.001.
Tardif, T., Fletcher, P., Liang, W., Zhang, Z., Kaciroti, N., & Marchman, V. A. (2008). Baby's first 10 words. *Developmental Psychology, 44*(4), 929–938. 10.1037/0012-1649.44.4.929.
Tong, W., Zhu, L., & Lo, K. (2019). Livelihood adaptation and life satisfaction among land-lost farmers: Critiquing China's urbanisation-driven land appropriation. Bulletin of Geography. *Socio-Economic Series, 46*(46), 149–161. 10.2478/bog-2019-0040.
Wake, M., Levickis, P., Tobin, S., Zens, N., Law, J., Gold, L., Ukoumunne, O. C., Goldfeld, S., Le, H. N., Skeat, J., & Reilly, S. (2012). Improving outcomes of preschool language delay in the community: Protocol for the language for learning randomised controlled trial. *BMC Pediatrics*, 12(1), 96. https://doi.org/10.gbb55g
Uccelli, P., Demir-Lira, Ö. E., Rowe, M. L., Levine, S., & Goldin-Meadow, S. (2019). Children's Early Decontextualized Talk Predicts Academic Language Proficiency in Midadolescence. *Child Development, 90*(5), 1650–1663. https://doi.org/10.1111/cdev.13034.
Wang, L., Liang, W., Zhang, S., Jonsson, L., Li, M., Yu, C., Sun, Y., Ma, Q., Bai, Y., Abbey, C., Luo, R., Yue, A., & Rozelle, S. (2019). Are infant/toddler developmental delays a problem across rural China? *Journal of Comparative Economics, 47*(2), 458–469. 10.1016/j.jce.2019.02.003.
Wang, Y. (2020). 2019 Chengdushi Longquanyiqu Guomin Jingji he Shehui Fazhan Tongji Gongbao. [2019 People's economic and social development statistical report of Longquanyi District, Chengdu municipality]. *Chengdu Bureau of Statistics*. http://www.cdstats.chengdu.gov.cn/htm/detail_186382.html.
Wang, Y., Hartman, M., Aziz, N. A., Arora, S., Shi, L., & Tunison, E. (2017). A systematic review of the use of LENA technology. *American Annals of the Deaf, 162*(3), 295–311. 10.1353/aad.2017.0028.
Wang, Y., Li, W., Xiong, J., Li, Y., & Wu, H. (2019). Effect of land expropriation on land-lost farmers' health: Empirical evidence from Rural China. *International Journal of Environmental Research and Public Health, 16*(16), 2934. 10.3390/ijerph16162934.
Weber, A., Fernald, A., & Diop, Y. (2017). When cultural norms discourage talking to babies: Effectiveness of a parenting program in rural Senegal. *Child Development, 88*(5), 1513–1526. 10.1111/cdev.12882.
Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science, 24*(11), 2143–2152. 10.1177/0956797613488145.
Wong, D. F. K., & Song, H. X. (2008). The Resilience of migrant workers in Shanghai China: The roles of migration stress and meaning of migration. *International Journal of Social Psychiatry, 54*(2), 131–143. 10.1177/0020764007083877.
Xie, Y., & Zhou, X. (2014). Income inequality in today's China. *Proceedings of the National Academy of Sciences, 111*(19), 6928–6933. 10.1073/pnas.1403158111.
Xu, D., Yapanel, U., & Gray, S. (2009). Reliability of the LENA(TM) language environment analysis system in young children's natural home environment (LENA Foundation Technical Report LTR-05-02; p. 16). *LENA Foundation*.
Yue, A., Wu, M., Shi, Y., Luo, R., Wang, B., Kenny, K., & Rozelle, S. (2017). The relationship between maternal parenting knowledge and infant development outcomes: evidence from rural China. *REAP Working Papers*. https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/317-_the_relationship_between_maternal_parenting_knowledge_and_infant_development_outcomes.pdf.
Zhan, S. (2011). What determines migrant workers' life chances in contemporary China? Hukou, social exclusion, and the market. *Modern China, 37*(3), 243–285. 10.1177/0097700410379482.
Zhang, W. (2007). Alternation of [n] and [l] in Sichuan dialect, Standard Mandarin and English: A single-case study. *Leeds Working in Linguistics and Phonetics, 12*, 156–173.
Zhang, X., Ma, Y., Feng, T., Zhang, V., Wu, X., Li, M., … Rozelle, S. (2023). The home language environment and early language ability in rural Southwestern China. *Frontiers in Psychology, 13*, 1010442. https://doi.org/10.3389/fpsyg.2022.1010442.
Zhang, Y., Xu, X., Jiang, F., Gilkerson, J., Xu, D., Richards, J. A., Harnsberger, J., & Topping, K. J. (2015). Effects of quantitative linguistic feedback to caregivers of young children: A pilot study in China. *Communication Disorders Quarterly, 37*(1), 16–24. 10.1177/1525740115575771.
Zhu, Y. (2017). In situ urbanization in China: Processes, contributing factors, and policy implications. *China Population and Development Studies, 1*(1), 45–66. 10.1007/BF03500917.