It takes a village: caregiver diversity and language contingency in the UK and rural Gambia

Abstract

Introduction. There is substantial diversity within and between contexts globally in caregiving practices and family composition, which may have implications for the early interaction's infants engage in. We draw on data from the [blinded] project, which longitudinally examined infants in the UK and in rural Gambia, West Africa. In The Gambia, households are commonly characterized by multigenerational, frequently polygamous family structures, which, in part, is reflected in the diversity of caregivers a child spends time with. In this paper, we aim to 1) evaluate and validate the Language Environment Analysis (LENA) for use in the Mandinka speaking families in The Gambia, 2) examine the nature (i.e., prevalence of turn taking) and amount (i.e., adult and child vocalizations) of conversation that infants are exposed to from 12-24 months of age and 3) investigate the link between caregiver diversity and child language outcomes, examining the mediating role of contingent turn taking.

Method. We obtained naturalistic seven-hour-long LENA recordings at 12, 18 and 24 months of age from a cohort of N=204 infants from Mandinka speaking households in The Gambia and N=61 infants in the UK. We examined developmental changes and site differences in LENA counts of adult word counts (AWC), contingent turn taking (CTT) and child vocalizations (CVC). In the larger and more heterogenous Gambian sample, we also investigated caregiver predictors of turn taking frequency. We hereby examined the number of caregivers present over the recording day and the consistency of caregivers across two subsequent days per age point. We controlled for children's cognitive development via the Mullen Scales of Early Learning (MSEL).

Results. Our LENA validation showed high internal consistency between the human coders and automated LENA outputs (Cronbach's alpha's all >.8). All LENA counts were higher in the UK compared to the Gambian cohort. In The Gambia, controlling for overall neurodevelopment via the MSEL, CTT at 12 and 18 months predicted CVC at 18 and 24 months. Caregiver consistency was associated with CTT counts at 18 and 24 months. The number of caregivers and CTT counts showed an inverted u-shape relationship at 18 and 24 months, with an intermediate number of caregivers being associated with the highest CTT frequencies. Mediation analyses showed a partial mediation by number of caregivers and CTT and 24-month CVC.

Discussion. The LENA provided reliable estimates for the Mandinka language in the home recording context. We showed that turn taking is associated with subsequent child vocalizations and explored contextual caregiving factors contributing to turn taking in the Gambian cohort.

Keywords: contingent turn taking, caregiving, language development, diversity

1 Introduction

Contingent, responsive caregiver-child interactions (e.g., interactions characterized by caregiver sensitivity, and reciprocal interactions) have long been highlighted as important for promoting early learning and shaping children's developmental outcomes (e.g., Bornstein & Manian, 2013; Dunst & Kassow, 2008). While children's language development is characterized by marked individual differences (Cristia et al., 2014; Hoff, 2006), distinct genetic (Mountford & Newbury, 2019) and environmental factors (e.g., quantity and quality of verbal input) have been shown to affect children's language development (McGillion et al., 2017; Hirsh-Pasek et al., 2015; Hoff, 2006; Rowe, 2008). Indeed, the quantity and quality (e.g., contingency of verbal input) of caregiver-child-directed speech have an impact far beyond child language development and have been found in numerous studies to be associated with both cognitive and socio-emotional development more broadly (Huttenlocher et al., 1991; Huttenlocher et al., 1998; Weisleder & Fernald, 2013). An abundance of evidence suggests that parent-child interactions, and specifically the nature of early verbal interactions during the first two years of life, play a crucial role in shaping children's developmental outcomes. However, evidence is, as per the majority of developmental research, strongly biased in favor of studies conducted in high-income settings (HICs, Draper et al., 2022).

1.1 Language acquisition and socioeconomic status: evidence from HICs and LMICs

Evidence from HICs shows a clear link between socioeconomic status (SES) and language development (Hart & Risley, 1995; Hoff, 2003; Hoff-Ginsberg, 1991; Pace et al., 2017; Schwab & Lew-Williams, 2016). In their landmark study, Hart and Risley (1995) estimated that by the age of three, US-children from higher-SES backgrounds had heard 30 million more words (including 144,000 more encouraging and 84,000 fewer prohibitive words) than children from lower-SES backgrounds. As reviewed in [blinded] this study has been critiqued on methodological grounds, however recent large-scale studies provide convergent evidence, finding SES to mediate the link between maternal verbal diversity and subsequent child language (Daneri et al., 2019). A growing body of evidence also reports that low SES is frequently associated with structural and functional differences in language neural systems compared with other neurocognitive domains (see Farah, 2017 for a review). Structurally, lower SES is associated with reduced grey matter in areas surrounding the left lateral sulcus, which underlies phonological, semantic, and syntactic components of language comprehension and production (Noble et al., 2015; Noble, Houston, Kan & Sowell, 2012). Furthermore, grey matter reductions have been observed in bilateral occipitotemporal regions involved in reading (Merz, Maskus, Melvin, He, & Noble 2019). Functionally, SES differences have been linked to activation in left inferior frontal, superior temporal and fusiform regions during language and reading tasks (Noble, Wolmetz, Ochs, Farah, & McCandliss, 2006; Raizada et al., 2008, Farah, 2017). While a link between SES on child language development in HICs is welldocumented, recent findings from low- and middle-income countries (LMICs) provide a more nuanced picture: a recent meta-analysis by Sania and colleagues (2019) highlights how environmental pressures often prevalent in LMICs impact child language development, as well as overall neurodevelopmental attainment. Their meta-analysis demonstrated that both parental factors (e.g., maternal and paternal education, maternal height) and child factors (e.g., birth weight, gestational age at delivery, episodes of anemia across childhood), as well as infrastructural risk factors (e.g., limited access to sanitation) were associated with children's language outcomes. Such, frequently poverty-related, risk factors however do not occur in isolation, and may be mediated by favorable resilience-building factors found in a child's psychosocial environment. For example, Prado and colleagues (2017) found that an SES gradient in child language development in samples in Ghana, Malawi, and Burkina Faso was consistently mediated by maternal caregiving. While a connection between poverty-associated risk factors and language development in LMICs has been established, far less is known about the broader contextual factors (both risk and protective) at play. As caregiving practices vary widely by culture, many questions remain with regard to how different care models affect 1) the amount and nature of early language input and 2) subsequent language trajectories of the child. This is particularly relevant given that the mechanisms that drive infants' language learning might be more varied across LMIC settings. Cassilas et al. (2019) report that at 3.63 minutes per hour, infants from a Tseltal Mayan community were exposed to lower levels of infant directed speech than were found in previous studies in the US, Canada and the Netherlands. While their findings were in line with other research in Mayan communities (Shneidman & Goldin-Meadow, 2012), both studies reported that key milestones of language acquisition were still met. Since the majority of the infant directed speech reported in these studies occurred during conversational turn taking, bidirectional turn taking may deserve special attention in addition to unidirectional adult-infant directed speech. As suggested by Rowe & Weisleder (2020), further in-depth, cross-cultural investigations aimed at better understanding the mechanisms of this association therefore may provide relevant insights as to which aspects of caregiving may foster bi-directional turn taking. Building on research from HICs highlighting the immense importance of infant-directed speech, such studies also hold the potential to understand the caregiving arrangements in which infant-directed speech and turn taking are most likely to occur, providing tangible targets for psychosocial interventions.

1.2 Blended family caregiving and implications for turn taking and child language acquisition

Previous investigations of how caregiver language input affects child development are overwhelmingly grounded in nuclear family contexts (e.g., Sear, 2016), typically prevalent in HICs. Emerging cross-cultural findings highlight the universality of the presence of turn taking from an early age: in a study of 684 mothers and their 5.5 month-old infants across Argentina, Belgium, Brazil, Cameroon, France, Israel, Italy, Japan, Kenya, South Korea, and the United

States, Bornstein and colleagues (2015) show that while overall rates of mother and infant vocalizations differed starkly across countries and were uncorrelated in terms of their frequency of occurrence, turn taking between mothers and infants was observed in all study sites from early infancy onwards. The question of what moderates turn taking may be answered by a closer look at caregiving practices. Globally, there is substantial diversity in caregiving practices and family composition. The parents may be supported by grandparents (Chung et al., 2020), older siblings (Mapendo et al., 2022) and other family members in childrearing. An exploration of such models and associated caregiving diversity in the context of child language exposure may be beneficial in three ways: first, a larger number of caregivers may be associated with a higher overall quantity of adult word input, which shows strong links with subsequent child language outcomes (Campisi et al., 2009; Caskey et al., 2011; Hart & Risley, 1995; Hoff, 2003; Hoff & Naigles, 2002; Huttenlocher et al., 2010; Rowe, 2008). Developmentally, there is some indication that the predictive utility of quantity and quality of verbal input may vary across different points in time. For example, the total number of words has been found to be particularly important during the second year of life, while the quality (e.g., the diversity and sophistication) of language input increases in importance from around age 2 (Rowe, 2012). However, the importance of adult-child language exposure represents a crucial predictor of children's language development (Romeo et al., 2018). Secondly, a greater number of caregivers may be associated with more diverse language input, which has been found to predict language outcomes over and above the quantity of words heard by the child (Jones & Rowland, 2017). Thirdly, a larger number of caregivers may also lead to a greater frequency of contingent turn taking instances, which have been shown to be positively associated with children's vocabulary growth (Donnelly & Kidd, 2021). The latter of these proposed mechanisms may deserve particular attention, as some studies have already highlighted that the mere quantity of speech *around* the child overall is less predictive than specific instances of infant-directed speech and turn taking. Shneidman and colleagues (2013) showed that toddlers exposed to speech from only one household member compared to several household members did not differ in language outcomes at 3 years of age, and neither did they encounter infant directed speech at a higher frequency. They did however find that the frequency of infant-directed speech, though not increased by the number of adults the child interacted with, was associated with language outcomes in both groups. This and similar studies (e.g., Weisleder & Fernalt, 2013) suggests that one potential mechanism by which the number of caregivers might affect children's developmental outcome is by the meaningful, contingent interactions a child is able to engage in with the multiple caregivers that support them, both their mother and others. Some limited evidence from day-care settings, has examined adult to child ratios in relation to language development. For example, Schaffer & Liddell (1984) report while there was a significant increase in overall words spoken by the adult when moving from a 1:1 to a 1:4 adult: child ratio, the speech directed at any one child dropped to below 1/3. This was also the case in an Italian setting where adult words directed at each child decreased when moving from a 1:1 to a 1:3 and finally a 1:7 ratio (Scopesi, 1990). More recent studies highlight the importance of context: while in the home settings typically a lower number of adults who engaged in a high number of one-to-one interactions was reported, this was only partially the case for day-care and home day-care settings, where dynamics of child-adult interactions may unfold in a more complex way (Soderstrom et al., 2018). Hereby, it was found that the rate of adult words spoken around (but not necessarily directed at) the child gradually decreased as the adult: child ratio increased.

What has received less attention in the literature though is, what happens when the ratio reverses, with more than one adult interacting and providing verbal input to any one child. Furthermore, it is of interest to include measures of child-directed speech and conversational turns, as key drivers of child language outcomes. Hereby, we propose that there may be an intermediate level at which number of caregivers will be associated with greatest developmental gains: namely where caregivers can give more contingent input than could be provided by a single primary caregiver when in the context of a time-demanding day (i.e., where a caregiver might also be occupied by household chores and farming duties) but balanced against the caregiving being too thinly dispersed across too many caregivers to prevent these interactions from developing into responsive, meaningful, and developmentally appropriate interactions. To capture how such interactions play out in day-to-day life, a move away from standardized lab-based measures provides distinct advantages, as they allow measurements of real-life interactions over prolonged periods of time.

1.3 Naturalistic, automated language recordings offer new insight

The relatively recent advent of long-form, automated language recordings, which can be used to record children and their families naturalistically in the home, provides new opportunities for such research, circumventing the need for extensive transcription, or the pragmatic use of short-form recordings (Soderstrom et al., 2012). Language Environment Analysis (LENA, Ford et al., 2008) represents one of the most widely used systems to obtain recordings of both adult word counts (AWC) and child vocalizations counts (CVC), as well as contingent turn taking (CTT) between the target child and a social partner. To date, several studies have tested the accuracy of the LENA system in languages other than English; these include Spanish (Weisleder & Fernald, 2013), French (Canault et al., 2016), Vietnamese (Ganek & Eriks-Brophy, 2018), Mandarin (Gilkerson et al., 2015), Korean (Pae et al., 2016), Swedish (Schwarz et al., 2017) and Dutch (Busch et al., 2017). Most of these studies have restricted analyses to the reliability of AWC and CVC, with a sub-group also validating CTT measures (e.g., Busch et al., 2017; Ganek & Eriks-Brophy, 2018; Gilkerson et al., 2015; Oetting, Hartfield, & Pruitt, 2009). Overall, these studies have reported significant correlations between the LENA system and human coders for adult and child vocalization counts, with correlation coefficients ranging from 0.64 to 0.92 for AWC (e.g., Canault et al., 2016;

Weisleder & Fernald, 2013; Xu et al., 2009) and from 0.71 to 0.77 for CVC (e.g., Canault et al., 2016; Bush et al., 2017). The few reliability studies that have analyzed CTT counts have reported mixed findings from no correlation at all (Oetting et al., 2009), to significant correlations only after removing outliers (Gilkerson et al., 2015; Pae et al., 2016), to low-medium correlations (Busch et al., 2017; Ganek & Eriks-Brophy, 2018). Overall, these correlations ranged from 0.14 to 0.72 (Busch et al., 2017; Ganek & Eriks-Brophy, 2018; Gilkerson et al., 2015). In light of an increasing interest in conversational measures, further work on validating LENA counts and particularly CTT counts is warranted, to expand the utility of this method for a broader population.

1.4 Current study and cohort

The current study draws on data from the [blinded] project which longitudinally followed N=204 infants in The Gambia and N=61 in the United Kingdom (UK) from birth to 2 years of age. For a description of the full protocol please see [blinded]. Globally, The Gambia is one of the lowest ranking countries regarding gross national income and life expectancy. A majority of the rural-living population support themselves through subsistence farming, with mothers often responsible for the majority of the household and farming duties while fathers are employed in farming and other roles. Families commonly live in extended, multi-generational households (Brotherton et al., 2021; Kea, 2013; Sear & Mace, 2009) with childcare being viewed as a shared responsibility among family members. Most frequently, parents are supported by grandmothers and the child's older sisters (Brotherton et al., 2021; Sear & Mace, 2009). Islamic beliefs feed into child-rearing practices, and a great emphasis is placed on religious and community values (Sosseh et al., 2023). In terms of shifts of children's engagement in early education, the introduction of free universal primary schooling over the past 10 years has led to an enrolment of 97% (CEICdata.com, 2018), with increasing availability of preschool education (Blimpo et al., 2015, 2022). Previous findings from our group provide evidence for developmental differences in the Gambian cohort compared to age matched infants from the UK, which may have implications for their language development. On a behavioral level, [blinded] showed that on the Mullen Scales of Early Learning (MSEL), infants from this population showed a decline in their scores relative to norm scores across all domains of development, between 5 and 24 months of age, including indices of expressive and receptive language development. Furthermore, examining neural indices of habituation and novelty detection from 5-18 months of age, we found that while in the UK group-level, novelty detection was observable from 5 months of age in both functional near infrared spectroscopy (fNIRS) measures [blinded] and electroencephalography (EEG) measures [blinded], these only became apparent in the Gambian cohort from around 18 months of age [blinded]. These indices were obtained in context of auditory tasks, assessing selective attention to either different sound categories (EEG task) or a change in speaker from female to male (fNIRS task). Due to prior literature highlighting the importance to selectively attend to novel speech sounds in the acquisition of language patterns (e.g., Barry et al., 2009; Bishop & McArthur, 2004), the behavioral sequelae of these early neural correlates warrant investigation. These prior findings highlight the need for a more in-depth investigation of 1) naturalistic language trajectories in these cohorts, and 2) an examination of key mechanisms underlying given developmental outcomes. Such a characterization may inform future intervention approaches, by providing reference data on key language acquisition parameters for this population and by examining a potential mechanism that shapes such outcomes, that can be targeted in specific parenting programs.

The aims of this study were threefold: first, since this is the first study to examine Mandinka speaking families, we aimed to analyze the reliability of the LENA system in our Gambian cohort. Secondly, we aimed to examine how LENA measures in both sites changes with age and investigate possible site differences between The Gambian and the UK cohorts. Thirdly, to understand which caregiving factors may promote turn taking in the Gambian cohort, we examined the number of caregivers present during the recording day and caregiver consistency across two subsequent days. We hereby hypothesized that 1) similar as in early neural measures associated with language development we would see lower child vocalization counts in The Gambia compared to the UK, 2) LENA measures of child vocalizations would increase with from 12 to 18 and 24 months in both cohorts and 3) LENA contingent turn taking (CTT) counts at earlier age points (12 and 18 months) would be positively associated with child vocalization counts (CVC) at subsequent age points (18 and 24 months) in both cohorts. Building on these analyses we further hypothesized that 3) caregiver consistency would be associated with CTT counts within age points and that 4) number of caregiver would mediate the association of concurrent CTT and subsequent CVC in the Gambian cohort. For the UK cohort, such detailed analyses were precluded by the much smaller sample size as well as a lack of variance in key measures of interest (i.e., number of caregivers). Data from the UK cohort will therefore be presented for the first three hypotheses only, alongside the more in-depth analyses based on the caregiving context in the Gambian cohort.

2. Methods

2.1 Participants

Families participated in this LENA study as part of the [blinded] project ([blinded]). For the Gambian arm of the study, expectant mothers were recruited antenatally during routine clinic visits at the [blinded] field station of the [blinded]. Families were excluded in cases where 1) the mother was below the age of 18, 2) the infant was born before 37 or after 42 weeks gestation, 3) the infant was diagnosed with any neurological deficits at postnatal checks. Furthermore, all families had to be from the Mandinka ethnic group to participate (meaning Mandinka was spoken the majority of the time in the home). This was done to avoid confounds from having to translate stimulus material of other studies into several languages, and to avoid the need for multiple validations of the LENA recordings. In the West Kiang region of The Gambia where this research took place, the majority of the population belongs to the Mandinka ethnic group (Hennig et al., 2017). Overall, N = 204 families were recruited and eligible at the first antenatal visit. Families were then seen when the infant was 7-14 days, 1 month, 5 months, 8 months, 12 months, 18 months, and 24 months old. LENA recordings were obtained when infants were 12, 18 and 24 months of age. Ethical approval was obtained from [blinded].

In the UK, families were recruited during antenatal clinic visits to the [blinded]. In total, 61 families were recruited and eligible at the first antenatal visit. All families lived either in [blinded] or within a 20-mile radius. Ethical approval was granted by the [blinded].

2.1.1. Sample size justification

We based target sample sizes (The Gambia N=200, UK N=60,) in reference to the main outcome measures of the [blinded] project, namely fNIRS and EEG. These indicated that sample sizes from 20 (moderate effect size) to 42 (small effect size) were sufficient to determine regions of significant cortical brain activation in response to stimuli. Hereby, the Gambian cohort was designed to be larger to allow within-cohort sub-group comparisons and individual differences analyses (for further information please see [blinded].

2.2 Measures

2.2.1 LENA

Procedure. Each participant was recorded for an average of 14 hours over the course of two consecutive days (i.e., 7 continuous hours per day) at home. In The Gambia, two researchers visited the caregivers in their own houses on each recording day to introduce the LENA measure to the families and equip the child with the acquisition device, which was worn by the child above any clothing in a special-made vest. Such a testing approach was required, as households most commonly did not have electricity, meaning that devices needed to be prepared and charged at the field station before being distributed to the surrounding villages. Caregivers were instructed to leave the device inside the vest pocket all day, except for nap and bath time. In those cases, caregivers were asked to keep the recorder as close as possible to the infant. All recordings took place between the hours of 8am and 5pm. At the end of the day, the researchers came back to retrieve the device and administered an interview about the caregivers the child was with across the recording day, and the main activities they participated in. Due to limited literacy of mothers, and limited time-telling using clocks, we opted to record data around different memorable times of the day, in reference to the timing of prayers (i.e., who was with the child before first/second/third prayer etc.). For transcription purposes, we selected audio samples from the second day recording to avoid potential biases error produced by novelty effects during the first day.

In the UK, LENA devices were distributed to families during home visits and parents received information on how to administer and remove the device overnight. Parents would then re-equip the child with the device and mail it back to researchers for analyses.

Validation. For the validation of LENA measures in the Mandinka language, we manually coded the same outputs that are provided from the automated LENA software, specifically AWC, CVC and CTT. For each age point, a subset of ten families, who had successfully completed at least one day-long recording was randomly selected. Based on prior literature, (e.g., Ganek & Eriks-Brophy, 2018; Pae et al., 2016), we selected two noncontinuous 5-minute segments per participant from the second day recording from hours where adult and child word counts were highest (Canault et al., 2016, Gilkerson et al., 2018), resulting in a total of 60 five-minute segments. Sections where the child did not produce vocalizations, for example during naps, were excluded. Each segment was coded by two native Mandinka speakers, who underwent training prior to the coding. During the training and coding, coders had a written set of instructions to follow for each estimate. For AWC, coders listened and transcribed the five-minute segments, including one syllable words (e.g., no), onomatopoeias (e.g., vroom, beep beep) and family-specific word forms (e.g., nyanyang for come and eat). Determinants (e.g., ning moto [this car], woo moto [that car]), prepositions (e.g., tabulo koma [behind the table]), and pronouns (e.g., na moto [my car]; ella moto [your car]) were counted as one word (Busch et al., 2017; Canault et al., 2016). Filling pauses (e.g., uh, um), laughing and crying sounds were not transcribed. Overlapping segments were not transcribed unless the coders could clearly understand the adult words. Coders were instructed to identify as "adults" any speaker whose voice sounded to be 12 years of age or above, due to the emergence of substantial vocal changes early in puberty in both males and females (Killian, 1999; Killian & Wayman, 2010) that make the voice become more adult-like. For CVC and CTT estimates, coders did not produce a transcription of the recording. Instead, in similar manner as the LENA software, coders were instructed to identify and count target-child vocalizations as well as adult vocalizations produced within the 5 seconds before and after a target-child vocalization, similar to prior validation studies (Gilkerson et al., 2008; Gilkerson et al., 2015; Ganek & Eriks-Brophy, 2018). This task was completed using the Audacity software. For the CVC, coders counted all vocalizations produced by the key child including vowel-like sounds (e.g., aaa, eee), one syllable sounds (e.g., ba, ma), one-syllable babble strings (e.g., bababa), several syllables babble strings, squeals (i.e., high pitch vowel-like sounds), and words. Key child vocalisations were counted as one word if they were separated by another vocalisation for <300ms (Oller et al., 2010). Vegetative sounds (e.g., coughing, sneezing) were not coded. Occurrences of overlapping speech and fixed signals (e.g., laughter, crying and screaming) were not included in analyses. Overlapping segments were identified when the key child and other speaker talked at the same time. For CTT, coders counted instances where key child vocalizations were followed by adult words and vice versa. We also counted as turns instances where key child vocalizations were followed by older children and vice versa, due to the high prevalence of the key child being surrounded by multiple other children. Coders were instructed to identify speech within the 5 seconds interval preceding and following target-child vocalizations and then count conversational turns. Sounds such as crying, laughing, and screaming produced by the key child were not counted as part of conversational turns. Turns where one of the speakers was shadowed by someone else's talk were included only if the transcriber was able to understand the words spoken, and if the talk was directed towards the key child.

2.2.2 Mullen Scales of Early Learning

The Mullen Scales of Early Learning (MSEL) are a measure of cognitive ability and motor development using five scales: Gross Motor, Visual Reception, Fine Motor, Expressive Language, and Receptive Language. The assessment was performed at the 5, 8-, 12-, 18-, and 24-month visits of the [blinded] project. Here, we draw on data from the 12- and 18-month age point to control associations between LENA outputs, as they may in part be driven by the general cognitive and language development of the infant. The administration of the MSEL as well as the adaptations made for use in the Gambian context are described in [blinded]. As in our and others prior work, we use both a verbal and non-verbal outcome measure on basis of the MSEL (Delehanty et al., 2018; [blinded]). We obtained average scores for infants' verbal development (verbal MSEL, based on the expressive and receptive language subscales) and non-verbal development (non-verbal MSEL, based on the fine motor and visual reception subscales).

2.3 Statistical analyses

All analyses were based on the second day of recording, to allow participants to get comfortable with the language recorder and obtain the most representative measures. Only families who had at least seven-hour long recordings were retained in analyses, and hourly LENA counts are presented throughout this paper.

For the validation of the LENA in the Mandinka language, we first calculated Cronbach's alpha between automated and manual counts. Next, we examined longitudinal changes in LENA counts per age within each site, via repeated measures analyses of variance (RM-ANOVAs). We then examined site differences via RM-ANOVAs with within factor age (12 months, 18 months, 24 months) and between factor site (Gambia / UK).

Subsequently, we used linear regressions per site to examine links between 1) CTT at 12 months and CVC at 18 months, controlled for verbal and non-verbal MSEL at 12 months, 2) CTT at 12 months and CVC at 24 months, controlled for verbal and non-verbal MSEL at 12 months, and 3) CTT at 18 months and CVC at 24 months, controlled for verbal and non-verbal MSEL at 18 months. Hereby, our analyses were controlled for the child's verbal and non-verbal MSEL at the time of the CTT.

To investigate in more depth the caregiving factors linked with turn taking in the Gambian cohort, we drew on the LENA interview data. Hereby, we obtained indices of caregiver consistency (i.e., the difference in numbers of unique caregivers from the first to the second recording day) and the number of unique caregivers present during the recording day. We then examined linear regressions between CTT and caregiver consistency within each age point. We also carried out linear regression in between number of caregivers and CTT at each age point. Lastly, we conducted mediation analyses to assess whether the number of caregivers mediated the link between LENA CTT and CVC counts longitudinally.

3 Results

Sample sizes for the LENA and MSEL in the Gambian cohort across age points are presented in Figure 1. There was no difference in infants sex ratios at either age point (p all >.25). A majority of parents reported having received no formal education (59.4% of mothers, 55% of fathers), with some reporting limited primary education (12.9 of mothers, 5.9% fathers), complete primary education (3.5% of mothers, 4.7% of fathers), some secondary education (18.2% of mothers, 7.1% of fathers) or complete secondary education (5.9% of mothers, 27.2% of fathers).

Figure 1. Sample sizes for LENA and MSEL per age point for the Gambian cohort and reasons for missing data. [color figure]

Figure 2. Sample sizes for LENA and MSEL per age point for the Gambian cohort and reasons for missing data. [color figure]

Sample sizes for the LENA and MSEL in the UK cohort across age points are presented in Figure 2. The majority of mothers and fathers in this cohort had received undergraduate or postgraduate education and reported working in higher managerial or professional jobs. The median reported household annual income was £60,000 - £79,999. Furthermore, in terms of caregiving arrangements, LENA interview data showed that infants were primarily looked after by their mothers, their fathers, and in few instances accompanied by a sibling, a grandparent and uncle or an aunt. Due to the overall smaller sample size and limited variability in caregiving arrangements, we limited our analyses here to longitudinal changes in the LENA measures and longitudinal associations between LENA indices. Before conducting cross-site comparisons of the LENA indices, we examined our validation results for the use of LENA in the Mandinka language.

3.1 LENA validation in the Mandinka language

Descriptive statistics of the manual counts and the LENA counts and their inter-rater reliability for the 60 transcribed samples can be found in Table 1.

Table 1. LENA indices relative to human coders on CVC, AWC and CTT measures at 12, 18 and 24 months.

	12 months			18 months			24 months		
	Manual x̄ ± SD	LENA x̄ ± SD	Cronbach's α	Manual x̄ ± SD	LENA x̄ ± SD	Cronbach's α	Manual x̄ ± SD	LENA x̄ ± SD	Cronbach's α
Child vocalization count (CVC)	29.10 ± 34.59	35.20 ± 33.95	0.973	31.4 ± 18.79	33.85 ± 15.5	0.873	52.00 ± 47.21	43.05 ± 23.99	0.817
Adult word count (AWC)	251.55 ± 187.24	254.85 ± 133.34	0.897	184.60 ± 179.90	144.05 ± 114.47	0.891	139.05 ± 94.45	176.70 ± 83.32	0.919

Contingent turn taking (CTT)	7.90 ± 4.25	10.25 ± 4.92	0.940	9.05 ± 7.236	7.30 ± 3.08	0.723	12.37 ± 8.69	12.30 ± 6.46	0.898

The reliability of manual codes against the automated LENA measures was found to be in the high range for all measures and age points for CVC and AWC (Cronbach's alpha's all > 0.8, Table 1). For CTT, reliability was also found to be moderate to high (Cronbach's alpha's all >0.7) Bland-Altman plots visualizing the inter-rater reliability can be seen in Figure 3.

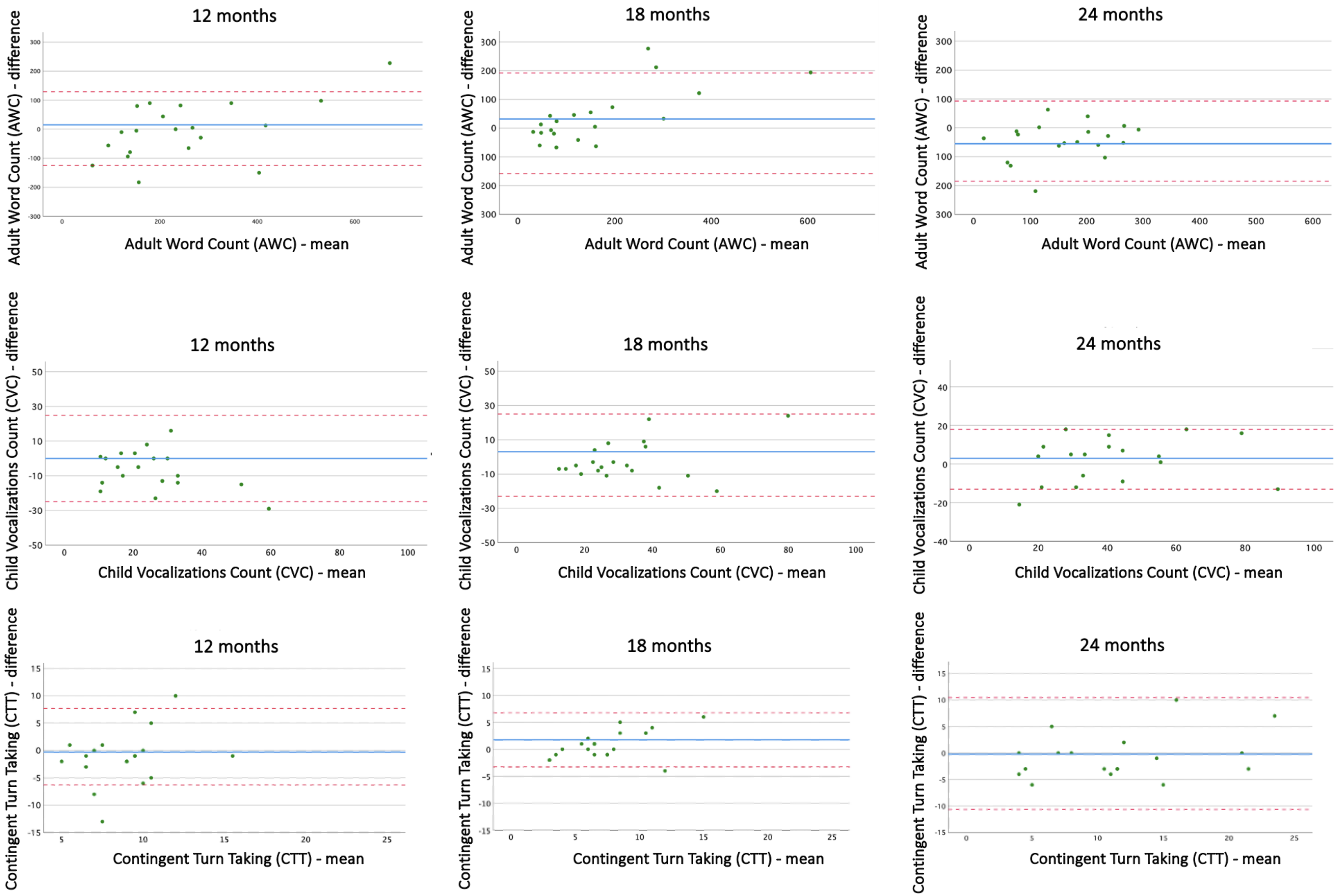

Figure 3. Bland-Altman plots showing averages and difference scores across human coders and automated codes for the Child Vocalization Count (CVC), Adult Word Count (AWC) and Contingent Turn Taking (CTT). These show good agreement between coders for all three counts. [color figure]

3.2 LENA changes with age in the Gambian and the UK cohorts

For the Gambian cohort, changes across the three age points in LENA counts of AWC, CTT and CVC are visualized in Figure 4. Per LENA estimate (i.e., AWC, CTT and CVC), data were entered in a RM-ANOVA with within factor age (12 months, 18 months, 24 months). AWC was found to decline with age ($F_{2,180}$ = 17.354, p <.001). There was no statistically significant change in the CTT count ($F_{2,180}$ = 1.851, p = .160), however the CVC showed a significant increase with age ($F_{2,180}$) = 31.918, p <.001).

For the UK cohort, AWC, CTT and CVC measures are visualized in Figure 5. There was a trend towards an age-related decrease in AWC ($F_{2,66}$ = 3.142, p =.05), and significant increases in both CTT ($F_{2,66}$ = 12.450, p<.001) and CVC ($F_{2,66}$ = 37.852, p<.001) measures.

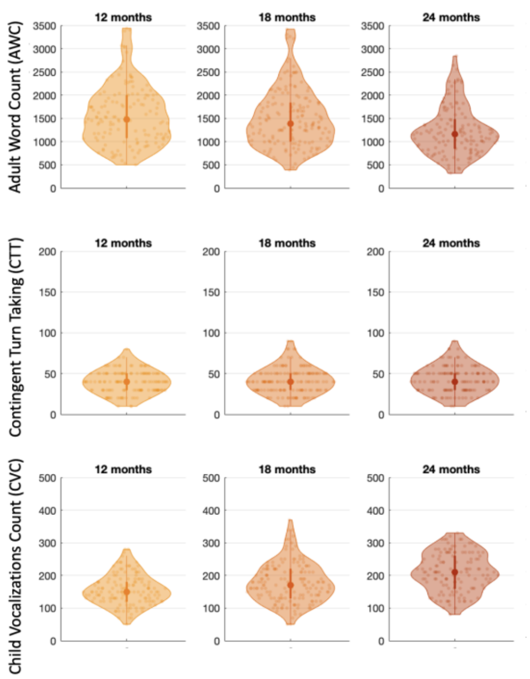

Figure 4. Distributions of Adult Word Count (AWC), Contingent Turn Taking (CTT) and Child Vocalization Count (CVC) at 12 months, 18 months, and 24 months in the Gambian cohort. ACW declined with age (p<.001), CTT showed no age-related change (p=.160) and CVC increased with age (p<.001). [color figure]

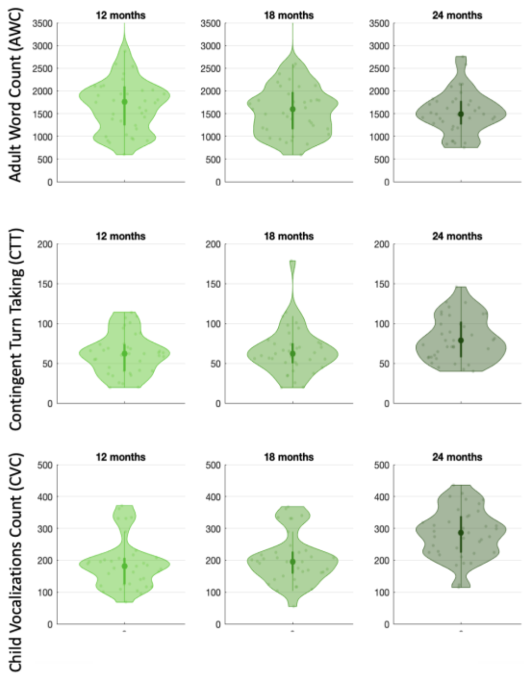

Figure 5. Distributions of Adult Word Count (AWC), Contingent Turn Taking (CTT) and Child Vocalization Count (CVC) at 12 months, 18 months, and 24 months in the UK cohort. There was a trend toward an age-related increase for AWC (p=.05), and significant increases for CTT (p<.001) and CVC (p<.001). [color figure]

3.3 Site differences between the UK and the Gambia in LENA AWC, CTT and CVC

To examine potential site differences in LENA counts and their change over time, we entered LENA measures into RM-ANOVA's with within factors age (12 months, 18 months, 24 months) and between factor site (UK, Gambia). Separate models were run for the AWC, CTT and CVC measures. For AWC, we found main effects for age ($F_{2,246}$=12.40, p <.001) and site ($F_{1,123}$ = 9.373, p=.003), but no age*site interaction ($F_{2,246}$=2.438, p=.121). The site effect was driven by overall higher mean AWC levels in the UK. For CTT, we also found main effects for age ($F_{2,246}$=20.954, p<.001) and site ($F_{1,123}$ = 80.338, p<.001), as well as an age*site interaction ($F_{2,246}$ = 14.275, p<.001). These effects were driven by higher overall CTT mean levels and a higher age-related CTT increase in the UK. For CVC, we found main effects for age ($F_{2,246}$ = 71.527, p<.001) and site ($F_{1,123}$ = 15.977, p<.001), as well as an age*site interaction ($F_{2,246}$ = 11.172, p<.001). As for the CTT measures, these results were due to higher mean levels of CVC and a higher age-related increase in the UK cohort. In sum, these analyses showed higher mean levels of the key LENA indicators in the UK compared to the Gambian cohort, and in case of CTT and CVC stronger age-related change in the UK compared to The Gambia.

3.4 Longitudinal associations of CTT and CVC in the UK and The Gambia

To explore whether CTT was most strongly associated with short-term gains in CVC, or whether CTT at one specific age point was most beneficial for later CVC, we carried out three linear regressions per site to examine links between 1) CTT at 12 months and CVC at 18 months, controlled for verbal and non-verbal MSEL at 12 months, 2) CTT at 12 months and CVC at 24 months, controlled for verbal and non-verbal MSEL at 12 months, and 3) CTT at 18 months and CVC at 24 months, controlled for verbal and non-verbal MSEL at 18 months. All analyses were controlled for the child's language development (via verbal MSEL) and general cognitive development (via non-verbal MSEL) at the time of the CTT measure to ensure increased turn taking was not just a by-product of parents interacting with a more verbose child.

In the Gambian cohort, 12-month CTT was a significant predictor of 18-month CVC (t=3.893, p<.001) and 24-month CVC (t=2.713, p=.008). Further, 18-month CTT predicted 24-month CVC (t=4.148, p<.001). Collectively, these findings highlight links between turn taking in early infancy and subsequent child vocalizations. Partial correlations based on these regression models are visualized in Figure 6.

In the UK, 12-month CTT was a significant predictor of 18-month CVC (t=9.270, p<.001) but not of 24-month CVC (t=1.072, p=.296). However, 18-month CTT was at trend level for 24-month CVC (t=1.778, p=.086). Partial correlations based on these regressions are visualized in Figure 7.

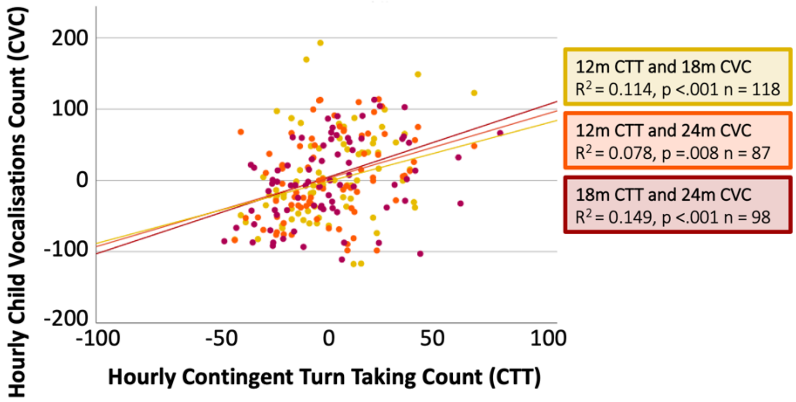

Figure 6. Partial correlations between contingent turn taking count (CTT) at 12 and 18 months and child vocalization count (CVC) at 18 and 24 months for the Gambian cohort. Correlations are controlled for verbal and non-verbal MSEL at the age point of the CTT measure. Moderate to strong longitudinal correlations were evident between CTT and CVC measures. [color figure]

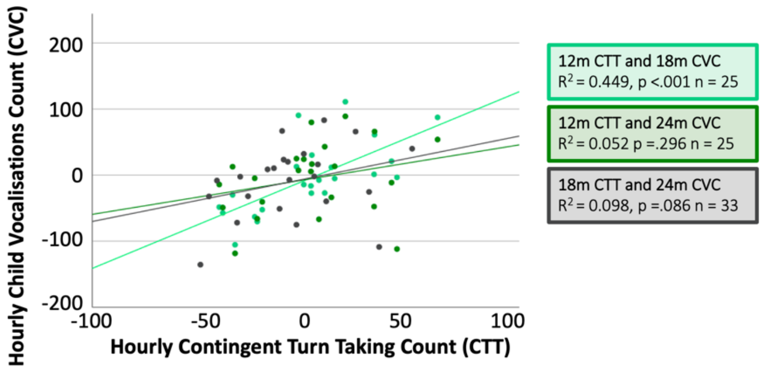

Figure 7. Partial correlations between contingent turn taking count (CTT) at 12 and 18 months and child vocalization count (CVC) at 18 and 24 months for the UK cohort. Correlations are controlled for verbal and non-verbal MSEL at the age point of the CTT measure. A correlation was found for 12-month CTT and 18-month CVC, but not for 12-month CTT and 24-month CVC. A trend was found for 18-month CTT and 24-month CVC. [color figure]

Since CTT was found to be associated with later CVC in the Gambian cohort, we next examined in more depth which caregiving factors in the Gambian context were associated with CTT, where CVC rates were overall lower compared to the UK. Due to a lack of statistical power in the UK cohort, as well as limited variance in the number of caregivers, these analyses were conducted solely for the Gambian cohort.

3.5 Caregiver consistency and infant turn taking

To assess caregiver consistency, we obtained a count of the number of adult caregivers who were reported to have been with the child on each recording day. The number of adult caregivers ranged from 1-21 ($\bar{x}_{12m}$ = 6.2, sd_{12m} =2.89, $\bar{x}_{18m}$ = 4.14 sd_{18m}=2.31, $\bar{x}_{24m}$=3.23 sd_{24m}=1.75). These most commonly included the mother, a co-wife, grandparents and aunts. Less frequently it was reported that fathers, uncles, and babysitters were with the child. We examined the association of caregiver consistency across the three age points, by calculating the absolute value for the difference score between number of unique adults present between the first and the second recording day. We found a linear association between caregiver consistency and CTT at 18 months ($F_{1,156}$ = 6.501, p=.012) and 24 months ($F_{1,130}$ = 9.212, p=.003), but not at 12 months ($F_{1,127}$ = 0.607, p =.437), with higher consistency in caregiver numbers being associated with higher CTT. Results are visualized in Figure 8.

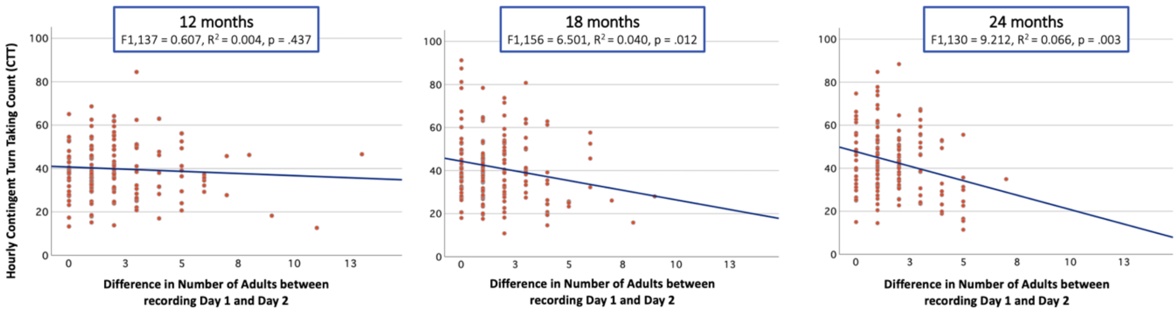

Figure 8. Linear associations between the Difference in Number of Adults between LENA recording Day 1 and Day 2, with Contingent Turn Taking (CTT) measures. Associations were significant at 18 months and 24 months, but not at 12 months. [color figure]

3.6 Number of caregivers and infant turn taking

To probe the relationship between the number of caregivers and children's language exposure, we entered the number of unique adults present during the recording day per age point into separate linear regression models with CTC as the outcome. This showed no significant linear associations at any of the age points. However, significant quadratic associations were found at 18 months ($F_{2,155}$ = 3.330, p=.038) and a trend at 24 months ($F_{2,121}$ = 3.068, p=.050) with those infants looked after by a medium number of caregivers (n = 4-5) showing highest turn taking over the testing day. This association was not apparent at 12 months of age ($F_{2,135}$ = .133, p = .875). As the 12-month association included two outlying values with regard to the number of adults present, we repeated the analysis having removed these two cases, however results remained non-significant ($F_{1,133}$ = .307, p = .736). Results are visualized in Figure 9.

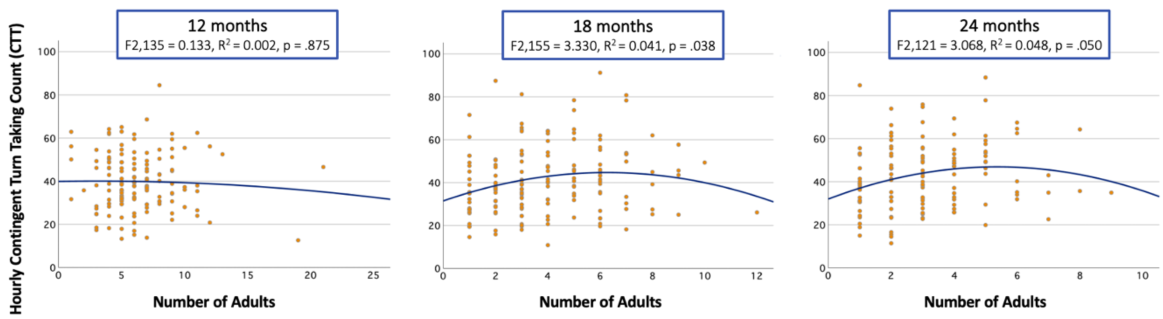

Figure 9. Quadratic associations between the Number of Adults with the child during the recording day and Contingent Turn Taking (CTT) measures. Associations were significant at 18 months, at trend level at 24 months, but not significant at 12 months. [color figure]

3.7 The mediating role of number of caregivers

Building on the above results, we examined whether the number of caregivers mediated the established relationship between contingent turn taking and child vocalizations. Mirroring the regression results above, none of the direct or indirect effects of 12-month CTT, 12-month number of caregivers, and 18-month CVC were significant. However, there was a direct effect of 12-month CTT and 24-month CVC with higher 12-month turn taking predicting higher 24-month CVC (b=1.258, p<.001), which was partially mediated by the number of caregivers present at 12 months (Figure 10).

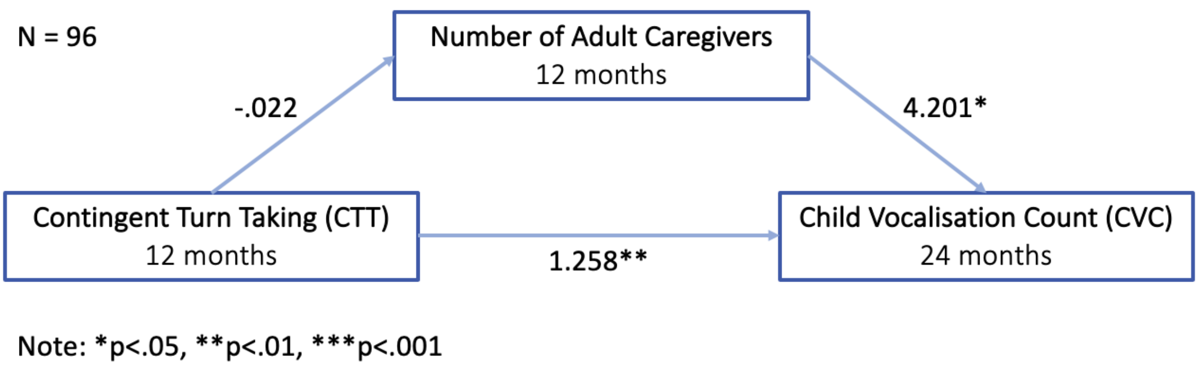

Figure 10. Mediation model showing association between 12-month CTT and 24-month CVC, with Number of Adult Caregivers at 12 months as the mediator. N represents the number of infants with data for all variables included in the model. [color figure]

There also was a direct effect of 18-month CTT and 24-month CVC with higher 18-month turn taking predicting higher 24-month CVC (b=.359, p<.001), which was mediated by number of caregivers present at 12 months (Figure 11).

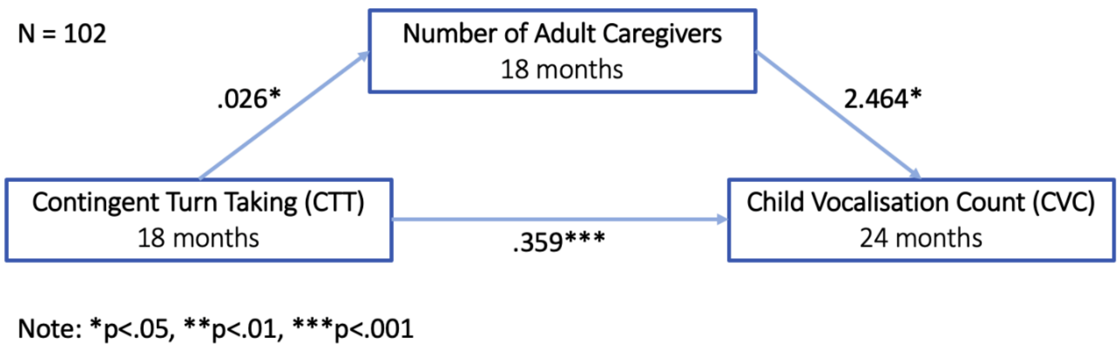

Figure 11. Mediation model showing association between 18-month CTT and 24-month CVC, with Number of Adult Caregivers at 18 months as the mediator. N represents the number of infants with data for all variables included in the model. [color figure]

In sum, these findings highlight the role of caregiver diversity for the relationship of turn taking and subsequent enhanced communication skills.

3.8 Interactions with other children and contingent turn taking

Since caregiving practices vary across countries, and older siblings and other children are involved in childcare in the West Kiang region of The Gambia, we also assessed associations between number of children present during the recording day and CTT. Number of siblings present ranged from 0-7 ($\bar{x}_{12m}$ = 2.23, sd_{12m} =1.82, $\bar{x}_{18m}$ = 2.75 sd_{18m}=1.88, $\bar{x}_{24m}$=2.60 sd_{24m}=2.19), non-siblings ranged from 0-8 ($\bar{x}_{12m}$ = 2.17, sd_{12m} =1.78, $\bar{x}_{18m}$ = 2.05 sd_{18m}=1.78, $\bar{x}_{24m}$=2.51 sd_{24m}=1.91). Due to complexities of the family structure in The Gambia, we examined the number of children in one joint variable, regardless of sibling status. This was done because it is common for a child to be surrounded by both full siblings (sharing a mother and a father) as well as half-siblings (sharing a father or mother only). Regression analyses did not show significant associations between the number of children present during the recording day for turn taking either within age points or longitudinally (p all >.21).

4 Discussion

This study aimed to examine how turn taking is longitudinally associated with children's vocalizations in two diverse cohorts. We obtained longitudinal naturalistic language recordings in infant cohorts in the UK and in rural Gambia across the second year of life. Our study represents the first application of LENA recordings in a Mandinka-speaking population, and, to the best of our knowledge, is also the first application of LENA recordings in sub-Saharan Africa more generally. First, we undertook a validation by comparing LENA indices to the codes of two Mandinka-speaking, Gambian coders. Reliability indices showed a level of agreement that was as good and, in some cases, above that of previous LENA validation studies for other languages.

Having validated the method for Mandinka in this community, we examined age-related changes in LENA measures in both cohorts. We found that in the Gambian cohort AWC declined with age, with a trend in the same direction found in the UK. While no differences were found in CTT in The Gambia, there was an increase across the age points in the UK. Both cohorts showed clear evidence for an increase in CVC across the second year of life. Additionally, we found site-differences in the overall frequencies of the LENA counts for AWC, CTT and CVC, with higher frequencies of all counts in the UK cohort. We also found site*age interaction effects for CTT and CVC, showing higher increases with age in these measures in the UK cohort. These findings indicate that there may be early differences in children's vocalizations across the two cohorts, as well as in the driving factors (namely adult vocalizations and turn taking), that are relevant for long-term language development.

We then examined the relevance of turn taking on the child's subsequent vocalization. We found that CTT was associated with CVC both over a short interval (i.e., from 12 to 18 months and from 18 to 24 months), as well as longer-term (i.e., from 12 to 24 months) in the Gambian cohort. We therefore cannot provide clear evidence supporting the notion that there is one age point at which turn taking is associated with the largest subsequent developmental gains, or whether such gains only become visible over longer follow-up periods. In the UK, only a significant association was found between 12-month turn taking and 18-month child vocalizations. Together with a trending association between 18-month turn taking and 24-month child vocalizations, this may indicate that turn taking provides benefits for immediate to short-term child vocalizations, but not at any specific age point across the second year or life. However, especially with recent, higher-powered studies identifying a link between LENA CVC and CTT (e.g., Donnelly & Kidd, 2021) more consistently in high income settings, we are cautious not to over-interpret our findings from the small sample measured in the UK.

Having identified contingent turn taking as a relevant predictor of child vocalizations in the Gambian cohort, we then examined some characteristics of caregiving that may affect turn taking frequencies within this cohort. We focused on caregiver consistency over the two days around when the LENA recording took place as well as the overall number of caregivers present during the recording day. We found that caregiver consistency was associated with higher rates of turn taking at 18 and 24 months. We also found that the number of caregivers showed an inverted u-shaped association with turn taking, such that an intermediate number of caregivers was associated with highest levels of turn taking. We followed this up by examining the potential mediating role of the number of caregivers for the link of turn taking and child vocalizations. Again, controlling for overall developmental outcomes, we found that the number of adults partially mediated the relationship of 12- and 18-months turn taking and 24-months child vocalizations. This was not the case in the relationship of 12- month turn taking and 18-months vocalizations. However, the direct effect between turn taking and child vocalizations remained significant in all models, highlighting that the association is only in part accounted for by the number of caregivers. Lastly, we found that the presence of other children did not play into the turn taking frequency.

Our findings expand on previous research in several key ways. First, they highlight that while child vocalizations were lower in the Gambian compared to the UK cohort, turn taking may be representing an important mechanism to increase child vocalizations and subsequent language outcomes. In The Gambia, we did not find evidence for any one closely-defined period of time over the second year of life where susceptibility to contingent caregiver-infant interactions is greatest. Our findings suggest that increases in turn taking may have positive implications for children’s vocalizations at any point between the child’s first and second birthday. In context of recent interventions in HICs promoting turn taking, it has become apparent that while short term-gains in turn taking and child vocalizations were measurable, these effects may wash-out over time (McGillion et al., 2017). Therefore, what may be required is more long-term support to families. We also showed that turn taking positively affected later vocalizations, even when controlling for overall developmental status of the child. This highlights that rather than being driven by the verbosity of precocious communication skills of the infant, turn taking has a unique association with longitudinal child vocalization, which in turn provide a valuable training ground for language acquisition. In context of LMICs, recent work also highlights the systemic socio-economic challenges that may affect parent's engagement with their infant (Weber et al., 2021). As has been shown in other contexts where infant mortality is high (e.g., Foley et al., 2021, Scheper-Hughes, 1997), with an infant-mortality rate of 3.9% (Jarde et al., 2021) in The Gambia, delayed or reduced engagement with and attachment to the infant may represent a protective factor for parents' own mental health.

With regard to the number of caregivers and their consistency over time, we found developmental differences in the Gambian cohort: 12-month turn taking was not associated with the consistency of caregiver from one day to the next, whereas such an association was found for the 18- and 24-month age points. At surface level, this may be attributed to the fact that up until their first birthday, mothers still provide the lion's share of the caregiving to the infant, with arrangements becoming increasingly more diverse from 1 year onwards. However, data in the present study do not fully support this, as the number of caregivers was within a similar range across age points. It may also be the case that infants' vocalizations are not yet sufficiently clear to be identified by anyone but the primary caregiver, which may change across the second year of life. More nuanced, in-depth analyses into the specific content of the interaction each caregiver engages the infant in are needed.

4.1 Strengths and limitations

Our findings need to be viewed in context of some strengths and limitations. First, we aimed to explore whether turn taking was most predictive over a shorter longitudinal follow- up period, or whether turn taking at any one specific age point appeared most predictive, linking in with a potential sensitive period in development. This question can however only partially be resolved, due to the limited number of age points studied. Further, our mediation models showed that the number of caregivers showed some associations with turn taking frequencies, which in turn predicted child vocalizations. However, these were only partially mediated, meaning that other factors (e.g., caregiver identity, their role in caregiving, nature and content of interactions) may warrant exploration in follow up analyses. These also may include further analyses on the quality of the caregiver-child interactions, for example by drawing on accompanying data gathered on maternal sensitivity. We further need to contextualize the null findings regarding the link between other children being with the study child and turn taking counts. While no associations were found, LENA recordings are not well-suited to distinguish between different speakers and may have not counted interactions with other children due to similarities in vocal input. Furthermore, we did not systematically record the age of children reported to have been present as part of the LENA interviews, which precluded a more detailed analysis. In this study, we further were only able to present limited analyses on the UK cohort, due to the smaller sample size and lack of variance on reported caregiver numbers. Future work on family structure and caregiving context needs to more systematically capture these factors in order to draw meaningful cross-context comparisons.

5 Conclusion

Contingent turn taking between infants and caregivers represents an important mechanism to shape developmental and learning outcomes in the long-term, and therefore warrants further investigations in previously understudied populations. Our study adds to the knowledge base by highlighting the utility of LENA measures to study infant development in sub-Saharan Africa. Implications of this work provide a foundation to better understand some of the relevant of the caregiving factors as they pertain to turn taking in a this context, highlighting the utility to improve our understanding of how these shape language development on a global scale.

6 Data availability

Data for this article will be shared subject to established data sharing agreements, see [blinded].

7 Competing interest

The authors declare no competing interests.

8 Acknowledgements and funding

We would like to thank all families who participated in this research. [funding information blinded]

References

Barry, J. G., Hardiman, M. J., & Bishop, D. V. (2009). Mismatch response to polysyllabic nonwords: A neurophysiological signature of language learning capacity. PloS one, 4(7), e6270.

Bishop, D. V., & McArthur, G. M. (2004). Immature cortical responses to auditory stimuli in specific language impairment: evidence from ERPs to rapid tone sequences. Developmental Science, 7(4), F11-F18.

[blinded]

Blimpo, M. P., Blimpo, M., Evans, D., & Lahire, N. (2015). Parental human capital and effective school management: evidence from The Gambia. World Bank Policy Research Working Paper, (7238).

Blimpo, M. P., Carneiro, P., Jervis, P., & Pugatch, T. (2022). Improving access and quality in early childhood development programs: Experimental evidence from the Gambia. Economic Development and Cultural Change, 70(4), 1479-1529.

Bornstein, M. H., & Manian, N. (2013). Maternal responsiveness and sensitivity reconsidered: Some is more. Development and psychopathology, 25(4pt1), 957-971.

Bornstein, M. H., Putnick, D. L., Cote, L. R., Haynes, O. M., & Suwalsky, J. T. (2015). Mother-infant contingent vocalizations in 11 countries. Psychological Science, 26(8), 1272-1284.

Brotherton, H., Daly, M., Johm, P., Jarju, B., Schellenberg, J., Penn-Kekana, L., & Lawn, J. E. (2021). “We All Join Hands”: Perceptions of the Kangaroo Method Among Female Relatives of Newborns in The Gambia. Qualitative Health Research, 31(4), 665-676.

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, 1921-1932.

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*, 1109-1124.

Chung, E. O., Hagaman, A., LeMasters, K., Andrabi, N., Baranov, V., Bates, L. M., ... & Maselko, J. (2020). The contribution of grandmother involvement to child growth and development: an observational study in rural Pakistan. *BMJ Global Health*, *5*(8), e002181.

Cristia, A., Seidl, A., Junge, C., Soderstrom, M., & Hagoort, P. (2014). Predicting individual variation in language from infant speech perception measures. Child development, 85(4), 1330-1345.

Daneri, M. P., Blair, C., Kuhn, L. J., FLP Key Investigators, Vernon-Feagans, L., Greenberg, M., ... & Mills-Koonce, R. (2019). Maternal language and child vocabulary mediate relations between socioeconomic status and executive function during early childhood. Child Development, 90(6), 2001-2018.

Donnelly, S., & Kidd, E. (2021). The longitudinal relationship between conversational turn-taking and vocabulary growth in early language development. Child Development, 92(2), 609-625.

Draper, C. E., Barnett, L. M., Cook, C. J., Cuartas, J. A., Howard, S. J., McCoy, D. C., ... & Yousafzai, A. K. (2022). Publishing child development research from around the world: An unfair playing field resulting in most of the world's child population under-represented in research. Infant and Child Development, e2375.

Dunst, C. J., & Kassow, D. Z. (2008). Caregiver sensitivity, contingent social responsiveness, and secure infant attachment. *Journal of Early and Intensive Behavior Intervention*, *5*(1), 40.

Farah, M. J. (2017). The neuroscience of socioeconomic status: Correlates, causes, and consequences. Neuron, 96(1), 56-71.

Foley, S., Hughes, C., Murray, A. L., Baban, A., Fernando, A. D., Madrid, B., ... & Eisner, M. (2021). Prenatal attachment: using measurement invariance to test the validity of comparisons across eight culturally diverse countries. *Archives of women's mental health*, *24*, 619-625.

Ford, M., Baer, C. T., Xu, D., Yapanel, U., & Gray, S. (2008). The lenatm language environment analysis system.

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.

Gilkerson, J., Zhang, Y., Xu, D., Richards, J. A., Xu, X., Jiang, F., ... & 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.

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

Hennig, B. J., Unger, S. A., Dondeh, B. L., Hassan, J., Hawkesworth, S., Jarjou, L., ... & Fulford, A. J. (2017). Cohort profile: the Kiang West Longitudinal Population Study (KWLPS)—a platform for integrated research and health care provision in rural Gambia. *International journal of epidemiology*, *46*(2), e13-e13.

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., ... & Suma, K. (2015). The contribution of early communication quality to low-income children's language success. Psychological science, 26(7), 1071-1083.

Hoff-Ginsberg, E. (1991). Mother-child conversation in different social classes and communicative settings. Child development, 62(4), 782-796.

Hoff, E. (2003). Causes and consequences of SES-related differences in parent-to-child speech.

Hoff, E. (2006). How social contexts support and shape language development. Developmental review, 26(1), 55-88.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. Developmental psychology, 27(2), 236.

Huttenlocher, J., Levine, S., & Vevea, J. (1998). Environmental input and cognitive growth: A study using time-period comparisons. Child Development, 69(4), 1012-1029.

Jarde, A., Mohammed, N. I., Gomez, P., Saine, P. C., D'Alessandro, U., & Roca, A. (2021). Risk factors of infant mortality in rural The Gambia: a retrospective cohort study. *BMJ paediatrics open*, *5*(1).

Jones, G., & Rowland, C. F. (2017). Diversity not quantity in caregiver speech: Using computational modeling to isolate the effects of the quantity and the diversity of the input on vocabulary growth. *Cognitive psychology*, *98*, 1-21.

[blinded]

[blinded]

[blinded]

[blinded]

Mapendo, F., Sando, M., Mosha, D., Yussuf, M., Tinkasimile, A., Lyatuu, I., ... & Bonyck, M. (2022). Older siblings' role in infant and young child feeding and care in rural Tanzania. *Tanzania Journal of Health Research*, *23*.

McGillion, M., Pine, J. M., Herbert, J. S., & Matthews, D. (2017). A randomised controlled trial to test the effect of promoting caregiver contingent talk on language development in infants from diverse socioeconomic status backgrounds. Journal of Child Psychology and Psychiatry, 58(10), 1122-1131.

Merz, E. C., Desai, P. M., Maskus, E. A., Melvin, S. A., Rehman, R., Torres, S. D., ... & Noble, K. G. (2019). Socioeconomic disparities in chronic physiologic stress are associated with brain structure in children. Biological psychiatry, 86(12), 921-929.

[blinded]

Mountford, H. S., & Newbury, D. F. (2019). The genetics of language acquisition. In International handbook of language acquisition (pp. 33-50). Routledge.

Noble, K. G., Houston, S. M., Brito, N. H., Bartsch, H., Kan, E., Kuperman, J. M., ... & Sowell, E. R. (2015). Family income, parental education and brain structure in children and adolescents. Nature neuroscience, 18(5), 773-778.

Noble, K. G., Houston, S. M., Kan, E., & Sowell, E. R. (2012). Neural correlates of socioeconomic status in the developing human brain. Developmental science, 15(4), 516-527.

Noble, K. G., Wolmetz, M. E., Ochs, L. G., Farah, M. J., & McCandliss, B. D. (2006). Brain–behavior relationships in reading acquisition are modulated by socioeconomic factors. Developmental science, 9(6), 642-654.

Oetting, J. B., Hartfield, L. R., & Pruitt, S. L. (2009). Exploring LENA as a tool for researchers and clinicians. *The ASHA Leader*, *14*(6), 20-22.

Oller, D. K., Niyogi, P., Gray, S., Richards, J. A., Gilkerson, J., Xu, D., ... & Warren, S. F. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences*, *107*(30), 13354-13359.

Pace, A., Luo, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2017). Identifying pathways between socioeconomic status and language development. Annual Review of Linguistics, 3, 285-308.

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.

Prado, E. L., Abbeddou, S., Adu-Afarwuah, S., Arimond, M., Ashorn, P., Ashorn, U., ... & Dewey, K. G. (2017). Predictors and pathways of language and motor development in four prospective cohorts of young children in Ghana, Malawi, and Burkina Faso. Journal of Child Psychology and Psychiatry, 58(11), 1264-1275.

Raizada, R. D., Richards, T. L., Meltzoff, A., & Kuhl, P. K. (2008). Socioeconomic status predicts hemispheric specialisation of the left inferior frontal gyrus in young children. Neuroimage, 40(3), 1392-1401.

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological science*, *29*(5), 700-710.

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.

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.

Scheper-Hughes, N. (1997). Demography without numbers. *Anthropological demography: toward a new synthesis*, 201-222.

Schwab, J. F., & Lew-Williams, C. (2016). Language learning, socioeconomic status, and child-directed speech. Wiley Interdisciplinary Reviews: Cognitive Science, 7(4), 264-275.

Schwarz, I. C., Botros, N., Lord, A., Marcusson, A., Tidelius, H., & Marklund, E. (2017). The LENATM system applied to Swedish: Reliability of the Adult Word Count estimate. In *Interspeech 2017, Stockholm, Sweden, 20-24 August, 2017* (pp. 2088-2092). The International Speech Communication Association (ISCA).

Sear, R. (2016). Beyond the nuclear family: an evolutionary perspective on parenting. Current Opinion in Psychology, 7, 98-103.

Soderstrom, M., Casillas, M., Bergelson, E., Rosemberg, C., Alam, F., Warlaumont, A. S., & Bunce, J. (2021). Developing a cross-cultural annotation system and MetaCorpus for studying infants' real world language experience. Collabra: Psychology, 7(1), 23445.

Soderstrom, M., Grauer, E., Dufault, B., & McDivitt, K. (2018). Influences of number of adults and adult: child ratios on the quantity of adult language input across childcare settings. First Language, 38(6), 563-581.

Sosseh, S. A., Barrow, A., & Lu, Z. J. (2023). Cultural beliefs, attitudes and perceptions of lactating mothers on exclusive breastfeeding in The Gambia: an ethnographic study. BMC Women's Health, 23(1), 1-15.

Weber, A. M., Diop, Y., Gillespie, D., Ratsifandrihamanana, L., & Darmstadt, G. L. (2021). Africa is not a museum: the ethics of encouraging new parenting practices in rural communities in low-income and middle-income countries. *BMJ global health*, *6*(7), e006218.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. Psychological science, 24(11), 2143-2152.

Xu, D., Yapanel, U., & Gray, S. (2009). Reliability of the LENA Language Environment Analysis System in young children's natural home environment. Boulder, CO: Lena Foundation, 1-16.