Infancy

WILEY | INFANCY THE OFFICIAL JOURNAL OF THE INTERNATIONAL CONGRESS OF INFANT STUDIES

RESEARCH ARTICLE

Time to Talk: Variability in Caregiver-Child Verbal Engagement During Everyday Activities Sampled From Daylong Recordings

Janet Y. Bang[1,2] | Arlyn Mora[1] | Mónica Munévar[1] | Anne Fernald[1] | Virginia A. Marchman[1]

[1]Department of Psychology, Stanford University, Stanford, California, USA | [2]Department of Child and Adolescent Development, San José State University, San Jose, California, USA

Correspondence: Janet Y. Bang (janet.bang@sjsu.edu)

Received: 23 September 2024 | **Revised:** 17 September 2025 | **Accepted:** 13 October 2025

Handling Editor: Alejandrina Cristia

Keywords: caregiver-child interaction | child-directed speech | daylong recordings | stability

ABSTRACT

Children learn language through interactions with others. To document variation in how caregivers engage verbally with their 2-year-old children, we sampled six 10-min segments of dense child-directed speech from naturalistic daylong audio recordings in an economically- and linguistically-diverse sample of English- ($n = 45$) and Spanish-speaking ($n = 45$) families. Segments of child-directed speech occurred during child-centered (e.g., booksharing, play) and adult-centered (e.g., cooking) activities, with substantial variation among families. During all activity types, child-directed speech was associated with one or more linguistic or interactive features that have been shown to facilitate language development, such as lexical diversity, mean length of utterance (MLU), or responsiveness to children. Moderate to strong stability within families was found, suggesting that caregivers in some families engaged with children more than in others, regardless of the activities in which they participated. Patterns were generally similar across language groups. This study extends previous research by using naturalistic daylong recordings to explore how activity type and caregivers' individual tendencies relate to children's early language experiences.

1 | Introduction

It is well-documented that there is considerable variation in the linguistic, conceptual, and interactional features of the speech that young children hear directed to them (Rowe and Snow 2019). Research on language acquisition has demonstrated links between variation in particular features of child-directed speech and their later language outcomes (Chater and Christiansen 2018; Shneidman and Goldin-Meadow 2012; Weisleder and Fernald 2013). However, previous studies have provided a limited picture of child-directed speech because they have been restricted to primarily English-speaking families or used short recordings with an experimenter present. More studies are needed that document the different ways in which caregivers talk to their children in the context of their everyday lives.

This study used daylong naturalistic audio recordings to sample the contexts in which child-directed speech naturally occurred during the daily activities of 24-month-old children from a diverse sample of families in the U.S. We examined two key sources of variation. First, we examined the types of activities in which caregivers and children engaged, building on prior studies that have explored the linguistic and interactional features of child-directed speech during different everyday activities, such as the number of different words, utterance length, or proportions of responses (Hoff-Ginsberg 1991). How frequently and how long caregivers engage in different types of activities

lead to variation among families in the opportunities they provide for language learning (Rowe and Weisleder 2020). Another source of variation is the caregivers themselves. Prior studies have shown that variation in the quantity and quality of child-directed speech is stable within caregivers, that is, some caregivers tend to talk more, use longer utterances, or be more responsive than others (Bornstein et al. 1999). Evidence for stability suggests that caregivers have different individual tendencies that contribute to variation among families, regardless of the variation in how caregivers and children spend their time. Few studies have explored stability using both linguistic and interactional features of speech in diverse populations. By assessing caregiver talk during naturalistic, everyday activities in a diverse sample of English- and Spanish-speaking families, our goal is to offer a more in-depth understanding of variation among families that is attributable to the activities in which families engage and the tendencies of individual caregivers.

1.1 | Expanding Our View of the Activities in Which Caregivers and Children Engage

Previous studies of caregiver-child interactions have typically examined short audio- or video-recordings while the caregiver and young child engages in structured or semi-structured activities in the home or in the laboratory. Thus, much of the literature has focused on child-directed speech that occurs during *child-centered activities* in which the caregiver and child are engaged in the same activity, such as play and booksharing, or routines, such as getting dressed (Hoff-Ginsberg 1991; Rosemberg et al. 2020, 2023; Tamis-LeMonda et al. 2019). While these types of child-centered activities may be prototypical contexts in which caregivers and children engage, the frequency and duration of these activities varies among families. Child-centered activities can be contrasted with *adult-centered activities* when the caregiver and child are doing different things, such as when the caregiver is folding laundry while the child plays nearby. Adult-centered activities are less well-studied in the literature, but have been noted in a few observational studies of child-directed speech (Dunn et al. 1977; Nelson 1985; Rosemberg et al. 2020, 2023). Moreover, families with toddlers report that these types of activities regularly occur (Cycyk and Hammer 2020). While it is clear that young children are likely to experience a mix of child- and adult-centered activities throughout the day, studies which use short recordings in structured contexts do not capture the natural ebb-and-flow between child-centered and adult-centered activities. Daylong recordings can provide a more naturalistic picture of the frequency of these types of activities in the daily lives of families.

Critically, features of tCDS vary systematically in different activities. In the classic study by Hoff-Ginsberg (1991), English-speaking caregivers and their children from low-to middle-class backgrounds were video-taped during prescribed one-on-one activities, such as eating a meal, playing with toys, and reading books. During toy play, caregivers were more likely to use directive utterances, while meal and dressing activities were more likely to contain conversation-eliciting utterances, such as questions. Multiple studies have shown that caregivers tended to produce more words during booksharing and grooming than during other activities such as mealtimes and play (Hoff-Ginsberg 1991; Tamis-LeMonda et al. 2019). Booksharing has been shown to expose children to more rare words and complex grammar, which are more likely to be found in book texts than in spontaneous speech (Massaro 2015; Montag 2019; Snow et al. 1976; Weizman and Snow 2001). In a more recent study using 4-hour-long audio recordings, Rosemberg et al. (2023) examined caregiver speech in Spanish-speaking families with 14-month-old children in Argentina. Speech during meals was associated with fewer tokens, but also with higher lexical diversity than other activities, consistent with earlier studies by Weizman and Snow (2001). These studies suggest that young children are systematically exposed to different features of language in caregiver-child interactions through their participation in different activities in daily life.

Given that young children engage with caregivers throughout the day, it is important to explore the frequency and duration of the everyday activities that children experience. Differences among families in the relative balance of activity type may have a substantial impact on the nature of children's opportunities for language learning (Montag et al. 2018; Rowe and Weisleder 2020). For example, children who experience relatively more child-centered activities may be exposed to a different experience than those who experience relatively more adult-centered activities. To date, only a few studies with daylong recordings have assessed the frequency of different types of activities and compared features of caregiver verbal engagement among activities (Rosemberg et al. 2020, 2023; Soderstrom and Wittebolle 2013).

In addition, few studies have examined how linguistic and interactional features vary across activities in families who speak languages other than English (cf. Bohn et al. 2024; Rosemberg et al. 2020, 2023). Studies that are restricted to a single language group or that use only convenience samples with mid-to highly-educated families in the U.S. or Europe provide a limited view of when, how, and how much caregivers engage with their young children (Blasi et al. 2022; Kidd and Garcia 2022). In our surrounding community, English is the primary language; however, Spanish is the most frequently used non-English language and approximately half of all non-English speaking children older than 5 years of age speak Spanish at home (U.S. Census Bureau, 2024).

To capture a more comprehensive picture, we conducted new analyses of daylong recording data from two previously-published studies of language processing efficiency that were conducted in parallel (Fernald et al. 2013; Marchman et al. 2020). English-speaking families were recruited in neighborhoods where families lived in less-densely populated areas than is typical of university-based samples. Spanish-speaking families lived in a more densely populated area of a metropolitan suburb, with the majority of the families having immigrated to the U.S. from countries where Spanish was the dominant language, most typically, Mexico. In both groups of families, maternal education levels ranged from less than college (e.g., elementary or some high school) to more than college educated. Our goal is not to compare these families directly, but rather to explore everyday activities with young children in a diverse set of participants, so as to obtain a broader view of the activities

during which children in a U.S. community hear the most speech from their caregivers. Because these families come from different language sub-communities associated with different systemic opportunities for socioeconomic mobility in the U.S., we also explored SES (Dailey and Bergelson 2022) and language group as potential moderators of main effects.

1.2 | Stability in Features of Child-Directed Speech

In addition to variation that derives from the activities in which caregivers and children engage, a second source of variation derives from the caregivers themselves. Some caregivers are likely to talk more and others less, regardless of the activities in which they are engaged. We can estimate stability within caregivers by computing how consistently or inconsistently caregivers, either individually or together within a family, are positioned relative to others on some measure, such as number of word tokens or types, over time or activity contexts (e.g., play time vs. meals) (Bornstein et al. 1999, 2017). By documenting stability in caregivers' behaviors, we can determine whether sources of variation in children's learning environments depend less on activity contexts and more on individual tendencies of caregivers to be more or less talkative.

There is some evidence of caregiver stability across activities in the literature. Bornstein et al. (1999) observed English-speaking caregivers in the U.S. from middle- to upper-class backgrounds during standardized play sessions at two ages. Those mothers who produced more word types when their child was 13 months also did so when their child was 20 months. Those who produced more word tokens and types during a structured home play session with their 13-month-olds were also more likely to do so during other daily routines (Tamis-LeMonda et al. 2017). In more recent work by Demir-Lira et al. (2019), those caregivers who produced more utterances during booksharing also did so during other activities.

Studies based on daylong recordings using automated estimates of adult speech have also found evidence for stability at the level of families. Using the LENA system (Gilkerson and Richards 2020), Gilkerson et al. (2017) reported that adult word counts (AWC) within English-speaking families in the U.S. on consecutive days were strongly inter-correlated. Such correlations based on recordings made farther apart in time were weaker, but still moderate. d'Apice et al. (2019) obtained three daylong LENA recordings within a week in 107 English-speaking families in the U.K. They reported weak stability within families in AWCs across hours within a day; however, correlations were higher when comparing across the three full days. These studies suggested that those children who experience more adult speech on one day are also likely to experience more adult speech on another day. Note that AWCs reflect the frequency of all clear speech spoken around the child by one or more caregivers. Thus, such estimates include speech from all of the adult caregivers that the child engaged with on a particular day, including both speech directed to the target child, that is, tCDS, and speech spoken nearby but not directed to that child, that is, non-tCDS. Relying only on automated estimates from daylong recordings limits the conclusions that can be made about the stability of child-directed speech within families throughout the day. Here, we base our analyses on transcriptions in order to explore evidence for stability in several different features of tCDS and non-tCDS utterances.

1.3 | The Current Study

In this study, we annotated and transcribed speech from the densest hour of the daylong recording for each family, operationalized as those 10-min segments with the highest LENA-generated AWC values judged to be primarily tCDS. We sampled the densest hour of adult speech to capture the features of the language that children hear when they experienced the most speech from adults, regardless of the time of day, what was happening, or who was present (Gilkerson et al. 2017; Rosemberg et al. 2023; Soderstrom and Wittebolle 2013). This sampling strategy allowed us to explore the different activities that were associated with dense periods of child-directed speech and the features of the speech that occurred during these most talkative periods of engagement. However, it does not provide a fully representative picture of the activities that occurred, nor how frequently, how long, or how much caregivers talked during a typical day. Since one or more adults may have engaged verbally with the child at different times in the day, the sampled segments could have captured speech from a single adult or one or more different adults. Typically, these adults were primary/secondary caregivers, but could also have been other adults, such as relatives who participate in the child's daily life. For ease of exposition, we refer broadly to these adults as "caregivers."

For each family, we extracted and transcribed the six 10-min segments in which the most tCDS occurred (total ~ 1 h per family). We first explored how frequently child- versus adult-centered activities occurred during these densest 10-min segments and logged their durations in the segment. This provided a descriptive overview of the relative duration of different activities that occurred when caregivers engaged with their child. We then explored the features of talk during each of those activities, correcting for the duration of the respective activity. Based on transcriptions of all speech, we computed several measures of caregiver speech that have been suggested to promote children's lexical and grammatical development, that is, number of word tokens, word types, MLU, responses, and imitations/expansions (Hirsh-Pasek et al. 2015; Rowe 2012; Weisleder and Fernald 2013; Weizman and Snow 2001). Finally, we conducted correlations among activity types in these measures to assess stability across families.

Our participants were a diverse sample of English- and Spanish-speaking families in the Western U.S, spanning a range of demographic, linguistic, and cultural backgrounds. Given the diversity of our sample, we explored two moderators. We explored language group, given the linguistic differences between groups, and SES, given its proposed relation to caregiver speech (Dailey and Bergelson 2022). Analyses were pre-registered (see Table S5 for minor changes) and study

documentation and manuals can be found on our OSF page (https://osf.io/byjfg/overview). Transcripts are publicly available on the CHILDES repository (https://talkbank.org/childes/access/Eng-NA/FMB_home.html). We asked the following questions:

1. How much of the densest hour of speech is spent engaging in different types of activities? Specifically, how is dense tCDS distributed over booksharing and other types of child- and adult-centered activities?
2. Do features of tCDS known to have positive effects on child outcomes (i.e., word tokens and types, MLU, responses, and imitations/expansions) vary as a function of activity during the densest hour?
3. How stable are features of tCDS and non-tCDS within families, regardless of activity?

2 | Method

2.1 | Participants

Participants were families with 24-month-old children living in a region of the Western U.S. that includes urban, suburban, and less densely populated, rural living contexts (see Fernald et al. 2013; Marchman et al. 2020 for recruitment information). Caregivers provided written informed consent prior to study participation. Studies were conducted according to the guidelines in the Declaration of Helsinki and all procedures were approved by the Institutional Review Board at Stanford University.

Families were screened as speaking either primarily English (n = 45) or Spanish (n = 45) and were both considered monolingual home learning environments. All inclusionary/exclusionary criteria, as well as the methods of data collection, coding, and transcription protocols were identical in both groups regardless of the primary language of the family. All caregivers in the English-speaking families reported that English was the primary home language. A comprehensive language background survey (Marchman and Martínez-Sussmann 2002) was available for a portion of the sample (n = 34), with families reporting an average of 98.94% of exposure to English (SD = 3.84%) across all individuals who interacted regularly with the child (e.g., caregivers, siblings). In the Spanish-speaking families, all caregivers reported that children were exposed predominantly to Spanish from caregivers. The comprehensive language background interview was available for all families, indicating that all children were exposed primarily to Spanish, with families reporting an average of 89.33% of exposure to Spanish (SD = 11.29%).

Demographic and health information was also obtained via parent report. All children were born full term and, at 24 months, were reported to have no vision/hearing loss, developmental delays, or neurodevelopmental concerns. As seen in Table 1, there was a generally equal distribution of boys and girls in both language groups. Ranges of educational (maternal years of education) and socioeconomic backgrounds (Hollingshead Index; 1975) were represented. In the English-speaking families, reported race/ethnicity was 33 non-Hispanic/White, 5 non-Hispanic/Mixed, 5 Hispanic/White, and 2 Hispanic/Mixed. In the Spanish-speaking families, all caregivers reported their ethnicity as Hispanic. Mothers were born in Mexico (n = 38, 84%), Central or South America (n = 4) or the U.S. (n = 3), with years in the U.S. ranging from 5 to 23 years (Mdn = 14 years). All children were born in the U.S.

Data were included for all English-speaking families with available LENA recordings at 24 months, 19 of whom were reported in Fernald et al. (2013). An additional 8 families were excluded because target children were twins (2), the total recording length was < 4 hours (2), the recording occurred only while the child was at daycare (1), or the child wore the LENA incorrectly (3). The 45 Spanish-speaking families in the current analysis were randomly selected from 93 families with complete data to arrive at a sample size equivalent to that of the English-language group and stratified to ensure the full range of available socioeconomic backgrounds. Post-hoc power analyses with G*Power (Faul et al. 2007) determined that our sample size (n = 90) was sufficiently powered (≥ 0.90) to detect medium to large effects.

2.2 | Study Protocol

Daylong audio recordings: Bilingual/bicultural English- or Spanish-speaking RAs asked families to provide a LENA recording (Gilkerson and Richards 2020) on a "typical day" that did not include special events. Children wore a specialized vest to hold the LENA device, but in all other ways, families were encouraged to engage as they normally would. Families were asked to turn on the recorder in the morning and let it turn off after 16 h. Families who requested to record on multiple days were asked to record so that all periods of the day were represented. Caregivers could turn off the recorder at any point or ask to delete the entire recording. All recordings were cleaned to

TABLE 1 | Demographic characteristics of participants from English- and Spanish-speaking families.

	English (n = 45)	Spanish (n = 45)
Child age (months)	24.09 (0.76), 23.0–26.	25.59 (0.62), 24.6–28.1
Child sex (M, F)	20, 22	23, 22
Hollingshead index[a]	40.93 (13.72), 9–64.5	27.83 (11.42), 8–63
Maternal education (years)[b]	14.51 (2.21), 10–18	12.11 (3.15), 6–18

Note: Child Age, Hollingshead Index, and Maternal Education are presented as M (SD), ± range.
[a]The Hollingshead Four Factor Index of Socioeconomic Status (Hollingshead 1975) ranges from 8 to 66.
[b]Number of years of education based on the US: high school graduate = 12 years; college graduate = 16 years; post college degree = 18 years.

exclude segments when the child was sleeping or the recorder was worn incorrectly (see Table S1). Most recordings occurred between 9:00 a.m. to 7:00 p.m. and all hours of the day were represented (see Supporting Information S1: Figure S1).

Selecting dense periods of tCDS: The LENA software provides automated estimates of adult word count (AWC) from adult speech that is "near and clear" to the child for each 5-min recording segment. In some cases, segments were less than 5 min if the recording was turned on/off prior during the segment. To select the six densest 10-min segments of tCDS, the automatically-calculated AWC values were added across every two continuous 5-min segments to better assess continuity over activities. These 10-min segments were then sorted from highest to lowest AWC. Bilingual/bicultural English- or Spanish-speaking RAs classified each 10-min segment, from highest to lowest AWC, as either primarily tCDS, where approximately ≥ 70% of caregiver words were addressed to the target child (either alone or as part of a group) or not. This process filtered out segments where the automated LENA AWC values reflected primarily speech to other individuals. RAs continued until six primarily-tCDS segments were identified per family, yielding ~1 h (M = 59.51 min, Mdn = 60.0, range = 53.55–60.0) of the recording with the densest tCDS. Note that these six 10-min segments were not necessarily contiguous, but were extracted from any part of the recording, regardless of when those segments occurred in the day (see Supporting Information S1: Figure S2). Using the LENA automated counts, AWC totals were computed in the selected six densest tCDS segments per family and AWC/hour values were computed across the full recording.

Activity duration coding: For each selected 10-min segment of the recording per family, English- or Spanish-speaking RAs annotated the start and stop times for all periods of tCDS with ELAN (version 5.0; Wittenburg et al. 2006). Activities were then identified based on utterance content, speech cues such as prosody (e.g., exaggerated intonation, pace, affective tone), and environmental cues (e.g., running water). All other periods of speech were considered non-tCDS. For each period of tCDS, coders identified the main caregiver activity as one of the following: (1) booksharing, (2) other types of child-centered activities (*playing, feeding, routines,* e.g., dressing or diaper changing, or *unstructured conversation*), or (3) adult-centered activities (e.g., caregivers were doing something different than the target child, such as chores while talking with the child). For each family, we computed the total duration of each activity type per segment, resulting in a possible maximum of six observations per activity type per family (see Supporting Information S1: Figure S3 for a visualization of the sampling and coding process).

Reliability: Prior to coding, one 10-min segment per family was selected to be independently double-coded by a second coder who did not have access to the initial codes. Because primary coders listened to all segments in a given family, while secondary coders listened only to one, interrater reliability estimates may be underestimating true reliability. Reliability on the identification (Cohen's κ English = 0.84, Spanish = 0.76) and duration of activities (M ICC: English = 0.74, Spanish = 0.90) was moderate to high (See Table S2 and online manual for more detail).

Transcribed measures of caregiver and child features of talk: Initial transcription of all audible speech in both English and Spanish was conducted by a transcription service employing fluent speakers of the respective language. Bilingual/bicultural RAs working in-house then later proofread and corrected each transcript, attending to speaker roles, utterance breaks, adherence to spelling conventions, and activity start and end times. We first derived three linguistic measures of the speech using FREQ and MLU in CLAN (MacWhinney 2000): *number of word tokens, number of word types* (the number of different lemmatized forms, e.g., "eat" and "eating" were both instances of "eat") and *mean length of utterance in words* (*MLUw*) per speaker. All word tokens/types were included regardless of the language, which was primarily English for the English corpus (99%) and primarily Spanish for the Spanish corpus (95%). Zero values were imputed by CLAN if caregivers had spoken at any point during the 10-min segment but were not speaking during the respective activity. Word tokens and types were normalized to a rate/min to compare activities with different durations. When non-tCDS utterances occurred in the background of an activity that was considered to be primarily child-directed speech, these utterances were transcribed, but not included in the analyses for word tokens and types. MLUw was computed in words, rather than morphemes, given the complexities of calculating morphemes in Spanish and to be consistent across languages (Gutiérrez-Clellen et al. 2000). Because MLUw was computed by utterances, all tCDS and non-tCDS utterances were included in the computation of MLUw. If no utterances occurred for a given speaker, MLU was treated as missing. Note that absolute differences in word token and type counts and MLUw are not directly comparable across languages because of variation in the morphosyntactic structure of English and Spanish (e.g., "Give it to me" is four words in English but one word in Spanish, "dámelo").

To capture the interactional features of speech, we used CHIP to compute the number of tCDS utterances that were *responses* to a target child's utterance, that is, that followed a child's vocalization (e.g., speech or laughter) within five utterances (Sokolov and MacWhinney 1990). CHIP also provides an automated output which sub-categorizes responses as *imitations/expansions,* the number of caregiver utterances containing at least one lemma in common with the prior child utterance. Frequencies were converted to proportions, correcting for the total number of child utterances during each activity. If a child did not vocalize, a proportion could not be calculated. Values greater than one indicated that adults were likely to respond, on average, more than once to each child utterance, while values less than one indicated that adults were likely to respond, on average, less than once to each child utterance.

3 | Results

3.1 | General Descriptives in Amount of Talk

Using the automated output from LENA, Figure 1 plots the sum of the AWCs in the six densest segments (correcting for total duration in the few cases when six segments did not total 60 min), ranked from lowest to highest. This figure illustrates the substantial variability in how much total speech 2-year-old

children experienced during the 6 densest 10 min segments (~1 h) in both groups of families. Analyses confirmed that these dense AWC values were significantly higher, on average, than AWC/hour averaged over the full recordings (English: t (44) = −7.93, $p < 0.001$, $d = 1.12$; Spanish: t (44) = −10.38, $p < 0.001$, $d = 1.53$). Nevertheless, AWC totals during the selected segments were strongly correlated with AWC/hour from the full recording (Full sample: $r = 0.73$, 95% CI [0.62, 0.82], English $r = 0.74$, 95% CI [0.57, 0.85], Spanish $r = 0.79$, 95% CI [0.64, 0.88]) (See Supporting Information S1: Figure S3). This suggests that the individual differences in amount of speech that we observed in our analyzed sample were generally representative of what we would have observed had we analyzed the entire recording.

3.2 | What Are Caregivers Doing When They Talk the Most to Their Child?

Table 2 describes the frequency and duration of the activities that occurred during the six densest 10-min segments in English- and Spanish-speaking families. Within 10-min segments, we found that multiple tCDS activities were often present ($M = 2.07$, SD = 0.86, range 1–5 activities). To provide more detailed descriptive statistics, we present data for booksharing and each of the other types of child-centered activities separately. Booksharing occurred in only about half of the families in each group. In contrast, all families in both groups had at least one instance of each of the other types of child-centered activities, adult-centered activities, and periods of non-tCDS.

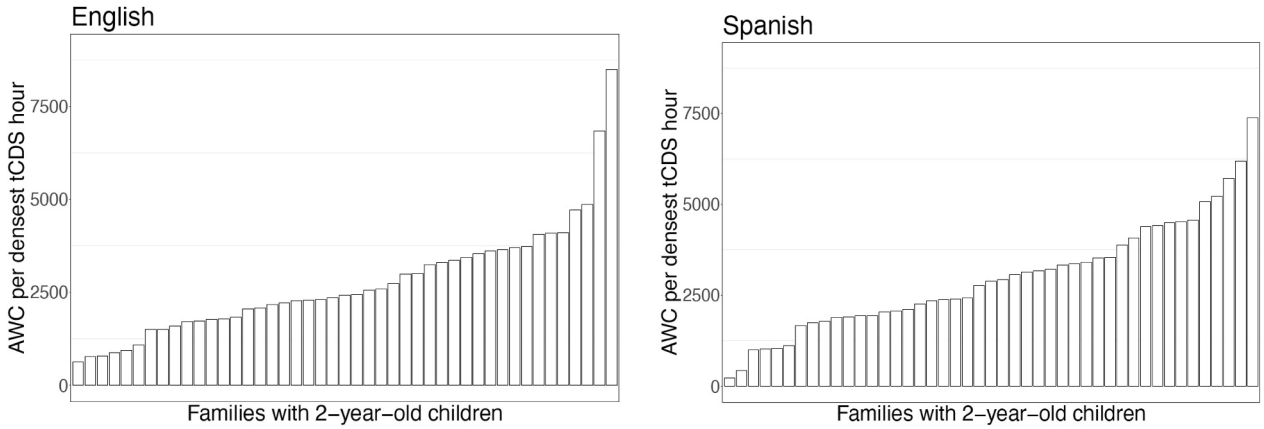

FIGURE 1 | Adult Word Counts (AWC) across the six densest tCDS segments, converted to a rate/hour, in English- ($n = 45$) and Spanish-speaking ($n = 45$) families, sorted from lowest to highest.

TABLE 2 | Number of families and duration per tCDS activity and non-tCDS time in English- and Spanish-speaking families.

	English		Spanish	
Activity	**Families *n***	**Duration (min) M (SD) ± range**	**Families *n***	**Duration (min) M (SD) ± range**
tCDS: Books	22	11.45 (9.20)	20	12.16 (10.36)
		2.09–39.92		1.15–45.31
tCDS: Other child-centered	45	24.21 (11.45)	45	23.80 (11.91)
		1.38–51.27		0.50–47.55
Play	39	10.51 (7.99)	37	8.60 (7.82)
		1.10–33.71		0.21–27.70
Feeding	31	5.79 (5.33)	31	7.32 (5.67)
		0.06–20.18		0.50–20.78
Routines	32	4.84 (5.07)	35	4.91 (3.88)
		0.04–25.64		0.15–14.15
Unst. Conv.	43	8.03 (6.03)	43	8.23 (6.92)
		0.11–24.32		0.16–24.99
tCDS: Adult-centered	45	7.80 (6.65)	45	11.15 (7.60)
		0.01–25.96		1.04–30.58
Non-tCDS	45	21.85 (9.81)	45	19.19 (10.35)
		5.28–42.71		0.36–43.87

Note: Durations are summed across segments, based on those families who contributed to each activity. Zeros were not imputed and therefore, the averages do not sum to 60 min.

Table 2 provides the time spent in each activity type in families who contributed to that activity type, derived by summing across all six segments to get a total duration by activity type. Total time per activity ranged widely from < 1 min to > 45 min across families. Although, we were unable to consistently identify specific speakers from the audio recording alone, we could note the range of speaker roles in both English- and Spanish-speaking families, including mothers, fathers, grandmothers, grandfathers, godmothers, aunts, uncles, and other unidentified adult males or adult females. About one-quarter to one-half of all activities included more than one adult speaker (34% of all activities; books = 20%, routines = 25%, play = 27%, food = 36%, unstructured conversation = 38%, adult-centered activities = 41%, non-tCDS = 52%).

To depict the relative distribution of each activity duration, Figure 2 illustrates the proportion of the total hour by activity type, with zeros imputed for those families where the activity was not present. On average, tCDS occurred between 6% and 15% of the time for each of the child-centered activities, with book-sharing comprising 9% of the time on average in both groups of families. Adult-centered activities, when the adult was talking to the child while engaged in another activity occurred, on average, 13% of the time in English-speaking families and 19% in Spanish-speaking families. Notably, even in these segments classified as having AWC values that reflected primarily child-directed speech, over 30% of the time, on average, caregivers were engaging verbally with other individuals (i.e., non-tCDS). Thus, tCDS was spread out over a variety of different activities and frequently occurred with non-tCDS speech to other children or adults. The amount of time spent during non-tCDS activities was highly variable across families. Note the striking similarity of these distributions across language groups.

Figure 3 illustrates that, in both groups, dense tCDS segments occurred at all hours of the day, even for activities which could typically be associated with bedtime, such as booksharing. Thus, families varied both in the types of activities in which they engaged and in when those activities occurred across the day.

3.3 | Do Features of Talk Vary by Activity Type in Everyday Speech?

We conducted hierarchical linear mixed models for each of the five speech measures using the lme4 package (Bates et al. 2015), summarized in Figure 4. Only families who engaged in the respective activity were included in the models. Family and segment were treated as random effects on the intercept for all models except for MLUw (Matuschek et al. 2017). For MLUw, we only included a random intercept of family; including segment resulted in a singular fit, suggesting that the model was overfit. We first explored models including a factor of activity type (*book-sharing*, *other child-centered activities*, *adult-centered activities*, *non-tCDS*), with child word tokens/hour as a covariate. Child tokens/hour was significant in all models (Fs > 11.16, ps < 0.001), except proportion imitations/expansions (F = 2.10, p = 0.15), indicating that child speech was related to most caregiver measures. Before interpreting any differences by activity type across groups, we tested if language group or SES were potential moderators of activity type. SES did not increase overall model fits as a moderator in any model (χ^2s < 3.80, ps > 0.15). In contrast, language group showed several significant moderation effects. Thus, we present models for all measures, with language group as a fixed effect and its interaction with activity type. We used the emmeans package (Lenth 2025) to calculate estimated marginal means (EMMs) (e.g., adjusted for covariates, fixed effects, random effects) and to conduct post-hoc pairwise comparisons for final models. EMMs and standard errors (SE) are provided in Table 3. See Table S3 and Table S4 for descriptive statistics and Supporting Information S1: Figure S5 for boxplots with raw data. Multiple

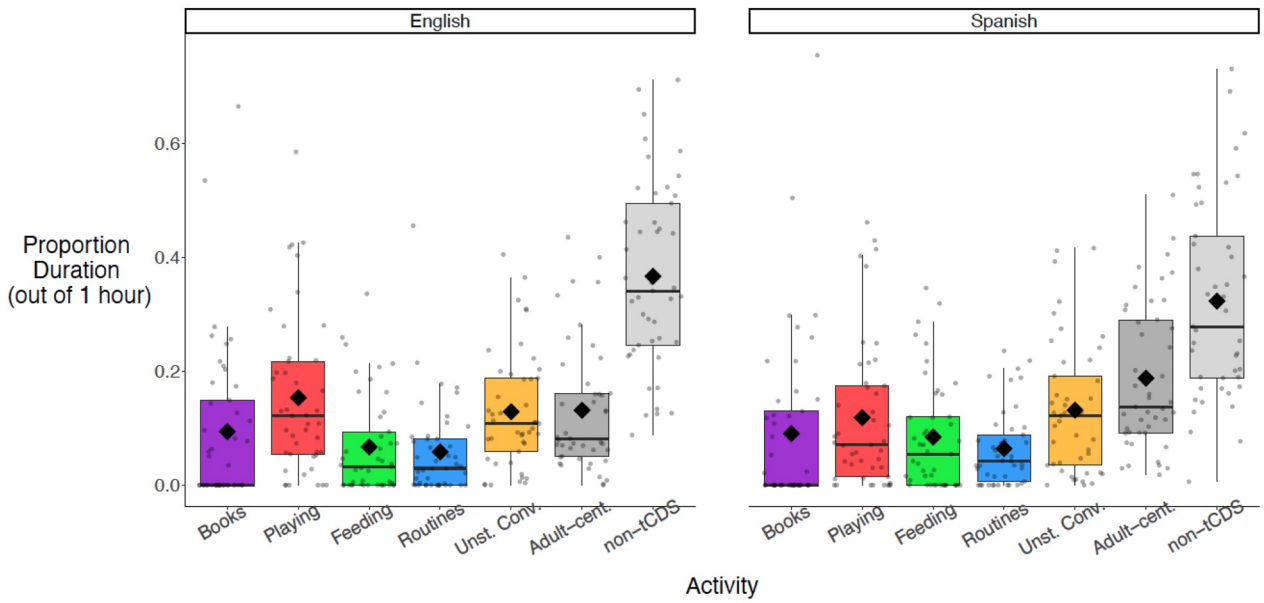

FIGURE 2 | Minutes per activity as a proportion of 1 h in English- (n = 45) and Spanish- (n = 45) speaking families. Zeros were imputed in families when those activities did not occur in order to better evaluate proportion of the total hour across all participants. Horizontal lines reflect median; diamonds represent mean; top and bottom of boxes reflect interquartile range. Dots reflect individual participants.

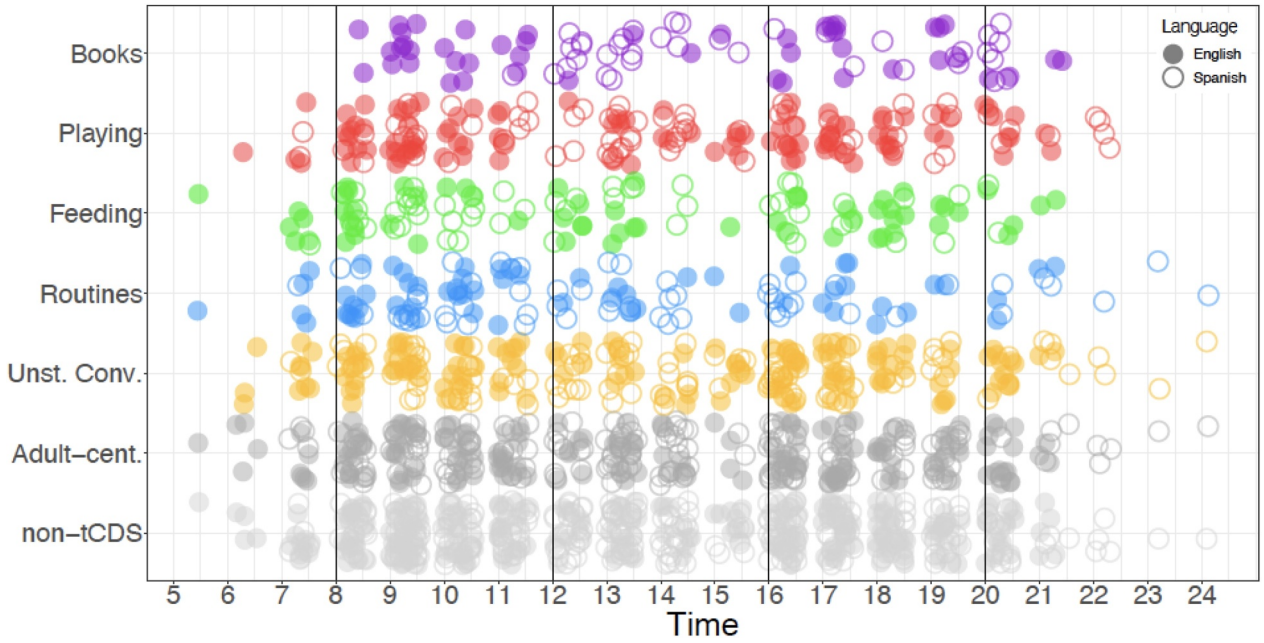

FIGURE 3 | Time of day by activity type for English-speaking ($n = 45$, solid circles) and Spanish-speaking ($n = 45$, open circles) families. Each dot represents one of the six audio segments per family. Dots are jittered to more easily view the distribution of all observations. See Table S1 for more information regarding the number of days for recording.

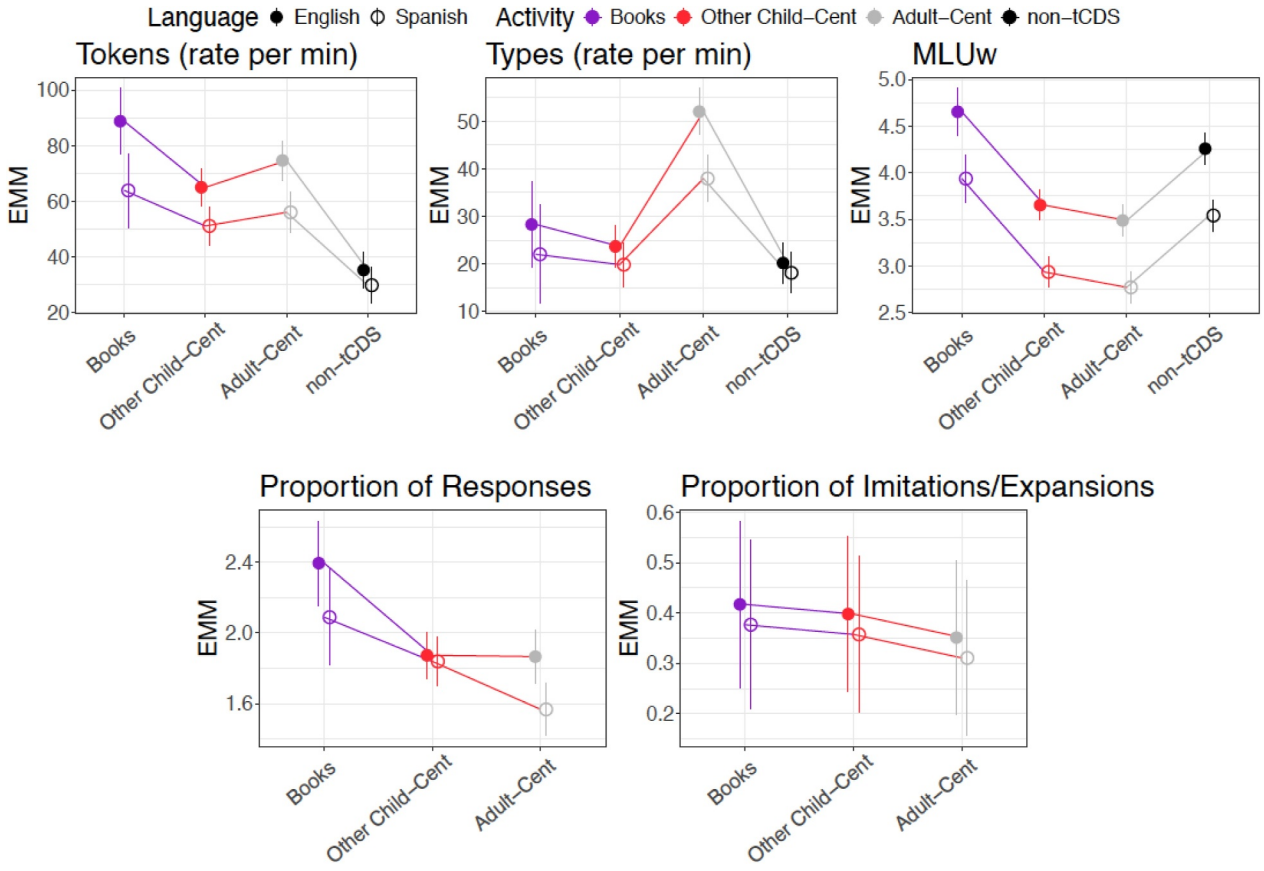

FIGURE 4 | Estimated marginal means (EMMs) by activity type based on final linear mixed effects models, controlling for children's total word tokens, for English- ($n = 45$, solid circles) and Spanish-speaking ($n = 45$, open circles) families. Error bars represent the 95% confidence interval.

TABLE 3 | Estimated marginal means (EMM) and standard errors (SE) per type of activity in English- (n = 45) and Spanish-speaking (n = 45) families.

Measure EMM (SE)					
Activity type English	**Tokens/min**	**Types/min**	**MLUw**	**Proportion responses**	**Proportion Imit/Exp**
tCDS					
Child-centered books	88.8 (6.01)	28.3 (4.64)	4.66 (0.13)	2.40 (0.12)	0.42 (0.08)
Other child-centered	65.0 (3.34)	23.7 (2.23)	3.65 (0.08)	1.87 (0.07)	0.40 (0.08)
Adult-centered	74.7 (3.59)	52.1 (2.49)	3.49 (0.09)	1.87 (0.08)	0.35 (0.08)
non-tCDS	35.3 (3.23)	20.2 (2.12)	4.26 (0.08)	—	—
Spanish					
tCDS					
Child-centered books	63.9 (6.77)	22.2 (5.26)	3.94 (0.13)	2.09 (0.14)	0.38 (0.08)
Other child-centered	51.1 (3.47)	19.8 (2.35)	2.93 (0.08)	1.84 (0.07)	0.36 (0.08)
Adult-centered	56.0 (3.62)	38.0 (2.50)	2.77 (0.09)	1.57 (0.08)	0.31 (0.08)
non-tCDS	29.8 (3.27)	18.1 (2.15)	3.54 (0.08)	—	—

Note: EMM and SE were derived from final models. For tokens/min, types/min, and proportion responses, models included interaction terms. For MLUw, the final model included activity and language as main effects, without the interaction term. For proportion of imitations/expansions, neither language nor activity type were significant fixed effects, however, for descriptive purposes we provide estimated marginal means from a model with both activity and language as main effects. A visual representation of these values can be seen in Figure 4.

comparisons were corrected using the Bonferroni method (alpha = 0.05), grouping simple main effects on activity and language. For pairwise comparisons, EMMs for booksharing and all other activities include only those participants who provided data to both types of activities, and for all other activity comparisons, EMMs include all participants.

The top panel of Figure 4 displays how each of the three linguistic measures (tokens/min, types/min, and MLUw) varied as a function of activity type for families from each language group. Note first that the rate of word tokens was highest in all tCDS compared to non-tCDS activities, on average, for both English- and Spanish-speaking groups (ts > 11.26, ps < 0.001). Thus, caregiver speech contained more words per minute in contexts when that speech was directed to the target child than when it was directed to others around the target child. There was a statistically significant interaction between activity type and language group (p = 0.02). Tokens/min was higher, on average, during booksharing compared to other child-centered activities in the English-speaking families, t (1391) = 3.98, p = 0.001. This pattern was similar, but not significant for the Spanish-speaking families, t (1407) = 1.88, p = 0.97. Note that books and other child-centered activities were not significantly different from adult-centered activities in tokens/min for both English- and Spanish-speaking families (ts < 2.68, ps > 0.12).

We saw a different pattern across activities for the measure of word types/min. In both language groups, types/min were highest in adult-centered activities, compared to other child-centered activities and non-tCDS (ts > 5.96, ps < 0.001). There was a statistically significant interaction between activity type and language group (p = 0.02). More word types/min were produced, on average, during adult-centered activities versus booksharing in English-speaking families, t (1413) = 4.71, p < 0.001), although this trend did not achieve statistical significance in the Spanish-speaking families, t (1349) = 2.84, p = 0.07. Note that the number of different word types/min did not reliably differ between book-sharing and other types of child-centered activities in either the English- or the Spanish-speaking families (ts < 0.95, ps = 1.0). Thus, both types of child-centered activities were associated with similar levels of lexical diversity and did not differ from adult speech during non-tCDS periods.

Turning now to MLUw, Figure 4 shows that across both language groups, utterances during booksharing were longer, on average, than during other child-centered, adult-centered, and non-tCDS activities, ts > 3.38, ps < 0.005. The interaction between activity type and language group was not statistically significant (p = 0.45). Interestingly, MLUs were higher in non-tCDS activities than in other child- and adult-centered activities (ts > 9.38, ps < 0.001), which were not statistically different from each other, t (1239) = 2.44, p = 0.09. Thus, out of all child- and adult-centered activities, utterances during booksharing were the longest, having average levels that were more similar to non-tCDS than to the other tCDS activities.

The bottom two panels in Figure 4 explored the interaction-based measures, responses and imitations/expansions. Note first that caregivers were more likely to respond to children's utterances during booksharing than in adult-centered activities in both groups (ts > 3.62, ps < 0.003). There was a statistically significant interaction between activity type and language group (p = 0.03). For English-speaking families, caregivers were also more likely to respond during booksharing than during other types of child-centered activities, t (804) = 4.31, p < 0.001. This trend was similar but not reliable in the Spanish-speaking families, t (796) = 1.80, p = 0.65. In English-speaking families, the proportion of child utterances to which an adult responded was similar across both child- and adult-centered activities, t

(777) = 0.10, $p = 1.0$. In contrast, caregivers in Spanish-speaking families were more likely to respond to a child utterance during other child-centered than during adult-centered activities, $t(769) = 3.47$, $p = 0.005$.

Finally, for the proportion of imitations/expansions, the main effect of activity type did not meet the prespecified level of statistical significance ($p = 0.055$) in our baseline model. Thus, while imitations and expansions were, on average, higher during book sharing than during other types of child-centered activities (books EMM = 0.40, SE = 0.08; other child-centered EMM = 0.38; SE = 0.08), and while other child-centered activities included more imitations and expansions than adult-centered activities (adult-centered EMM = 0.33, SE = 0.08), no differences were statistically reliable. Moreover, the main effect of language and the language by activity interaction terms did not significantly increase overall model fit (χ^2s > 1.96; ps > 0.16).

3.4 | How Stable Are Features of Speech to Children Within Families Across Activity?

To answer this question, we conducted zero-order Pearson's correlations across activities for each measure, collapsing across groups to ensure sufficient power, minimize spurious findings, and evaluate individual differences. We considered an $r > 0.3$ as moderate and an $r > 0.5$ as a large effect size, and results should be considered primarily descriptive.

We examined stability by considering comparisons across activity types for each of the different measures. As shown in Table 4 and Figure 5, correlations for all measures between booksharing and other child-centered activities were generally moderate to strong (rs = 0.23–0.67), suggesting that most features indicated stability within families across these two activity types. Thus, caregivers who tended to talk more, use longer utterances, and respond more to children's utterances during booksharing also tended to do so during other types of child-centered activities, like play or dressing (types per min was the only measure with effects under 0.3, $r = 0.23$). Correlations were somewhat weaker between booksharing and adult-centered activities (rs = 0.02–0.41), and between other types of child-centered and adult-centered activities (rs = 0.08–0.69). Correlations with periods of non-tCDS were small to large for word tokens and MLUw (rs = 0.24–0.54), and generally weaker for word types (rs = 0.01–0.22). When considering each feature, effect sizes for tokens/min (rs = 0.24–0.65) and MLUw (rs = 0.26–0.69) were generally stronger, while types/min (rs = 0.01–0.23), proportion of responses (rs = 0.21–0.36), and imitations/expansions, (rs = 0.02–0.32) were generally weaker. Thus, there was evidence of moderate to strong stability across child-centered activities (i.e., booksharing vs. other child-centered activities) and in some measures, particularly, word tokens and MLUw.

4 | Discussion

What are caregivers doing when they talk to their child and how do linguistic and interactional features of everyday speech differ across activities and across families? Are families stable in how they talk to their child regardless of what they are doing? Here, we shed new light on these questions by using daylong recordings to describe the landscape of activities that naturally occur day-to-day in a linguistically-diverse sample of families in the Western U.S. We discuss our key findings and their implications.

First, we found that caregivers directed speech to young children across many types of activities throughout the day, including both child-centered activities, when the child was the focus and adult-centered activities when caregivers were focused on their own activities. On average, only about 50% of the 60-min sample in each family consisted of child-centered activities. Notably, even during these densest segments of tCDS, caregivers sometimes did not engage verbally with the target child, but did so with other adults or children (i.e., non-tCDS). These patterns corroborate previous findings noting the ebb and flow of speech to and around children in naturalistic environments (Rosemberg et al. 2020, 2023; Soderstrom and Wittebolle 2013; Tamis-LeMonda et al. 2017). Activities tended

TABLE 4 | Zero-order correlations between activities for linguistic and interactional features of caregiver speech for English- and Spanish-speaking families.

	tCDS activities			tCDS versus non-tCDS activities		
	Books × Other-CC (n = 42)	**Books × AC (n = 42)**	**Other-CC × AC (n = 90)**	**Books × non-tCDS (n = 42)**	**Other CC × non-tCDS (n = 90)**	**AC × non-tCDS (n = 90)**
Word tokens/min	**0.65**	0.29	0.26	**0.34**	**0.51**	0.24
Word types/min	0.23	0.14	0.08	0.01	0.22	0.19
MLUw	**0.67**	**0.41**	**0.69**	0.26	**0.54**	**0.44**
Prop resp	**0.36**	0.21	0.27	—	—	—
Prop imit/exp	**0.32**	0.02	**0.31**	—	—	—

Note: Effect sizes > 0.3 are in bold.

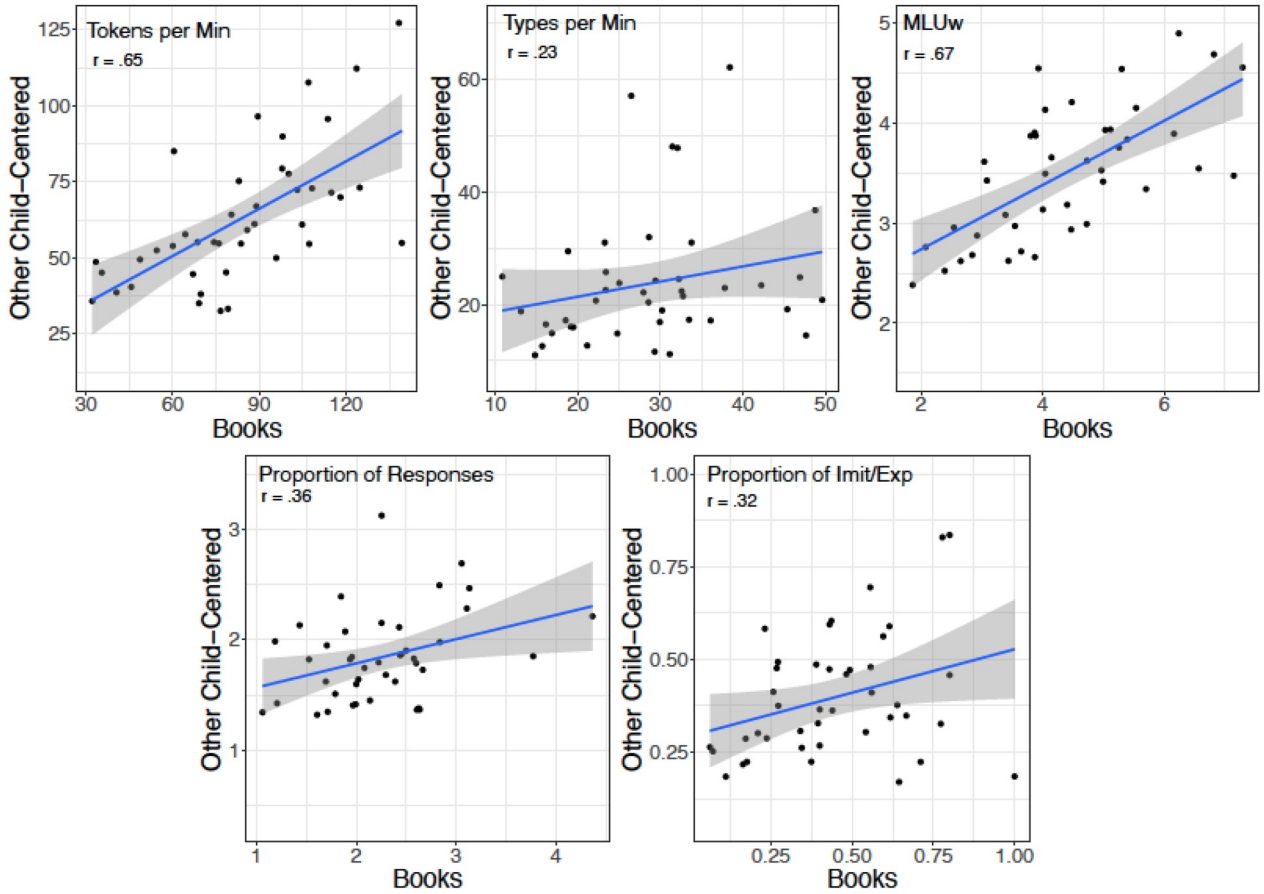

FIGURE 5 | Scatterplots of the associations between booksharing and other-child-centered activities by speech feature for the English- and Spanish-speaking families who engaged in booksharing ($n = 42$). Values represent Pearson's r and shaded areas represent 95% confidence intervals.

to be short and varied. Additionally, activities were layered such that one or more caregivers engaged verbally with the child while doing many things (Gilkerson et al. 2017), interspersed with talking to others apart from the target child. The relatively small proportion of time spent in booksharing and play was an important reminder that prototypical, child-centered activities likely represent a small slice of the everyday verbal interactions between caregivers and young children in many families (Demir-Lira et al. 2019; Soderstrom and Wittebolle 2013). Note that these patterns were very similar in our English- and Spanish-speaking families, revealing common patterns in the daily lives of families with two-year-old children.

Second, building on prior work focusing on child-centered activities (Hoff-Ginsberg 1991; Soderstrom and Wittebolle 2013; Tamis-LeMonda et al. 2019), we observed that children experienced different linguistic and interactive cues that are potentially valuable for language learning when caregivers were engaging in booksharing or other types of child-centered activities, compared to when caregivers were engaging in adult-centered activities. In general, all child-centered activities tended to include fewer word types and more verbal responses to children's utterances than adult-centered activities. It is possible that caregivers were more likely to repeat words during child-centered activities, including booksharing, lowering the levels of lexical diversity compared to adult-centered activities. This may be advantageous, as repetition may be helpful for word learning and building word meaning. Repetition can occur in various ways including self-repetition in variation sets, expansions, and dialogic repetition (Brown 1998; Fusaroli et al. 2023; Küntay and Slobin 1996), as well as when children are exposed to the same words across varied sentential contexts (Horst 2013; Schwab and Lew-Williams 2016; Newman et al. 2016; Tamis-LeMonda et al. 2019). More frequent responses by caregivers may reflect the fact that they are tuning into the child's interests and providing predictable, consistent frames that can reinforce learning and provide a "conceptual springboard" for word learning (Tamis-LeMonda et al. 2019, 2149). More frequent responses to child speech may result in fewer word types. Caregivers may be repeating or expanding on familiar words that the child is also using during the interaction, thereby, supporting the acquisition of meaning (Nelson 1985). Future studies which directly assess caregiver responses together with measures of repetition are needed to explore these possible scenarios.

While there were few differences between booksharing and other types of child-centered activities in terms of lexical diversity (Tamis-LeMonda et al. 2019), booksharing was associated with more word tokens, reliably in the English-speaking families (Tamis-LeMonda et al. 2019; Weizman and Snow 2001) and trending in the Spanish-speaking families. Moreover, caregivers produced longer utterances and were also more likely to respond to their child's utterances during booksharing than during adult-centered activities. These findings support many

prior studies suggesting that booksharing affords optimal opportunities for back and forth verbal engagement (Hoff-Ginsberg 1991; Montag 2019; Snow et al. 1976; Sosa 2016). However, our observation of similar levels of lexical diversity in booksharing and other types of child-centered activities contrasts with some previous findings (e.g., Demir-Lira et al. 2019; Hoff-Ginsberg 1991). These differences could be due to the fact that the children in our study were different ages or because our families read books or engaged in child-centered activities that were more likely to encourage repetition than in previous studies. In addition, our measure is a rate (per minute) computed within a segment, rather than an absolute count or type/token ratio. We initially chose to normalize our measures by minutes per activity per segment in order to allow direct comparisons across activities that varied in duration. However, this approach highlights the fact that our activities were bounded within each 10-min segment, thus capturing clusters of interactions throughout the day. The properties of these bounded interactions may be different from those of the samples used in previous studies.

We also found evidence that many positive features of caregiver speech were seen during adult-centered activities. There were few differences between adult- and child-centered activities in word tokens and MLUw. Moreover, adult-centered activities were more lexically diverse, on average, than other types of child-centered activities, including booksharing. Adult-centered activities, such as shopping or caring for another child, involve a greater variety of actions and objects, and therefore, may elicit a greater variety of words and word referents than activities such as play or dressing. When adults comment on their own activities, children have more opportunities to be exposed to familiar and unfamiliar words than in routine child-centered activities. Thus, the role of speech to children during adult-centered activities should also be valued when considering language learning opportunities. In the reality of daily life, rich and supportive periods of tCDS can occur during a broad array of activities that reflect, in part, what caregivers need to get done in a given day, such as meals and shopping, as well as child-centered activities in which the child is the focus.

Our strategy also allowed us to directly compare features of tCDS with the speech children hear that is not directed to them (i.e., non-tCDS or "overheard"). We found longer average MLUw in non-tCDS than in child-centered activities and in adult-centered activities in both groups of families, although average MLUw in non-tCDS was not longer than in booksharing. This suggests that overhead speech may also provide important cues to learning (Foushee et al. 2016; Foushee and Srinivasan 2024; Girouard-Hallam and Norris 2024).

Third, we found that features of tCDS and non-tCDS were stable within families. Our findings extend previous work with primarily middle-to higher-SES English-speaking families (e.g., Bornstein et al. 1999) by examining evidence for stability in English- and Spanish-speaking families from a diverse sample. The evidence for stability in children's language environments suggests that families tend to behave relatively consistently throughout the day, maintaining their relative rank ordering compared to other families, even if they are engaging in different activities. Thus, while some activities may be more or less likely to elicit different linguistic or interactional features of speech, on average, the individual tendencies of caregivers may be relatively stable for some features. For example, even though booksharing resulted in the longest utterances on average, we found within-family correlations on most measures between booksharing and other types of child-centered activities. This is consistent with the proposal that some families use longer or shorter utterances than other families during all activities in which they engage with the child. The strong correlation between AWC in the densest segments and AWC/hour averaged over the entire daylong recording also suggests stability within individual families (Ferjan Ramírez and Hippe 2024). Note that evidence for stability in daylong naturalistic recordings is particularly striking since activities may have involved different caregivers (e.g., mother during booksharing, grandfather during play), unlike in laboratory or structured home-based observational studies.

However, evidence for stability varied depending on the activities and measures selected. The largest effects were seen when comparing booksharing to other child-centered activities and when using measures of word tokens and MLUw, both of which may reflect verbosity rather than qualitative features of language interaction. It is important to note that examining stability between booksharing and other types of activities was only possible in the half of the families who engaged in this activity on the recording day. Thus, the strength of these relations reflected only these families, and not our entire sample. Word types and responsiveness showed less stability, suggesting that these aspects of tCDS may be more influenced by the types of activities in which caregivers engage in with their children than those measures which reflect overall amount of speech or talkativeness. Note that evidence for stability does not suggest that caregivers are not able to change. Instead, acknowledging stability can provide insights into the preferences or tendencies that caregivers bring to the learning context and may help to identify particular areas of caregivers' strengths and their potential for change. Further, supporting caregivers in how they engage verbally in one activity may allow positive behaviors to generalize to other activities.

4.1 | Similarities and Differences Between Language Groups

In our English- and Spanish-speaking samples, we found many patterns that were similar. For example, there was a similar range in AWC counts between the most and least talkative families, families in both groups divided their time in similar ways across child- and adult-centered activities, and activities occurred across the day. It was also notable that booksharing occurred in only half the families in each group. Importantly, many of the features of tCDS by activity type were remarkably similar across groups. It could be that the lack of these group differences reflect the wide within-group variation in the measures used here. It is also possible that any systematic differences reflect small, rather than moderate to large effects, that could not be reliably observed in our sample.

At the same time, our results revealed a few minor differences in how features of tCDS varied across activity type by language group, specifically, in word tokens, lexical diversity, and proportion responses. We are cautious to interpret between-language differences in number of word tokens and types, because they may result from linguistic factors (e.g., "Give it to me" in English is four words but one-word in Spanish, "dámelo") that are unrelated to the variation in caregiver behaviors that we are trying to capture. Language differences during booksharing could also be due to the types of books (e.g., picture vs. narrative; monolingual vs. bilingual) shared by the families. While most Spanish-speaking families read books in Spanish, some read bilingual or English books which may have altered interactions in ways that resulted in fewer word tokens or more asides about word meanings or pronunciation (Read et al. 2021). Ongoing analyses are exploring the types of books that families chose in relation to the linguistic features of the speech occurring during booksharing episodes (Read et al. 2023).

4.2 | Limitations

We sampled activities during the six densest 10-min periods of recording (by AWC) that were determined to be tCDS, but did not analyze the many moments in the day that were characterized by sparser tCDS, speech provided by other children, or periods that were predominately non-tCDS (de Barbaro and Fausey 2022; Ferjan Ramírez and Hippe 2024). This sampling strategy may have resulted in an overrepresentation of particular types of activities (e.g., see Demir-Lira et al. 2019 for book sharing estimates) and did not consider how time of day might also impact caregiver speech (Casillas et al. 2019). Moreover, because our sample was limited to 10-min segments that were aggregated for each family from many parts of the day, we could not analyze how continuity of activity might impact our findings. Though our goal was to be systematic and consistent in our judgments about the start and end of activities, the lack of video data limits our understanding of the language environment context. We could not provide a comprehensive picture of the frequency and duration of different activity types, since we systematically omitted those that may have occurred in periods of less dense tCDS. Additionally, while normalizing by activity duration (e.g., types/min) allowed us to compare across activities of different lengths, the use of other types of measures may have revealed different patterns. In fact, further analyses have indicated that our episodes of booksharing were higher in lexical diversity compared to non-booksharing episodes when using measures that are based on units of words that are not normalized by time (e.g., measure of textual lexical diversity (MTLD), lexical density, lexical frequency, and lexical contextual diversity, https://osf.io/q26jx/overview; Read et al. 2023). These issues point to how different measures of lexical diversity may be capturing different aspects of talk and the challenges associated with interpreting the construct of lexical diversity across studies that use different measures.

Our families represented only a small portion of the many diverse populations living in the U.S. who care for 2-year-old children. We invite replication of our findings with samples from other sociocultural and language communities. The specific types of activities that would be considered as "child-centered" versus "adult-centered" may vary within and between different communities. Other factors not examined, such as speech heard from media and the number of adult caregivers or older siblings (Bohn et al. 2024; Cycyk and Hammer 2020), were also not tested in relation to speech measures. Future studies should explore patterns of tCDS in different familial contexts, as well as longitudinally across development. The patterns that we observed might shift over time, as practices adapt to changing family needs. While our occupation- and education-based measure of SES was not a significant moderator in our analyses, we did not explore other SES-related factors which could have moderated our effects, for example, stress or financial insecurity (Ellwood-Lowe et al. 2021; Roby and Scott 2022). We also chose to include commonly-used measures in the field with proposed links to child outcomes, however, there are numerous lexical, grammatical, conceptual, and pragmatic measures of caregiver speech that may provide additional insights into theoretically-important and clinically-relevant outcomes (Bestgen 2025; Fergadiotis et al. 2015; Hirsh-Pasek et al. 2015; Tamis-LeMonda et al. 2019). Finally, we did not link features of caregiver verbal engagement to child outcomes, a critical area of future research.

5 | Conclusion

This study provided a broad picture of the activities in which caregivers and their 24-month-old children engaged verbally during their everyday experiences. Our findings showed that families engaged with young children in multiple ways. Naturally-occurring speech to young children occurred during a variety of child-centered, as well as adult-centered, activities. Prototypical child-centered activities were more likely to involve back-and-forth conversations, yet adult-centered activities were characterized by more lexical diversity. Booksharing was associated with longer utterances, but was not different from other types of child-centered activities in lexical diversity and some aspects of caregiver responsiveness. Given that booksharing was uncommon in about half of the families, it will be fruitful to explore other ways that families can support early language interactions. This study also showed that there are relatively stable individual differences among caregivers, especially in word tokens and MLUw, regardless of the activities in which they engage. Overall, findings revealed that variation in child-directed speech is both attributable to what families do and the tendencies of individual families. Moreover, these findings illuminated several ways in which talking to children during a variety of everyday activities offers rich potential learning opportunities.

Author Contributions

Janet Y. Bang: conceptualization, formal analysis, investigation, methodology, project administration, supervision, writing – original draft, writing – review and editing. **Arlyn Mora:** investigation, methodology, project administration, writing – review and editing. **Mónica Munévar:** investigation, methodology, project administration, writing – review and editing. **Anne Fernald:** conceptualization, funding acquisition, investigation, methodology, writing – review and editing.

Virginia A. Marchman: conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, supervision, writing – review and editing.

Acknowledgments

We are grateful to the children and parents who participated and the research assistants who helped code and transcribe: Jessica Magallón, Nadia Segura, Shriya Anand, Sophia Gamboa, Marisol Rodriguez, Maria Lopez, Stephen Lopez, Jesús Esquivel-Barrientos, Laura Jonsson, Kalpana Gopalkrishnan, Maribel Mercardo, Tami Alade, Jaqueline De Paz-Romero, Lesly Leon, Alice Articia, Julia Briones-Avila, and Elizabeth Sanchez. We also thank our colleagues for helpful discussions including Sarah Surrain, Michael C. Frank, Kirsten Read, Alvin Tan, Talya Gómez, and the staff of the Language Learning Laboratory at Stanford University. This work was supported by grants from the National Institutes of Health (Grants R01 HD42235, R01 DC008838, R01 HD092343, 2R01 HD069150), the Schusterman Foundation, the David and Lucile Packard Foundation, the Bezos Family Foundation, and the Stanford Maternal and Child Health Research Institute. The authors declare no conflicts of interest with regard to the funding source for this study.

Data Availability Statement

The pre-registration and codebook for this study can be seen on Open Science Framework (https://osf.io/2uc3d). All data and code are available on GitHub (https://github.com/janetybang/TimetoTalk). Transcripts are available on the CHILDES data repository (https://talkbank.org/childes/access/Eng-NA/FMB_home.html).

References

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67, no. 1: 1–48. https://doi.org/10.18637/jss.v067.i01.

Bestgen, Y. 2025. "Estimating Lexical Diversity Using the Moving Average Type-Token Ratio (MATTR): Pros and Cons." *Research Methods in Applied Linguistics* 4, no. 1: 100168. https://doi.org/10.1016/j.rmal.2024.100168.

Blasi, D. E., J. Henrich, E. Adamou, D. Kemmerer, and A. Majid. 2022. "Over-Reliance on English Hinders Cognitive Science." *Trends in Cognitive Sciences* 26, no. 12: 1153–1170. https://doi.org/10.1016/j.tics.2022.09.015.

Bohn, M., W. F. Da Silva Vieira, M. Giner Torréns, et al. 2024. "Mealtime Conversations Between Parents and Their 2-Year-Old Children in Five Cultural Contexts." *Developmental Psychology* 60, no. 7: 1255–1268. https://doi.org/10.1037/dev0001688.

Bornstein, M. H., D. L. Putnick, and G. Esposito. 2017. "Continuity and Stability in Development." *Child Development Perspectives* 11, no. 2: 113–119. https://doi.org/10.1111/cdep.12221.

Bornstein, M. H., C. S. Tamis-LeMonda, and O. M. Haynes. 1999. "First Words in the Second Year: Continuity, Stability, and Models of Concurrent and Predictive Correspondence in Vocabulary and Verbal Responsiveness Across Age and Context." *Infant Behavior and Development* 22, no. 1: 65–85. https://doi.org/10.1016/S0163-6383(99)80006-X.

Brown, P. 1998. "Conversational Structure and Language Acquisition: The Role of Repetition in Tzeltal." *Journal of Linguistic Anthropology* 8, no. 2: 197–221. https://doi.org/10.1525/jlin.1998.8.2.197.

Casillas, M., P. Brown, and S. C. Levinson. 2019. "Early Language Experience in a Tseltal Mayan Village." *Child Development* 91, no. 5: 1819–1835. https://doi.org/10.1111/cdev.13349.

Chater, N., and M. H. Christiansen. 2018. "Language Acquisition as Skill Learning." *Current Opinion in Behavioral Sciences* 21: 205–208. https://doi.org/10.1016/j.cobeha.2018.04.001.

Cycyk, L. M., and C. S. Hammer. 2020. "Beliefs, Values, and Practices of Mexican Immigrant Families Towards Language and Learning in Toddlerhood: Setting the Foundation for Early Childhood Education." *Early Childhood Research Quarterly* 52: 25–37. https://doi.org/10.1016/j.ecresq.2018.09.009.

Dailey, S., and E. Bergelson. 2022. "Language Input to Infants of Different Socioeconomic Statuses: A Quantitative Meta-Analysis." *Developmental Science* 25, no. 3: e13192. https://doi.org/10.1111/desc.13192.

d'Apice, K., R. M. Latham, and S. von Stumm. 2019. "A Naturalistic Home Observational Approach to Children's Language, Cognition, and Behavior." *Developmental Psychology* 55, no. 7: 1414–1427. https://doi.org/10.1037/dev0000733.

de Barbaro, K., and C. M. Fausey. 2022. "Ten Lessons About Infants' Everyday Experiences." *Current Directions in Psychological Science* 31, no. 1: 28–33. https://doi.org/10.1177/09637214211059536.

Demir-Lira, E., L. R. Applebaum, S. Goldin-Meadow, and S. C. Levine. 2019. "Parents' Early Book Reading to Children: Relation to Children's Later Language and Literacy Outcomes Controlling for Other Parent Language Input." *Developmental Science* 22, no. 3: e12764. https://doi.org/10.1111/desc.12764.

Dunn, J., C. Wooding, and J. Hermann. 1977. "Mothers' Speech to Young Children: Variation in Context." *Developmental Medicine and Child Neurology* 19, no. 5: 629–638. https://doi.org/10.1111/j.1469-8749.1977.tb07996.x.

Ellwood-Lowe, M. E., R. Foushee, and M. Srinivasan. 2021. "What Causes the Word Gap? Financial Concerns May Systematically Suppress Child-Directed Speech." *Developmental Science* 25, no. 1: e13151. https://doi.org/10.1111/desc.13151.

Faul, F., E. Erdfelder, A.-G. Lang, and A. Buchner. 2007. "G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences." *Behavior Research Methods* 39, no. 2: 175–191. https://doi.org/10.3758/bf03193146.

Fergadiotis, G., H. H. Wright, and S. B. Green. 2015. "Psychometric Evaluation of Lexical Diversity Indices: Assessing Length Effects." *Journal of Speech, Language & Hearing Research* 58, no. 3: 840–852. https://doi.org/10.1044/2015_JSLHR-L-14-0280.

Ferjan Ramírez, N., and D. S. Hippe. 2024. "Estimating Infants' Language Exposure: A Comparison of Random and Volume Sampling From Daylong Recordings Collected in a Bilingual Community." *Infant Behavior and Development* 75: 101943. https://doi.org/10.1016/j.infbeh.2024.101943.

Fernald, A., V. A. Marchman, and A. Weisleder. 2013. "SES Differences in Language Processing Skill and Vocabulary Are Evident at 18 Months." *Developmental Science* 16, no. 2: 234–248. https://doi.org/10.1111/desc.12019.

Foushee, R., T. L. Griffiths, and M. Srinivasan. 2016. "Lexical Complexity of Child-Directed and Overheard Speech: Implications for Learning." In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*, 1697–1702.

Foushee, R., and M. Srinivasan. 2024. "Infants Who Are Rarely Spoken to Nevertheless Understand Many Words." *Proceedings of the National Academy of Sciences* 121, no. 23: e2311425121. https://doi.org/10.1073/pnas.2311425121.

Fusaroli, R., E. Weed, R. Rocca, D. Fein, and L. Naigles. 2023. "Caregiver Linguistic Alignment to Autistic and Typically Developing Children: A Natural Language Processing Approach Illuminates the Interactive Components of Language Development." *Cognition* 236: 105422. https://doi.org/10.1016/j.cognition.2023.105422.

Gilkerson, J., and J. A. Richards. 2020. *A Guide to Understanding the Design and Purpose of the LENA System*. LENA Foundation. https://www.lena.org/wp-content/uploads/2020/07/LTR-12_How_LENA_Works.pdf.

Gilkerson, J., J. A. Richards, and K. J. Topping. 2017. "The Impact of Book Reading in the Early Years on Parent–Child Language Interaction." *Journal of Early Childhood Literacy* 17, no. 1: 92–110. https://doi.org/10.1177/1468798415608907.

Girouard-Hallam, L. N., and M. N. Norris. 2024. "Overheard and Understood: A Systematic Review of Children's Learning From Overhearing." *Translational Issues in Psychological Science*. https://doi.org/10.1037/tps0000399.

Gutiérrez-Clellen, V. F., M. A. Restrepo, L. Bedore, E. Peña, and R. Anderson. 2000. "Research to Practice. Language Sample Analysis in Spanish-Speaking Children: Methodological Considerations." *Language, Speech, and Hearing Services in Schools* 31, no. 1: 88–98. https://doi.org/10.1044/0161-1461.3101.88.

Hirsh-Pasek, K., L. B. Adamson, R. Bakeman, et al. 2015. "The Contribution of Early Communication Quality to Low-Income Children's Language Success." *Psychological Science* 26, no. 7: 1071–1083. https://doi.org/10.1177/0956797615581493.

Hoff-Ginsberg, E. 1991. "Mother-Child Conversation in Different Social Classes and Communicative Settings." *Child Development* 62, no. 4: 782–796. https://doi.org/10.2307/1131177.

Hollingshead, A. B. 1975. "The Four Factor Index of Social Status." Unpublished Manuscript. https://sociology.yale.edu/sites/default/files/files/yjs_fall_2011.pdf#page=21.

Horst, J. S. 2013. "Context and Repetition in Word Learning." *Frontiers in Psychology* 4: 149. https://doi.org/10.3389/fpsyg.2013.00149.

Kidd, E., and R. Garcia. 2022. "How Diverse Is Child Language Acquisition Research?" *First Language* 42, no. 6: 14272372110664. https://doi.org/10.1177/01427237211066405.

Küntay, A., and D. Slobin. 1996. "Listening to a Turkish Mother: Some Puzzles for Acquisition." In *Social Interaction, Social Context, and Language: Essays in Honor of Susan Ervin-Tripp*, edited by D. I. Slobin, J. Gerhardt, A. Kyratzis, and J. Guo, 1st ed., 265–286. Erlbaum.

Lenth, R. 2025. "Emmeans: Estimated Marginal Means, Aka Least-Squares Means." *R package version 1.10.7*. https://CRAN.R-project.org/package=emmeans.

MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk*. 3rd ed. Lawrence Erlbaum Associates.

Marchman, V. A., V. N. Bermúdez, J. Y. Bang, and A. Fernald. 2020. "Off to a Good Start: Early Spanish-Language Processing Efficiency Supports Spanish- and English-Language Outcomes at 4½ Years in Sequential Bilinguals." *Developmental Science* 23, no. 6: e12973. https://doi.org/10.1111/desc.12973.

Marchman, V. A., and C. Martínez-Sussmann. 2002. "Concurrent Validity of Caregiver/Parent Report Measures of Language for Children Who Are Learning Both English and Spanish." *Journal of Speech, Language, and Hearing Research* 45, no. 5: 983–997. https://doi.org/10.1044/1092-4388(2002/080.

Massaro, D. W. 2015. "Two Different Communication Genres and Implications for Vocabulary Development and Learning to Read." *Journal of Literacy Research* 47, no. 4: 505–527. https://doi.org/10.1177/1086296X15627528.

Matuschek, H., R. Kliegl, S. Vasishth, H. Baayen, and D. Bates. 2017. "Balancing Type I Error and Power in Linear Mixed Models." *Journal of Memory and Language* 94: 305–315. https://doi.org/10.1016/j.jml.2017.01.001.

Montag, J. L. 2019. "Differences in Sentence Complexity in the Text of Children's Picture Books and child-directed Speech." *First Language* 39, no. 5: 527–546. https://doi.org/10.1177/0142723719849996.

Montag, J. L., M. N. Jones, and L. B. Smith. 2018. "Quantity and Diversity: Simulating Early Word Learning Environments." Supplement, *Cognitive Science* 42, no. S2: 375–412. https://doi.org/10.1111/cogs.12592.

Nelson, K. 1985. *Making Sense: The Acquisition of Shared Meaning*. Academic Press.

Newman, R. S., M. L. Rowe, and N. Bernstein Ratner. 2016. "Input and Uptake at 7 Months Predicts Toddler Vocabulary: The Role of Child-Directed Speech and Infant Processing Skills in Language Development." *Journal of Child Language* 43, no. 5: 1158–1173. https://doi.org/10.1017/S0305000915000446.

Read, K., J. Bang, A. Tan, V. Marchman. 2023. The Rich Language of Book Sharing With English- and Spanish-speaking Families. M. Castellana (Chair), Opportunities for learning from picture books and shared reading in monolingual and bilingual environments [Symposium].

Read, K., P. Contreras, and H. Martinez. 2021. "Tres Formas: Shared Reading Practices With Three Types of Spanish and English Dual-Language Learning Preschoolers." *Bilingual Research Journal* 44, no. 3: 360–380. https://doi.org/10.1080/15235882.2021.1994485.

Roby, E., and R. M. Scott. 2022. "Financial Concern Reduces Child Directed Speech in a Socioeconomically Diverse Sample." *Scientific Reports* 12, no. 1: 9173. https://doi.org/10.1038/s41598-022-13177-2.

Rosemberg, C. R., F. Alam, C. P. Audisio, M. L. Ramirez, L. Garber, and M. J. Migdalek. 2020. "Nouns and Verbs in the Linguistic Environment of Argentinian Toddlers: Socioeconomic and context-related Differences." *First Language* 40, no. 2: 192–217. https://doi.org/10.1177/0142723719901226.

Rosemberg, C. R., F. Alam, M. L. Ramirez, and M. I. Ibañez. 2023. "Activity Contexts and Child-Directed Speech in Socioeconomically Diverse Argentinian Households." *International Journal of Early Childhood* 55, no. 1: 1–25. https://doi.org/10.1007/s13158-022-00345-8.

Rowe, M. L. 2012. "A Longitudinal Investigation of the Role of Quantity and Quality of Child-Directed Speech in Vocabulary Development: Child-Directed Speech and Vocabulary." *Child Development* 83, no. 5: 1762–1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x.

Rowe, M. L., and C. E. Snow. 2019. "Analyzing Input Quality Along Three Dimensions: Interactive, Linguistic, and Conceptual." *Journal of Child Language* 47: 1–17. https://doi.org/10.1017/S0305000919000655.

Rowe, M. L., and A. Weisleder. 2020. "Language Development in Context." *Annual Review of Developmental Psychology* 2, no. 1: 201–223. https://doi.org/10.1146/annurev-devpsych-042220-121816.

Schwab, J. F., and C. Lew-Williams. 2016. "Repetition Across Successive Sentences Facilitates Young Children's Word Learning." *Developmental Psychology* 52, no. 6: 879–886. https://doi.org/10.1037/dev0000125.

Shneidman, L. A., and S. Goldin-Meadow. 2012. "Language Input and Acquisition in a Mayan Village: How Important Is Directed Speech?" *Developmental Science* 15, no. 5: 659–673. https://doi.org/10.1111/j.1467-7687.2012.01168.x.

Snow, C. E., A. Arlman-Rupp, Y. Hassing, J. Jobse, J. Joosten, and J. Vorster. 1976. "Mothers' Speech in Three Social Classes." *Journal of Psycholinguistic Research* 5, no. 1–19: 1–20. https://doi.org/10.1007/bf01067944.

Soderstrom, M., and K. Wittebolle. 2013. "When Do Caregivers Talk? The Influences of Activity and Time of Day on Caregiver Speech and Child Vocalizations in Two Childcare Environments." *PLoS One* 8, no. 11: e80646. https://doi.org/10.1371/journal.pone.0080646.

Sokolov, J. L., and B. MacWhinney. 1990. "The CHIP Framework: Automatic Coding and Analysis of Parent-Child Conversational Interaction." *Behavior Research Methods, Instruments, & Computers* 22, no. 2: 151–161. https://doi.org/10.3758/BF03203138.

Sosa, A. V. 2016. "Association of the Type of Toy Used During Play With the Quantity and Quality of Parent-Infant Communication." *JAMA Pediatrics* 170, no. 2: 132. https://doi.org/10.1001/jamapediatrics.2015.3753.

Tamis-LeMonda, C. S., S. Custode, Y. Kuchirko, K. Escobar, and T. Lo. 2019. "Routine Language: Speech Directed to Infants During Home Activities." *Child Development* 90, no. 6: 1–18. https://doi.org/10.1111/cdev.13089.

Tamis-LeMonda, C. S., Y. Kuchirko, R. Luo, K. Escobar, and M. H. Bornstein. 2017. "Power in Methods: Language to Infants in Structured and Naturalistic Contexts." *Developmental Science* 20, no. 6: e12456. https://doi.org/10.1111/desc.12456.

U.S. Census Bureau. 2024. "Language Spoken at Home." American Community Survey, ACS 5-Year Estimates Subject Tables." Table S1601, 2022 [Dataset]. https://data.census.gov/table/ACSST5Y2022.S1601?g=050XX00US06085.

Weisleder, A., and A. Fernald. 2013. "Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary." *Psychological Science* 24, no. 11: 2143–2152. https://doi.org/10.1177/0956797613488145.

Weizman, Z. O., and C. E. Snow. 2001. "Lexical Input as Related to Children's Vocabulary Acquisition: Effects of Sophisticated Exposure and Support for Meaning." *Developmental Psychology* 37, no. 2: 265–279. https://doi.org/10.1037/0012-1649.37.2.265.

Wittenburg, P., H. Brugman, A. Russel, A. Klassman, and H. Sloetjes. 2006. "ELAN: A Professional Framework for Multimodality Research." In *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*. https://archive.mpi.nl/tla/elan.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting Information S1: infa70051-sup-0001-suppl-data.docx.