



## Conversation Initiation of Mothers, Fathers, and Toddlers in their Natural Home Environment

Mark VanDam<sup>a,b,\*</sup>, Lauren Thompson<sup>a</sup>, Elizabeth Wilson-Fowler<sup>c</sup>, Sarah Campanella<sup>a</sup>, Kiley Wolfenstein<sup>a</sup>, Paul De Palma<sup>d</sup>

<sup>a</sup> Speech & Hearing Sciences, Elson S. Floyd College of Medicine, Washington State University, 412 E. Spokane Falls Blvd, Spokane, WA, 99202, USA

<sup>b</sup> Hearing Oral Program of Excellence of Spokane, HOPE School, 502 E. 5th Ave., Spokane, WA 99202, USA

<sup>c</sup> Communication Sciences and Disorders, Eastern Washington University, 310 North Riverpoint Blvd, Box B, Spokane, WA, 99202, USA

<sup>d</sup> Computer Science, School of Engineering and Applied Science, Gonzaga University, 502 E. Boone Ave., Spokane, WA, 99258, USA

### ARTICLE INFO

#### Keywords:

Conversation initiation  
Conversational dynamics  
Daylong vocal analysis

### ABSTRACT

In a conversational exchange, interlocutors use social cues including conversational turn-taking to communicate. There has been attention in the literature concerning how mothers, fathers, boys, and girls converse with each other, and in particular who initiates a conversation. Better understanding of conversational dynamics may deepen our understanding of social roles, speech and language development, and individual language variability. Here we use large-scale automatic analysis of 186 naturalistic daylong acoustic recordings to examine the conversational dynamics of 26 families with children about 30 months of age to better understand communication roles. Families included 15 with boys and 11 with girls. There was no difference in conversation initiation rate by child sex, but children initiated more conversations than mothers, and mothers initiated more than fathers. Results support developmental theories of the different and variable roles that interlocutors play in a social context.

### 1. Introduction

There has been sustained interest and theoretical attention concerning the role of social interactions in development (Vygotsky, 1978; Rogoff, 1990; Tomasello, 1992, 2003). Much of the emphasis has been on the roles of early vocal development and joint attention in young children learning to communicate and engage in conversation. The scientific interest in these subjects has a long pedigree, extending back to the early twentieth century (Ronjat, 1913; Smith, 1926, 1935; Adamson et al., 2014; Bornstein et al., 2008; Bloom et al., 1996; Barton and Tomasello, 1991; Oller, 2000; Pinker, 2003). In a conversational exchange, interlocutors use a complex set of linguistic rules (such as how plurals are formed or the word order of phrases) and social cues to communicate (Schacter, 1979; Tomasello and Farrar, 1986). Among those are gestures (Krauss et al., 1996), gaze (Cary, 1978; Sidnell, 2007), modulation of behavior (Wiener et al., 1972; Ochs and Schieffelin, 2017), and many well-documented linguistic and communicative behaviors such as turn-taking and conversational repair (Uno, 2017; Yang and Heeman, 2010; Grothendieck, et al., 2010; Sidnell, 2007; Black and Logan, 1995). For example, children have been shown to use longer utterances when engaged in conversation with adults as compared with conversations with other children (Smith, 1935). Vocal exchanges between children and their caregivers lay fundamental

\* Corresponding author. Department of Speech & Hearing Sciences, Elson S. Floyd College of Medicine, Washington State University, 412 E. Spokane Falls Blvd., Spokane, WA 99202, Tel. 509.358.7599

E-mail address: [mark.vandam@wsu.edu](mailto:mark.vandam@wsu.edu) (M. VanDam).

<https://doi.org/10.1016/j.csl.2021.101338>

Received 12 March 2020; Received in revised form 11 June 2021; Accepted 30 November 2021

Available online 2 December 2021

0885-2308/© 2021 Elsevier Ltd. All rights reserved.

groundwork for eventual conversation and communication (Bateson, 1975; Oller, 2000; Tomasello, 2008; Oller et al., 2013). These early vocal and conversational exchanges may help to better understand the influence of sex, race, or social differences in the course of speech and language development (Phillips 1973; Huttenlocher et al., 1991; Clark, 2013).

Some researchers have identified sex differences in the ways boys and girls use language. A recent study looking at 13,783 European children's communicative gestures, word comprehension, word production, and onset of early two-word utterances found that girls were generally ahead of boys. It also found that there was no more statistical variance among boys than girls (Eriksson et al., 2012). It is clear that relative differences among boys is far from settled science. Arden and Plomin (2006), for instance, find greater variance among boys in intelligence, which Bjorklund and Causey (2017) interpret as boys being more variable in verbal skills. Whatever the differences between boys and girls, they may further develop into sex-specific speech and behavior patterns throughout the toddler and preschool years (Kaushanskaya et al., 2013). Other researchers have suggested that young children may adapt to or emulate certain speech and conversational characteristics of their parents, learning from both fathers' and mothers' communication strategies (Langman and Shi, 2020; Maltz and Borker, 2012; Leaper, 1991). Sex differences may also be the result of girls' increased use of symbolic activities related to language (Brandlistuen et al., 2020; Toivainen et al., 2017; Bates et al., 1979; Bornstein et al., 1999).

A relatively recent area of interest in the study of language use is conversation initialization (Tomasello, 1988; Hoff, 2006; Ayers and Sonandre, 1999; Pancsofar and Vernon-Feagans, 2006). The person who initiates a conversation is more likely to choose the topic or frame the content of the exchange (Goffman, 1974; Linell, 1998; Uno, 2017), is perceived to be more confident and more accurate (Wright and Carlucci, 2011), and may increase the ability to influence interlocutors (Gabbert et al., 2006; Nguyen et al., 2021). Children who initiate an interaction are more likely to elicit a response or increase the overall rate of conversations or interactions (Tomasello et al., 2005; Zheng et al., 2020). Conversations initiated by children have been shown to be longer than conversations initiated by adults (Cazden, 1970). Toddlers have been shown to have more mature vocal utterances when they initiate conversations as opposed to when their mothers do the initiating (Ko et al., 2016). Communicative function of conversational initiations may also differ by sex. Girls may initiate conversations to establish or maintain relationships (Maltz and Borker, 2012; Trevarthen and Aitken, 2001), while boys may initiate conversations to establish dominance or attract and maintain an audience (Hyde, 2014; Kugler et al., 2018; Maltz and Borker, 2012). Another notable factor in this discussion is the influence of cultural context on interpersonal and conversational exchanges. It has been shown, for example, that working- and middle-class mothers and fathers do not engage in the same types of conversations with their children (Yu, Bonawitz, and Shafto, 2019). Other studies have shown communicative effects of social class (Hart and Risley, 1995), geographical and social differences in families (Sperry, Sperry, and Miller, 2018), and socio-economic status (Hoff, 2006; Huttenlocher, Waterfall, Vasilyeva, Vevea, and Hedges, 2010). We are not aware of studies that have looked explicitly at rate of conversation initiation by social factors. These results point to substantial, but varied, reported effects of conversational initiation, the subject of the present work.

In addition to the direct role that children play, mothers and fathers contribute to and participate in the communicative development of their children. Mothers and fathers also differ in what role they play in conversations (Maltz and Borker, 2012; Lundy, 2003). Fathers have been shown to initiate (and respond to) fewer conversations than children (Hiadek and Edwards, 1984). On the other hand, fathers appear to be more physical in interactions with their children than do mothers (Power and Parke, 1982; Lamb et al., 2017). There are other differences. Mothers engage in more frequent and longer conversations with their children than fathers do (Golinkoff and Ames 1979; Craig and Evans 1991), regardless of communicative disabilities of their children (Kristensen et al., 2020). Fathers contribute language that is more formal in tone, has greater lexical diversity, is shorter, and incorporates problem-solving and linguistic play (Tomasello et al., 1990; Rondal, 1980). Mothers, on the other hand, contribute language that is more comment-driven and child-content oriented (Aries, 1987).

Some studies have examined conversational characteristics of mothers and fathers to sons and daughters. The reports conflict. Whiting and Edwards (1988) indicate that mothers focus more conversational attention on their sons than on their daughters. Two later studies indicate that mothers focus more conversational attention on their sons (Leaper et al., 1998; Muchnik and Stavans 2009). Still others have found that mothers converse equally with their sons and daughters (VanDam et al., 2016). Children have also been shown to communicate differently with mothers and fathers (Leaper, 1991; VanDam and De Palma, 2014), engaging in more overall conversation with mothers as compared with fathers (Craig and Evans, 1991; De Palma and VanDam, 2017). In short, the differential conversational style of mothers and fathers with their daughters and sons is an open research area.

There is a long history of researchers using technology to address questions of language use and behavior of the sort discussed above. Vocal dosimetry has been used, albeit on small-scale and sparingly, in the 20<sup>th</sup> century (Harkness, 1975, 1977; Korman, 1984; Wells, 1979; Hart and Risley, 1995, 2003), but began to make serious strides in the early 2000s (Gilkerson and Richards, 2008) due to improved hardware, data storage, and processing techniques, among other advances (VanDam and De Palma, 2019; Beckman, et al., 2017; Dubey, et al., 2017). As a result of the sea-change in technology including the LENA (Gilkerson and Richards, 2008), the Virtual Speech Kitchen (Metze et al., 2013), and massive increases in available data such as the HomeBank data project (VanDam et al., 2016), there now exist the tools and raw data to address questions old and new using hardware and techniques that were unavailable to earlier research endeavors.

This report relies on data collected by the LENA device (Gilkerson and Richards, 2008; LENA Research Foundation, Boulder, CO). The system consists of a small audio recorder, worn by a child, and associated software to process the resulting daylong recordings. The child goes about his or her daily routine, collecting a naturalistic audio record of the day from the child's perspective. The automatic processing software outputs an uninterrupted audio file of the day and a computer file diarizing segment boundaries and labels corresponding to communicative events such as who is talking, when the TV/radio is turned on, and silence.

This work seeks to investigate the psychosocial phenomenon of conversational exchanges using modern computational hardware and processing techniques. It adds to the contributions of other researchers using similar techniques. For example, one study used

automatic methods to analyze over 13,000 hours of audio to look at communication dynamics in families with a child with autism (Warlaumont, Richards, Gilkerson, and Oller, 2014). Another looked at how very young infants' vocalizations elicited vocal responses from their caregivers (Prezter, Lopez, Walle, and Warlaumont, 2019). Still others have examined infants' sensitivity to parent conversations (Casillas and Frank, 2017), and another captured over 8000 conversational turns to examine the temporal characteristics of turn-taking (Hilbrink, Gattis, and Levinson, 2015). These point to a growing literature that leverages automatic tools to gain insight into human behavioral areas.

In this study, we make use of large-scale data collection and automatic speech processing that estimates vocal characteristics or vocal doses of talkers in a wild, ecologically valid environment. More specifically, we examine conversation initiation using data extracted from daylong recordings collected in a family setting. The following research questions are posed: (1) do toddler girls initiate more conversations than toddler boys? (2) do parents initiate more conversations than their toddlers? (3) do mothers initiate more conversations than fathers? (4) what proportion of family utterances and from which family members do not elicit a response or conversation? (5) which family members are most likely to make unrequited utterances? The empirical results are then discussed in terms of the extant literature. Finally, the advantages and limitations of the approach used in this study are presented.

## 2. Method

### 2.1. Participants

Twenty-six families participated. All families included a mother and father living with their toddler. There were 15 families with toddler boys and 11 families with toddler girls. Families in this study participated over the course of about a year, contributing 5 to 9 total daylong recordings each ( $M = 7.1$  recordings,  $SD = 0.5$  recordings), totaling 186 daylong recordings. Ages of all children at the time of recording ranged from 25-37 months, and the range was not different between boys and girls (boys' range: 25-36 months; girls' range: 25-37 months;  $t = 0.085$ ,  $d.f. = 184$ ,  $p = 0.533$ ). Given the total age range, variability in number of total recordings, and limited distributional differences between boys' and girls' ages, we pooled within-family data to obtain a mean age (and other within-family data across the sampled ages as described below). The mean age of boys ( $M = 31.0$  months,  $SD = 3.9$  months) was not different from the mean age of girls ( $M = 29.6$  months,  $SD = 4.3$  months) ( $t = 1.254$ ,  $d.f. = 24$ ,  $p = 0.111$ ).

Seven of the boys and 6 of the girls were only-children, and the remainder had up to three additional siblings. All children in the study (including the siblings) were reported by parents to be typically developing. All children passed a hearing screen. Self-reported socio-economic status (SES) was collected for the families using a 12-point scale of mother's education (1 = completed elementary school, 2 = completed junior high, 3 = received General Education Diploma (high school equivalence), 4 = completed high school, 5 = completed 1 or more years of technical/vocational school, 6 = completed technical/vocational school, 7 = completed one or more years of university/college, 8 = bachelor's degree, 9 = completed one or more years of graduate school, 10 = master's degree, 11 = course work completed for PhD, but no dissertation; law degree without bar; medical degree without internship completed, 12 = Ph.D.; law degree with bar; medical degree with internship completed). The pooled SES for all families was estimated ( $M = 8.2$ ,  $SD = 2.0$ ; range: 4-12). A two-tailed, unpaired  $t$ -test indicated that SES for families with boys ( $M = 7.9$ ,  $SD = 2.4$ ; range: 4-12) was not different from families with girls ( $M = 8.4$ ,  $SD = 1.2$ ; range: 4-12) ( $t = 0.633$ ,  $d.f. = 24$ ,  $p = 0.532$ ). Of the 15 families with boys, one family identified as Black, two as Hispanic/Latino, and the remainder White; of the 11 families with girls, one family identified as Hispanic/Latino, one as Asian/Pacific Islander, and the remainder as White. These demographics are similar to those reported in the literature (Haas, 2008; Vabel, Cohen, Leonard, Glymour, Duarte, and Yen, 2018), but exact comparisons are not possible due to different methods. All families spoke English as their primary language and report speaking only English at home. All families lived in the United States. Recordings were collected from families in Georgia, Illinois, Michigan, Oregon, Washington, Arizona, Idaho, Iowa, Montana, California, Florida, and New Jersey.

### 2.2. Materials and equipment

We employed a commercially-available, wearable audio recorder and associated processing software, the Language ENvironment Analysis (LENA; LENA Research Foundation, Boulder, CO, USA). The LENA comprises hardware to collect an *in situ* audio recording from the perspective of the child and software to process the audio offline using automatic speech processing (ASP) techniques. In this case, the software does a probabilistic mapping from audio signal to labels. The labels indicate the talker rather than text, as in conventional automatic speech recognition. The recording hardware is worn by the child in a specialized shirt or vest. An uninterrupted, unprocessed, daylong audio file is collected. The audio file is then uploaded to a desktop computer and processed by proprietary speech processing techniques to assess properties of the auditory environment. The output of the system is a 16-bit, 16kHz, pulse-code modulated (PCM), lossless WAV file of the audio and an XML-coded report of the results of the ASP procedure. The system has been described in detail in various reports in the literature (see, Gilkerson and Richards, 2008; Oller et al., 2010; Gilkerson et al., 2017).

There has been considerable attention in the empirical literature paid to the validity of the LENA software labeling (Warren et al., 2010; Oller et al., 2010; Canault et al., 2016; VanDam and Silbert 2016; Gilkerson et al., 2017; Busch et al., 2018; Ganek and Eriks-Brophy, 2018; VanDam and De Palma, 2019; Bulgarelli and Bergelson, 2019; Jones, Plesa Skwerer, Pawar, Hamo, et al., 2019). The evidence in the extant literature suggests support for the use of the LENA diarization output, although a number of shortcomings have also been pointed out. For example, a recent systematic review of the literature found 33 studies that looked at LENA automatic performance, finding correlations between automatic labeling and human labeling from  $r=0.36-.79$  (Cristia, Bulgarelli, and Bergelson,

2020). Another study looking at the automatic performance of the LENA system (Cristia, Lavechin, Scaff, Soderstrom, Rowland, Räsänen, Bunce, and Bergelson, 2020) found that the automated “conversational turn count” (CVC) generated by the system was among the lower performing variables when compared with human ratings. Though we do not use the CVC counts directly, our own custom-built software, described below, uses LENA-generated classifications. In another report (Cristia, Lavechin, et al., 2020), the authors point out that regardless of absolute and relative error rates, “whether LENA results are “good enough” depends largely on the goals of each particular study” (p. 16).

The diarization output includes onset and offset times which segment the recording and assign specific output labels to each segment. Among other labels, the present work was interested in the temporal order of the talker labels. These indicate live human speech production from the mother, the father, and the child wearing the recording device. The LENA classification system requires that segments are a minimum of 600ms in duration, are identified solely as a live voice, and are not interrupted by silence or noise greater than 800ms (described in Oller et al., 2010).

LENA segments and labels the daylong recording. We use these segments and labels to construct conversation initiation counts (a subject we introduce here and describe in detail below). For example, adjacent segments (of any length) of the mother followed by the child counts as a mother-initiated conversation. A similar sequence begun by the mother, but which included an intervening noise segment before the child segment would not be counted as a conversation. It should be noted that the software attempts to identify merely adult female (“FAN”), adult male (“MAN”), and target child (“CHN”), among other labels such as those corresponding to vocalizations of other children, environmental noise, and silence. The LENA system also identifies segments as adult female, adult male, and target child (“FAF,” “MAF,” and “CHF,” respectively) which are probabilistically “far” from the statistical likelihood model (hence the “F” in the category label contrasting with “N” indicating “near” to the model in other labels). These lower probability “far” segments were not used in the current study. We reasonably infer that the labels “FAN,” “MAN,” and “CHN” correspond to mother, father, and child, respectively. It is possible, for example, that an adult male who is not the child’s father is labeled as “MAN” and would ultimately be included in our category corresponding to “father.” In practice, these errors are surely present in the data, but we assume (1) the real presence of talkers who are not the mother or father are relatively rare, (2) language learning, social and conversational dynamics, and development may be minimally sensitive to a parent-non-parent distinction, and (3) there are substantial known valid data (see VanDam and Silbert, 2016; Silbert et al., 2013) that are minimally influenced by relatively rare events.

To determine who initiates a conversation, a custom parser was developed using MATLAB (VanDam, 2020). The custom scripts take XML files (known to LENA researchers as “ITS files”) as input, computes the conversational turn count, and generates summary values of the contingencies described immediately below. LENA ASP routines generate a record of the ordered segment labels. As noted above, these are used to determine the presence and order of conversations. The result is a record of the initiation of conversational turns for each talker. We identified each segment bearing the LENA-generated labels corresponding to mother, father, and target child, as described above, and assessed whether the segment immediately following was another of the target labels. Sequences of talker labels uninterrupted by other labels were considered a conversation, and only the first talker in a conversation was counted as the talker who initiated. This was important in determining “conversation initiation” apart from turns within a conversation. For example, a sequence of child-mother-child would count as a single conversation initiated by the child and not count for a conversation initiation by the mother. We do not here report “conversational turns” in which each talker exchange, regardless of utterance, is counted as a “turn.” Here we identify each “conversation” and report on the talker who initiates, regardless of how many turns follow, granted that at least one turn is required for a conversation, or it is classified as an unrequited utterance. Our analysis provides no information on the semantic or lexical content of the speech. This is to say, *talker adjacency* is operationalized as a proxy for a more complex, fuller understanding of *conversation*.

It is worth noting here the tradeoff between *in situ* generated data of the kind described in this study and that collected and coded in laboratory setting. The present approach allows for massive amounts of data, far more than can be collected in a laboratory. There is also a tradeoff with coding inaccuracy resulting from judgment calls made in manual coding done by live human judges. The data here are processed by a computer. On the other hand, the interpretation of conversation is necessarily less rich. The data used in this paper are a subset the Cougar Corpus (VanDam, 2018) described in detail in VanDam et al (2016).

### 2.3. Procedure

Families were recruited as part of a larger study from childcare facilities and preschools using flyers, email invitations, and word-of-mouth. Families were provided written, informed consent to participate in the study. The study was reviewed and approved by the human subjects committee and institutional review boards of the first author’s university. Parents provided written responses to a detailed intake form which included demographic and family information. Researchers stated to the families that the goal of the study was to collect regular monthly recordings for one year. This goal was not achieved by any family due to scheduling. Instead, families contributed less often, actually contributing from 5 to 9 total recordings in one 12-month span. We opted to retain the overall duration of 12 months, even though contributions were less frequent than anticipated, to provide a uniform developmental window within observations. All recordings took place in the homes and natural daily environments of participating families without the presence of the authors or researchers. Prior to each scheduled recording day, families were provided with a shirt or vest with a custom pocket designed to fit the LENA audio recorder. Each family was instructed to begin recording when the child woke in the morning and end recording at bedtime in the evening. Since, by design, the child carries the recorder, the recordings are from the auditory perspective of the child during his or her waking hours.

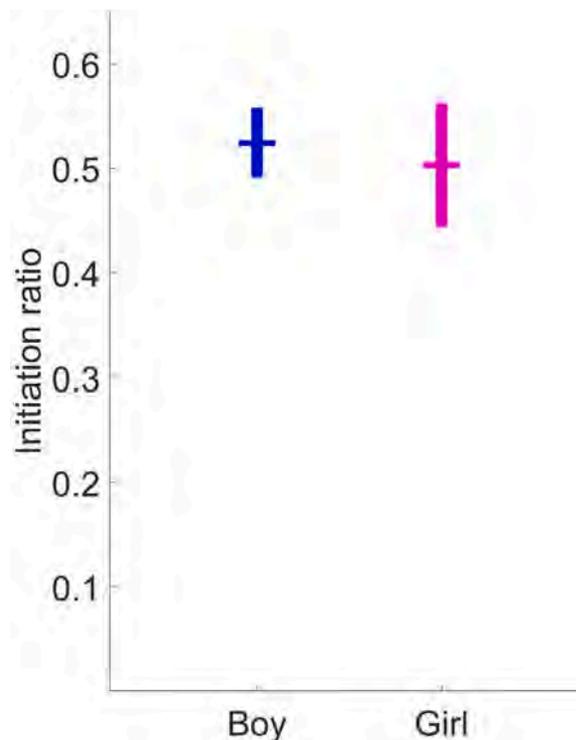
We asked families to record during a day that was typical for them and when parents and the child were all present. We did not request continuous presence of all family members throughout the day only that family members were active participants during some

part of the recording day. We asked that recordings be completed on a day when the family was usually present and not on a day when, for example, the mother was out of town on a business trip or the child spent the bulk of the day at a relative's house. Families submitted written logs spanning half-hour blocks over the entire recording time. These described typical activities and active participants. Families also indicated, in retrospect, whether the day went as planned, how typical the recording day was, and the general health of family members. This log was intended to screen for prospective sickness or unusual daily activities or routines. None were indicated. No recordings were screened or excluded as a result of the daily activity log.

After each daylong recording was collected, the recorder and daily activity log were retrieved by researchers and processed offline in the laboratory. LENA Pro software (V3.1.4-136 r9771) was used to process recordings and output the XML formatted diarized results, as described above. The proprietary LENA software processes the daylong audio recordings seeking to identify speech-like utterances by identifying high-energy periods bounded by low-energy periods, categorize those utterances by talker, and perform other cascading analyses looking to recover details of the articulatory mechanisms responsible for the acoustic output (Oller et al., 2010; Gilkerson & Richards 2008; Gilkerson et al., 2017). The system uses acoustic rhythm, spectral characteristics, prosody, and duration as key acoustic variables of interest (details are given in Oller et al., 2010).

#### 2.4. Data and analysis

Raw data from each recording was processed and summary estimates of conversation initiations by talker were tallied. The number of conversation initiations by each talker was then divided by the total number of conversations to obtain the proportion of total conversations initiated. The sum of each talker's total initiations sums to 1.0. Utterances from each family member that were not followed by another utterance were also tallied to provide an estimate of unrequited speech. The proportion of unrequited utterances was determined as the number of utterances neither preceding nor following another family member's utterance, divided by the total number of utterances from that talker. A within-family average was then computed across all recording for each family, collapsing across multiple recordings to reduce the possibility of an outlier affecting the data. The data subjected to statistical analysis consisted of one observation for each family at each variable of interest. Unpaired, two-sample, two-tailed *t*-tests were used to test for difference in sample means and an ANOVA was conducted with the proportion of unrequited utterances as the dependent variable. The alpha-criterion is nominally set to the traditional threshold of 0.05, but in all cases, we report exact probability values or thresholds substantially smaller than the alpha criterion. Figures below showing error bars indicate the mean and the 95% confidence interval about the mean.



**Figure 1.** Conversations of boys and girls. Girls do not initiate conversations more than boys. Mean conversation initiation ratios and 95% confidence intervals about the mean are plotted.

### 3. Results

Boys' and girls' mean rate of conversation initiation was not shown to differ ( $M_{\text{boys}} = 0.524$ ,  $SD_{\text{boys}} = 0.060$ ,  $M_{\text{girls}} = 0.503$ ,  $SD_{\text{girls}} = 0.087$ ;  $t(24) = 0.724$ ,  $p = 0.475$ ). Rate of conversation initiation is shown below in Figure 1.

Because boys and girls did not differ in their rate of conversation initiation, all children were pooled into a single category. Mothers', fathers', and children's mean rate of conversation initiation was different for all pairwise comparisons. Children's rate of conversation initiation ( $M_{\text{children}} = 0.515$ ,  $SD_{\text{children}} = 0.072$ ) was significantly greater than their mothers ( $M_{\text{mothers}} = 0.365$ ,  $SD_{\text{mothers}} = 0.091$ ;  $t(50) = 6.597$ ,  $p < 10^{-7}$ ) and significantly greater than their fathers ( $M_{\text{fathers}} = 0.117$ ,  $SD_{\text{fathers}} = .060$ ;  $t(50) = 21.516$ ,  $p < 10^{-26}$ ). Mothers rate of conversation initiation was significantly greater than fathers ( $t(50) = 11.520$ ,  $p < 10^{-14}$ ). Rate of conversation initiation is shown below in Figure 2.

Individual family conversation initiation ratio is shown in Figure 3 below. In all but two families, mothers initiated more conversations than fathers.

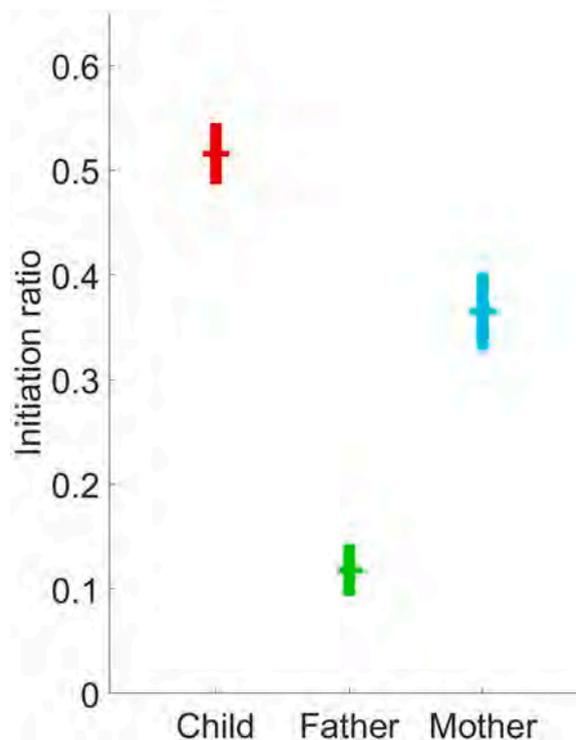
A two-way ANOVA was conducted that examined the effect of child sex and family member initiating an utterance on the proportion of unrequited utterances. Sex of the target child (boy, girl) and family member initiating (mother, father, child) were the independent variables. The model was significant ( $F(5, 72) = 20.002$ ,  $p < 10^{-11}$ ). There was a small main effect of sex ( $F(1, 72) = 9.043$ ,  $p = 0.004$ ), with boys showing a higher proportion of unrequited utterances ( $M = 0.59$ ,  $SD = 0.06$ ) compared to girls ( $M = 0.56$ ,  $SD = 0.05$ ). There was a main effect of family member initiating ( $F(2, 72) = 41.980$ ,  $p < 10^{-12}$ ), with mothers ( $M = 0.58$ ,  $SD = 0.03$ ), fathers ( $M = 0.51$ ,  $SD = 0.04$ ), and children ( $M = 0.63$ ,  $SD = 0.05$ ) mutually statistically different in all pair-wise comparisons shown by LSD posthoc tests (all  $ps < 0.002$ ). There were no interaction effects between the factors ( $F(2, 72) = 0.734$ ,  $p = 0.483$ ).

### 4. Discussion

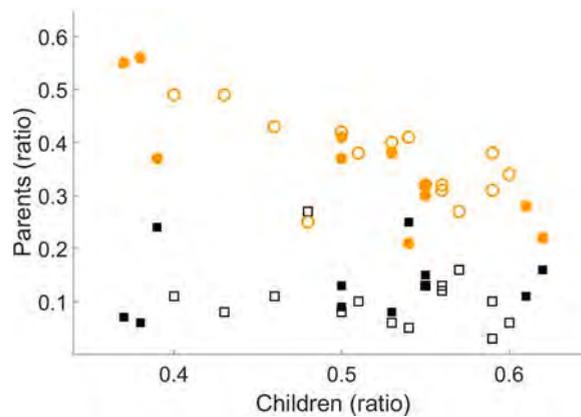
This work (1) did not find a difference in toddler boys' and girls' rate of conversation initiation, (2) demonstrated that, regardless of child sex, children initiate more conversations than their parents, (3) demonstrated that mothers initiate more conversations than fathers, and (4) showed that boys have more unrequited utterances than girls, children have more unrequited utterances than their parents, and mothers have more unrequited utterances than fathers.

#### 4.1. Theoretical implications

The finding that toddler boys and girls, at least in this sample, initiate conversations at the same rate may appear surprising and



**Figure 2.** Family conversations. Family conversations are initiated most often by children, then by mothers, and least frequently by fathers. All pairwise comparisons are significantly different. Mean conversation initiation ratios and 95% confidence intervals about the mean are plotted.



**Figure 3.** Individual family conversations. Conversations by mothers, fathers, boys and girls in relation to each other. Mothers are plotted in orange markers and fathers are plotted in black markers. Parents' conversation initiation ratio is given on the ordinate for both mothers and fathers. Children's conversation initiation ratio is plotted on the abscissa. Boys are plotted in open markers and girls in filled markers. Each within-family mother-father pair is represented by the unique child value on the abscissa. The sum of the mother, father, and child ratios must equal 1.00 for all families.

deserves further investigation, given the extensive empirical literature on the differential behavior of boys and girls (Arden and Plomin, 2006; Aries, 1987; Bjorklund and Causey, 2017; Eriksson et al., 2012; Hoff, 2006; Huttenlocher et al., 1991; Whiting and Edwards, 1988). Many of the reports in the literature showing how family members interact with each other have used more traditional measures such as human transcription, whereas we employed automatic methods to make observations. This study however, looked explicitly at a communicative skill (i.e., conversation initiation), while not examining the linguistic or semantic content of the utterances. Obviously, the automatic methods employed here are the source of this trade-off, with a number of consequences outlined below. Encouragingly, there are a growing number of studies that use automatic methods similar to those used here to examine relationships among conversations and factors such as reading skills (Merz et al., 2020) and cultural differences (Ganek, Smyth, Nixon, and Eriks-Brophy, 2018).

The finding that children initiate more conversations than their parents may support children's latent interest in conversation. When children initiate conversations, others talk more (Tomasello et al., 2005), talk longer (Cazden, 1970), and the children themselves use more mature vocal utterances (Ko et al., 2016). There have been a number of reasons offered as to why children initiated more conversations. These include persuading adults to act (Bartsch, Wright, and Estes, 2010), asserting a right to choose the topic of conversation (Linell, 1998), and showing confidence (Wright and Carlucci, 2011). From a developmental perspective, these notions suggest that children may be motivated to engage more and on their own terms (for example by dictating the topic of conversation). These actions likely lead to more joint attention and mutual engagement, thus bootstrapping development.

There has been some historical interest in the empirical behavioral literature looking at the role of fathers in the development of their children (Lamb, 2004), and in the development of language and communication in particular (Gleason, 1975; Perlmann and Gleason, 1993; McLaughlin et al., 1983; Tamis-LaMonda et al., 2013). Though it is important to point out that this study does not attempt to interpret the content of family speech, it does support the view that fathers speak less frequently to their children than do mothers. The question, of course, is why this should be the case. There is evidence that differential input from parents may be due to different, possibly tacit, goals of mothers and fathers. Adapted from early 20<sup>th</sup> century insights by Vygotsky (1978, 1986), the "apprenticeship model" suggests that learners, as apprentices, may learn differentially from different agents—the masters—in their socio-cultural spheres (Wells, 1980; Wertsch, 1985). A similar model referred to as the "bridge hypothesis" (Gleason, 1975; Mannle and Tomasello, 1987; Rondal, 1980) suggests that fathers, in particular, serve as a bridge to the outside world. In the case of vocal development, children may learn different speech and language use details from mothers, fathers, and other interlocutors in their social contexts. For example, mothers may use or model certain conversational styles which children learn to associate with the social role of their mothers. If fathers model this style less often or not at all, the association may be weaker with fathers and favor another style that models, for example, the outside world. This approach may also interact with statistical regularities in language learning in which children are sensitive to distributional characteristics of the input (Hintzman, 1984). For example, Saffran et al (1996) show that very young babies detect phonological regularities in the input and use those to segment lexical items. Others have argued that the statistical regularities in the communicative signal may give children insight into the social world and have an influence on development (Ruffman et al., 2012); Tasimi, 2020). It is possible, then, that young language learners are also sensitive to input characteristics that are stored and accessed in terms of social interactions. These may include who initiates a conversation and what role initiation plays in the development of social interactions. The present work offers a description of differences between mothers' and fathers' communication with their children. Mothers, for instance, initiate more conversations than fathers in a typical family setting, while others tend to have fewer unrequited utterances than mothers or children. Children may key into these statistical patterns and consequently develop an understanding of the role of specific interlocutors in communicative exchanges. Children, for example, may recognize the statistical regularity that fathers tend to elicit a response at a higher rate than they or their mothers. This could modulate

their expectations of conversational exchanges including attention, development of pragmatics, or language learning (Clark, 2018).

The present findings show that different rates of conversation initiation do not necessarily elicit a response, but they also document individual variability and may indicate complexity in communicative interactions that have not been thoroughly documented in the literature. It is particularly interesting that here we show boys elicited fewer responses to their utterances than girls. Boys got no response to 59% of utterances and girls to 56% of utterances. The absolute difference is relatively small, and this result could be examined in greater detail in additional work. It is possible that differences by sex are either uniquely revealed by the automatic methods or are an artifact of the automatic methods themselves. This warrants further attention. Eliciting a response has been shown to affect rate of conversation (Tomasello et al., 2005), maturity of vocalization (Ko et al., 2016), and vocal coordination (Abney, Warlaumont, Oller, Wallot, and Kello, 2016). These instances of vocal and social development may help to explain differences between boys and girls in the time course of their development. From this, we conclude that there are productive theoretical issues of the present work both in the face value of the findings and in further directions suggested by the findings.

#### 4.2. Practical implications

This study employed computational techniques to analyze 186 daylong recordings, totaling 1915.8 total hours of raw audio in which 19070 conversations were identified. In addition to the content of conversational exchange initiations, this study demonstrates that modern technology and very large datasets can address issues in communicative development. This project, then, is a demonstration of the new tools and methods available which, to date, have been only explored in a limited fashion. The principal advantages of this approach have to do with both size and situation. The size of the dataset examined exceeds that which may be examined under laboratory conditions. The situation of collection, home recordings, suggests a higher ecological validity than in the laboratory. These attributes, among others, suggest methodological extensibility with clear long-term potential.

Another practical application includes considering conversation initiation more explicitly as a dependent variable to better understand populations not explicitly sampled here. There may be interest in younger or older children, children from varying socio-cultural backgrounds, families who speak languages other than or in combination with English, or children with disabilities. In particular for children with communicative disorders or disabilities, the automatic assessment of conversation initiation may lead to earlier or differential diagnosis, alternative educational approaches, or more targeted intervention strategies. For example, it has been shown that fathers talked more to their boys with hearing loss and mothers talked more to their girls with hearing loss (VanDam et al., 2016; Kristensen, Sundby, Hauge, and Löfvkist, 2020). These sex differences in communicative use are currently only beginning to be explored. If the automated technique used here is applied to a broader sample, new insights into language and development could be explored.

#### 4.3. Limitations

Language use and conversation are complex human behaviors. The goal of the present work was focused on specific aspects of conversation initiations and the computational methods used to look at these phenomena, but there are contributions to communication not considered here.

The technology used in this work, the LENA, makes objective estimates of vocal and conversational characteristics of interlocutors, but the operationalization may not strictly parallel human-transcribed data or the ultimate behavioral phenomenon of interest here. For example, it has been shown grouping talkers by categories such as “man,” “woman,” and “child” are obtained through different decision channels for computers versus humans (VanDam and Silbert, 2016). In particular, automatic speech processing algorithms appear to use fundamental frequency and durational characteristics in a different way than human judges. These process differences impact the eventual labels applied to stimuli. It is a limitation of the present work that we rely on the specific labels that are automatically generated by the software. If those labels lack construct or content validity, the ability to draw conclusions is limited.

There have been many reports on the consistency, accuracy, precision, and validity of the LENA technology in a variety of domains (Oller et al., 2010; Warren et al., 2010; Xu et al., 2009; Gilkerson et al., 2017). Those reports typically find support for the labeling consistency of the diarization results, but still performance is different from human transcribers. For example, the LENA automatic labeling procedures have been shown to be biased: female utterances were mislabeled as male in 22% of instances but male utterances were mislabeled as female in only 3% of utterances, and female utterances were mislabeled as children in 6% of utterances while male utterances were mislabeled as children in only 1% of utterances (Cristia, Lavechin et al., 2020). Cristia and colleagues further caution that LENA labels corresponding to the father may be confused with other talkers or non-speech categories at a higher rate than children and mothers. These reports suggest fathers’ utterances are over-labeled, and children and mothers are more often confused than children and fathers. Although we cannot directly compare our results with the data in Cristia, Lavechin, et al., the direction of those trends in their report would likely bolster our findings with respect to fathers (possibly increasing the magnitude of the effects demonstrated here), but may reduce the magnitude of effects between children and mothers shown here. Further, there are only limited published reports on the fidelity of utterance segmentation procedures which rely on LENA-generated data (Oller et al., 2010; Cristia, Lavechin et al., 2020; Xu, Yapanel, Gray, 2009). Overall, the literature regarding utterance segmentation procedures is sparse and does not include the level of detail necessary to easily test the methods or attempt to reproduce critical results.

Another utterance segmentation issue, as pointed by Cristia and colleagues, is that there is a lack of one-to-one correspondence between the eventual segmentation of LENA methods and the more general understanding of what constitutes useful units of speech, communication, or language (Cristia, Lavechin et al., 2020). There is much current attention in the literature on improving the technology, including work on computational methods that use modern, and more transparent approaches (Metze et al., 2013;

Schuller et al., 2017; Räsänen, Seshadri, Lavechin, Cristia, and Casillas, 2020; Lavechin, Bousbig, Bredin, Dupoux, and Cristia, 2020) and commercially available audio (and video) recording hardware (Casillas, Brown, and Levinson, 2019).

It is possible that there are additional structures in the data that were not considered here, but which may play a role in how family conversational dynamics play out. For example, there is some evidence that fathers interact differently with their sons than their daughters (Siegal, 1987; De Palma and VanDam, 2017). Other evidence suggests that mothers and fathers may be differentially sensitive to specific details of a child's ability or disorder (Konstantareas et al., 1988; Flipen and Watson, 2015). Parents and children accommodate to one another as a function of age and maturity (Tamis-LaMonda and Bornstein, 1991; Vigil et al., 2005), which may factor into communicative styles. We do not have the data to thoroughly address contextual and social influences of language and conversational use, but social factors are known to influence communication (Hoff, 2006). In future work we may consider a wider range of variables known to influence speech and development, which may in turn reveal additional structures in communication.

It is also important to note that the operationalization of "conversation" here certainly lacks a deeper understanding of human interaction, including joint attention, semantic or topical qualities of communication, and other factors. It should also be noted that any characterization of an initiation necessarily includes or assumes a conversation. Here we use ordinal-dependent talker transitions as a proxy for conversations, but there are at least three issues with that approach. First, we do not have direct evidence that actual content was exchanged. For example, the method used here would not distinguish between two talkers having a meaningful conversation and two talkers vocalizing in a shared auditory environment but unrelated to one another. Although unrelated vocalizations likely occur, we do not access the (meaningful) content of the recordings, so no firm evidence is at hand. In the case of unrelated vocalizations, the automatic methods employed here may have counted a conversation and subsequent initiation in a real-world circumstance in which no interpersonal communication occurred, a situation that would not reasonably be included in a more perfect analysis. Second, there are temporally dependent characteristics of conversational and turn-taking exchanges that are more nuanced approaches to determining and understanding exchanges (Levinson and Torriera, 2015). The current approach used exclusively ordinal relationship of talkers, but another, not particularly computationally heavy approach, might take into account the timing of utterances, pauses, overlapping vocals, and other interactional characteristics. Third, there may be communicative events that would reasonably be considered conversations that are not accounted for here because they do not follow the temporal pattern used to define conversations. For example, a mother may give a directive to a child and the child might give a gestural response such as a head shake or shrug. This would be recorded in the audio record not as an exchange, thus not as a conversation, while nevertheless a communicative exchange could be understood to have taken place. In that case of a child engaging non-verbally, the exchange might be tallied as an unrequited utterance from the mother, no conversation initiation, and no child input. All of these conclusions may belie the true communicative event that took place. It is not immediately clear if the bulk of missed characterizations would tend to over- or under-estimate overall conversation initiations. There is not a clear remedy to this issue given the methods undertaken here. Taking these details into account may give a finer understanding of interactional vocal behavior.

#### 4.4. Summary and conclusion

This work studies conversation initiation in a home setting among mothers, fathers, and their toddlers, where the mean age of boys was 31.0 months and the mean age of girls was 29.6 months. Conversations were collected through audio recording devices worn by the children throughout their waking hours. Twenty-six families contributed 186 recordings, for a total of 1915.8 hours. The recordings were processed using the LENA speech classification software which labels the start and stop times of conversational participants. Speaker adjacency (e.g., a child speaking followed by an adult speaking) was taken as a proxy for a conversation. Using this definition, we examined 19070 conversations using custom-build software.

We found that boys and girls (1) initiate conversation at the same rate and (2) initiate more conversations than their parents, regardless of sex of the child or sex of the parent. We found, further, that (3) mothers initiate more conversations than fathers. We also report here that (4) boys elicit fewer conversational responses, children fewer than their parents, and mothers fewer than fathers.

Large scale, daylong audio recordings and automatic processing have recently enjoyed increasing recognition and productivity in the scientific literature. The ability to capture and represent complex interactional details within a family and within the family's natural environment using objective tools is relatively new and presents researchers with the opportunity to examine old questions with new methodology and new questions altogether. The promise of this set of tools is to better understand and describe high-level social and linguistic features of interactions. There remain certain limitations that future studies and greater application may resolve to sharpen and improve the tools.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgements

This work was supported by the National Science Foundation, SBE-RIDIR 1539133, 1539129, 1539010; the National Institutes of Health, NIDCD R01DC009569, DC009560-01S1; WSU Seed Grant, 124172-001; and the Washington Research Foundation.

## Reference

- Adamson, L.B., Bakeman, R., Deckner, D.F., Nelson, P.B., 2014. From interactions to conversations: The development of joint engagement during early childhood. *Child Dev* 85 (3), 941–955.
- Arden, R., Plomin, R., 2006. Sex differences in variance of intelligence across childhood. *Personality Individ. Diff.* 41 (1), 39–48.
- Aries, E., 1987. Gender and communication. In: Shaver, P., Hendrick, C. (Eds.), *Sex and Gender*. Sage Publications, Newbury Park, CA, pp. 149–176.
- Ayres, J., Sonandre, D.A., 1999. The effect of “initiating”; on communication apprehension and other variables in initial interactions between able-bodied persons and persons with a disability. *Commun Res Reports* 16 (2), 138–146.
- Barton, M.E., Tomasello, M., 1991. Joint attention and conversation in mother-infant-sibling triads. *Child Dev* 62 (3), 517–529.
- Bartsch, K., Wright, J.C., Estes, D., 2010. Young children’s persuasion in everyday conversation: Tactics and attunement to others’ mental states. *Soc. Devel.* 19 (2), 394–416. <https://doi.org/10.1111/j.1467-9507.2009.00537.x>.
- Bates, E., Benigni, L., Bretherton, L.L., Camaioni, L., Volterra, V., 1979. The emergence of symbols: Cognition and communication in infancy. Academic Press, New York, NY.
- Bateson, M.C., 1975. Mother-infant exchanges: the epigenesis of conversational interaction. *Annals New York Acad. Sci.* 263 (1), 101–113.
- Beckman, M.E., Plummer, A.R., Munson, B., Reidy, P.F., 2017. Methods for eliciting, annotation, and analyzing databases for child speech development. *Computer Speech and Lang* 45, 278–299.
- Bjorklund, D., Causey, K., 2017. *Children’s thinking: Cognitive development and individual thinking*, 6th ed. Sage, Los Angeles, CA.
- Black, B., Logan, A., 1995. Links between communication patterns in mother-child, father-child, and child-peer interactions and children’s social status. *Child Dev* 66 (1), 255–271.
- Bloom, L., Margulis, C., Tinker, E., Fujita, N., 1996. Early conversations and word learning: Contributions from child and adult. *Child Dev* 67 (6), 3154–3175.
- Bornstein, M.H., Haynes, O.M., Pascual, L., Painter, K.M., Galperin, C., 1999. Play in two societies: Pervasiveness of process, specificity of structure. *Child Dev* 70 (2), 317–331.
- Bornstein, M.H., Tamis-LeMonda, C.S., Hahn, C.S., Haynes, O.M., 2008. Maternal responsiveness to young children at three ages: Longitudinal analysis of a multidimensional, modular, and specific parenting construct. *Dev. Psychol.* 44 (3), 867–874.
- Brandlistuen, R.E., Flato, M., Stoltenberg, C., Helland, S.S., Wang, M.V., 2020. Gender gaps in preschool age: A study of behavior, neurodevelopment and pre-academic skills. *Scand. J. Pub. Health*, 1403494820944740. <https://doi.org/10.1177/1403494820944740>.
- Bulgarelli, F., Bergelson, E., 2019. Look who’s talking: A comparison of automated and human-generated speaker tags in naturalistic daylight recordings. *Behav. Res. Meth.* <https://doi.org/10.3758/s13428-019-01265-7>.
- Busch, T., Sangan, A., Vanpoucke, F., van Wieringen, A., 2018. Correlation and agreement between Language ENvironment Analysis (LENA™) and manual transcription for Dutch natural language recordings. *Behav. Res. Meth.* 50 (5), 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. *Behav. Res. Meth.* 48 (3), 1109–1124.
- Cary, M.S., 1978. The role of gaze in the initiation of conversation. *Soc. Psychol.* 41 (3), 269–271.
- Casillas, M., Frank, M.C., 2017. The development of children’s ability to track and predict turn structure in conversation. *J Memory Lang* 92, 234–253.
- Cazden, C.B., 1970. The neglected situation in child language research and education. In: Williams, F. (Ed.), *Language and poverty*. Markham, Chicago.
- Clark, E.V., 2018. Conversation and language acquisition: A pragmatic approach. *Lang. Learn. Devel.* 14 (3), 170–185. <https://doi.org/10.1080/15475441.2017.1340843>.
- Clark, E.V., 2013. *First language acquisition*, 2nd Ed. Cambridge University Press, Cambridge, UK.
- Craig, H.K., Evans, J.L., 1991. Turn exchange behaviors of children with normally developing language: The influence of gender. *J. Speech Lang. Hearing Res.* 34 (4), 866–878.
- Cristia, A., Bulgarelli, F., Bergelson, E., 2020. Accuracy of the Language Environment Analysis system segmentation and metrics: A systematic review. *J. Speech, Lang. Hear. Res.*, 63 (4), 1093–1105. [https://doi.org/10.1044/2020\\_JSLHR-19-00017](https://doi.org/10.1044/2020_JSLHR-19-00017).
- Cristia, A., Lavechin, M., Scaff, C., Soderstrom, M., Rowland, C., Räsänen, O., Bunce, J., Bergelson, E., 2020. A thorough evaluation of the Language Environment Analysis (LENA) system. *Behav. Res. Meth.* 1–30. <https://doi.org/10.1017/CBO9781107415324.004>.
- De Palma, P., and VanDam, M., 2017. Using automatic speech processing to analyze fundamental frequency of child-directed speech stored in a very large audio corpus. In 2017 Joint 17th World Congress of International Fuzzy Systems Association and 9th International Conference on Soft Computing and Intelligent Systems (IFSA-SCIS). doi:10.1109/IFSA-SCIS.2017.8023224.
- Dubey, H., Sangwan, A., Hansen, J.H.L., 2017. Using speech technology for quantifying behavioral characteristics in peer-led team learning sessions. *Computer Speech and Lang* 46, 343–366.
- Eriksson, M., Marschik, P.B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg, S., Gallego, C., 2012. Differences between girls and boys in emerging language skills: Evidence from 10 language communities. *British J. Dev. Psychol.* 30 (2), 326–343.
- Flippin, M., Watson, L.R., 2015. Fathers’ and mothers’ verbal responsiveness and the language skills of young children with autism spectrum disorder. *Amer. J Speech-Lang. Pathol.* 24 (3), 400–410.
- Gabbert, F., Memon, A., Wright, D.B., 2006. Memory conformity: Disentangling the steps toward influence during a discussion. *Psychon. Bull. Rev.* 13 (3), 480–485.
- Ganek, H.V., Eriks-Brophy, A., 2018. A concise protocol for the validation of Language ENvironment Analysis (LENA) conversational turn counts in Vietnamese. *Commun. Disord. Quart.* 39 (2), 371–380.
- Ganek, H., Smyth, R., Nixon, S., Eriks-Brophy, A., 2018. Using the Language ENvironment Analysis (LENA) system to investigate cultural differences in conversational turn count. *J. Speech, Lang., Hear. Res.* 61 (9), 2246–2258.
- Gilkerson, J., Richards, J.A., 2008. The LENA natural language study. LENA Foundation, Boulder, CO. Retrieved from: <https://www.lena.org/wp-content/uploads/2016/07/LTR-02-2.Natural.Language.Study.pdf>.
- Gilkerson, J., Richards, J.A., Warren, S.F., Montgomery, J.K., Greenwood, C.R., Kimbrough Oller, D., Paul, T.D., 2017. Mapping the early language environment using all-day recordings and automated analysis. *Amer. J. Speech-Lang. Pathol.* 26 (2), 248–265.
- Gleason, J.B., 1975. Fathers and other strangers: Men’s speech to young children. In: Dato, D.P. (Ed.), *Developmental psycholinguistics: Theory and applications*. Georgetown University Press, Washington, DC.
- Goffman, E., 1974. *Frame analysis: An essay on the organization of experience*. Harvard University Press, Cambridge, MA.
- Golinkoff, R.M., Ames, G.J., 1979. A comparison of fathers’ and mothers’ speech with their young children. *Child Dev* 50 (1), 28–32.
- Harkness, S., 1975. Cultural variation in mothers’ language. *Word* 27, 495–498.
- Harkness, S., 1977. Aspects of social environment and first language acquisition in rural Africa. In: Snow, C.E., Ferguson, C.A. (Eds.), *Talking to children: Language input and acquisition*. Cambridge University Press, Cambridge, UK, pp. 309–316.
- Hart, B., Risley, T.R., 1995. Meaningful differences in the everyday experience of young american children. Paul H. Brookes Publishing, Baltimore.
- Hart, B., Risley, T., 2003. The early catastrophe. *Amer. Educator* 27 (4), 6–9.
- Haas, S., 2008. Trajectories of functional health: The ‘long arm’ of childhood health and socioeconomic factors. *Social Sci. & Med.* 66 (4), 849–861. <https://doi.org/10.1016/j.socscimed.2007.11.004>.
- Hiadek, E., Edwards, H., 1984. A comparison of mother-father speech in the naturalistic home environment. *J. Psycholing. Res.* 13, 321–332.
- Hilbrink, E.E., Gattis, M., Levinson, S.C., 2015. Early developmental changes in the timing of turn-taking: a longitudinal study of mother–infant interaction. *Frontiers Psych* 6, 1492. <https://doi.org/10.3389/fpsyg.2015.01492>.
- Hintzman, D.L., 1984. MINERVA 2: A simulation model of human memory. *Behav. Res. Meth. Instruments Comp.* 16 (2), 96–101.
- Hoff, E., 2006. How social contexts support and shape language development. *Dev. Rev.* 26 (1), 55–88.

- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., Lyons, T., 1991. Early vocabulary growth: Relation to language input and gender. *Dev. Psychol.* 27 (2), 236–248. <https://doi.org/10.1037/0012-1649.27.2.236>.
- Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., Hedges, L.V., 2010. Sources of variability in children's language growth. *Cog Psych* 61, 343–365. <https://doi.org/10.1016/j.cogpsych.2010.08.002>.
- Hyde, J., 2014. Gender similarities and differences. *Ann. Rev. Psychol.* 65 (1), 373–398.
- Jones, R.M., Plesa Skwerer, D., Pawar, R., Hamo, A., Carberry, C., Ajodan, E.L., Caulley, D., Silverman, M.R., McAdoo, S., Meyer, S., Yoder, A., Clements, M., Lord, C., Tager-Flusberg, H., 2019. How effective is LENA in detecting speech vocalizations and language produced by children and adolescents with ASD in different contexts? *Autism Res* 12 (4), 628–635. <https://doi.org/10.1002/aur.2071>.
- Kaushanskaya, M., Gross, M., Buac, M., 2013. Gender differences in child word learning. *Learn. Indiv. Diff.* 27, 82–89.
- Ko, E.S., Seidl, A., Cristia, A., Reimchen, M., Soderstrom, M., 2016. Entrainment of prosody in the interaction of mothers with their young children. *J. Child Lang.* 43 (2), 284–309.
- Konstantareas, M.M., Mandel, L., Homatidis, S., 1988. The language patterns mothers and fathers employ with their autistic boys and girls. *Appl. Psycholing.* 9 (4), 403–414.
- Korman, M., 1984. Adaptive aspects of maternal vocalizations in differing contexts at ten weeks. *First Lang* 5, 44–45.
- Krauss, R.M., Chen, Y., Chawla, P., 1996. Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? In: Zanna, M. (Ed.), *Advances in experimental social psychology*. Academic Press, San Diego, CA, pp. 389–450.
- Kristensen, N.M., Sundby, C.F., Hauge, M.N., Löfvikist, U., 2020. Female caregivers talk more to 18-56-months-old children with and without hearing impairment than male caregivers measured with LENA™ – A cross-sectional pilot study. *Internat. J. Ped. Otorhinolaryngology* 130, 109809.
- Kugler, K.G., Reif, J.A.M., Kaschner, T., Brodbeck, F.C., 2018. Gender differences in the initiation of negotiations: A meta-analysis. *Psychol. Bull.* 144 (2), 198–222.
- Lamb, M.E. (Ed.), 2004. The role of the father in child development. John Wiley and Sons, Hoboken, NJ.
- Lamb, M.E., Pleck, J.H., Charnov, E.L., Levine, J.A., 2017. A biosocial perspective on paternal behavior and involvement. In: Altman, J. (Ed.), *Parenting across the life span: Biosocial dimensions*. Routledge, New York, NY, pp. 111–142.
- Langman, J., Shi, X., 2020. Gender, language, identity, and intercultural communication. In: Jackson, J. (Ed.), *The Routledge Handbook of Language and Intercultural Communication, Second Edition*. Routledge, New York, NY, pp. 211–226.
- Lavechin, M., Bousbib, R., Bredin, H., Dupoux, E., and Cristia, A., 2020. An open-source voice type classifier for child-centered daylong recordings. arXiv preprint arXiv:2005.12656. URL: <https://arxiv.org/abs/2005.12656>.
- Leaper, C., 1991. Influence and involvement in children's discourse: Age, gender, and partner effects. *Child Dev* 62 (4), 797–811.
- Leaper, C., Anderson, K.J., Sanders, P., 1998. Moderators of gender effects on parents' talk to their children: A meta-analysis. *Dev. Psychol.* 34 (1), 3–27.
- Levinson, S.C., Torreira, F., 2015. Timing in turn-taking and its implications for processing models of language. *Front. Psychol.* 6, 731. <https://doi.org/10.3389/fpsyg.2015.00731>.
- Linell, P., 1998. *Approaching dialogue: Talk, interaction and contexts in dialogical perspectives*. John Benjamins Publishing, Amsterdam.
- Lundy, B.L., 2003. Father-and mother-infant face-to-face interactions: Differences in mind-related comments and infant attachment? *Infant Behav. Dev.* 26 (2), 200–212.
- Maltz, D.N., Borker, R.A., 2012. A cultural approach to male-female miscommunication. In: Monaghan, L., Goodman, J.E., Robinson, J.M. (Eds.), *A cultural approach to interpersonal communication: Essential readings*. Wiley and Sons, Malden, MA, pp. 169–177.
- Mannle, S., Tomasello, M., 1987. Fathers, siblings, and the bridge hypothesis. *J. Child Lang.* 6, 23–42.
- McLaughlin, B., White, D., McDevitt, T., Raskin, R., 1983. Mothers' and fathers' speech to their young children: similar or different? *J. Child Lang.* 10 (1), 245–252.
- Merz, E.C., Maskus, E.A., Melvin, S.A., He, X., Noble, K.G., 2020. Socioeconomic Disparities in Language Input Are Associated With Children's Language-Related Brain Structure and Reading Skills. *Child Dev* 91 (3), 846–860.
- Metze, F., Fosler-Lussier, E., and Bates, R., 2013. The Speech Recognition Virtual Kitchen. *Proc. Interspeech*, 1858–1860.
- Muchnik, M., Stavans, A., 2009. Telling the same story to your child: Mothers' versus fathers' storytelling interactions. *Women Lang* 32 (1), 60–69.
- Nguyen, T., Schleihauf, H., Kayhan, E., Matthes, D., Vrtička, P., Hoehl, S., 2021. Neural synchrony in mother-child conversation: Exploring the role of conversation patterns. *Social Cog. Affect. Neurosci.* 16 (1-2), 93–102. <https://doi.org/10.1093/scan/nsaa079>.
- Ochs, E., Schieffelin, B.B., 2017. *Acquiring conversational competence*. Routledge, London.
- Oller, D.K., 2000. The emergence of the speech capacity. Psychology Press, New York, NY.
- Oller, D.K., Buder, E.H., Ramsdell, H.L., Warlaumont, A.S., Chorna, L., Bakeman, R., 2013. Functional flexibility of infant vocalization and the emergence of language. *Proc. Nat. Acad. Sci.* 110 (16), 6318–6323.
- 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. *Proc. Nat. Acad. Sci.* 107 (30), 13354–13359.
- Pancsofar, N., Vernon-Feagans, L., 2006. Mother and father language input to young children: Contributions to later language development. *J. Appl. Dev. Psych.* 27 (6), 571–587.
- Perlmann, R.Y., Gleason, J.B., 1993. The neglected role of fathers in the children's communicative development. *Sem. Speech Lang.* 14 (4), 314–324.
- Phillips, J.R., 1973. Syntax and vocabulary of mothers' speech to young children: Age and sex comparisons. *Child Dev* 44 (1), 182–185.
- Pinker, S., 2003. *The language instinct: How the mind creates language*. Penguin, London, UK.
- Power, T.G., Parke, R.D., 1982. Play as a context for early learning. In: Laosa, L.M., Sigel, I.E. (Eds.), *Families as learning environments for children*. Springer, Boston, MA, pp. 147–178.
- Pretzer, G.M., Lopez, L.D., Walle, E.A., Warlaumont, A.S., 2019. Infant-adult vocal interaction dynamics depend on infant vocal type, child-directedness of adult speech, and timeframe. *Infant Behav Dev* 57, 101325.
- Räsänen, O., Seshadri, S., Lavechin, M., Cristia, A., Casillas, M., 2020. ALICE: An open-source tool for automatic measurement of phoneme, syllable, and word counts from child-centered daylong recordings. *Behav Res Meth* 1–18. <https://doi.org/10.3758/s13428-020-01460-x>.
- Rogoff, B., 1990. *Apprenticeship in thinking: Cognitive development in social context*. Oxford University Press, Oxford, UK.
- Rondal, J.A., 1980. Fathers' and mothers' speech in early language development. *J. Child Lang.* 7 (2), 353–369.
- Ronjat, J., 1913. *Le développement du langage observé chez un enfant bilingue*. Librairie Ancienne H. Champion, Paris.
- Ruffman, T., Taoupepeau, M., Perkins, C., 2012. Statistical learning as a basis for social understanding in children. *Brit J Devel Psych* 30 (1), 87–104. <https://doi.org/10.1111/j.2044-835X.2011.02045.x>.
- Saffran, J.R., Aslin, R.N., Newport, E.L., 1996. Statistical learning by 8-month-old infants. *Science* 274 (5294), 1926–1928.
- Schachter, F.F., 1979. *Everyday mother talk to toddlers: Early intervention*. Academic Press, New York, NY.
- Schuller, B., Steidl, S., Batliner, A., Bergelson, E., Krajewski, J., Janott, C., ... and Warlaumont, A.S., 2017. The interspeech 2017 computational paralinguistics challenge: Addressee, cold and snoring. *Computational Paralinguistics Challenge (CompParE)*, *Proc. Interspeech*, 3442–3446.
- Sidnell, J., 2007. Comparative studies in conversation analysis. *Ann. Rev. Anthro.* 36, 229–244.
- Siegal, M., 1987. Are sons and daughters treated more differently by fathers than by mothers? *Dev. Rev.* 7 (3), 183–209.
- Silbert, N., Linck, J.A., VanDam, M., 2013. Multilevel models, covariates, and controlled factors in experimental speech research: Unified analyses of highly structured data. *Proc. Meet. Acoust.* 19 (1), 060029.
- Smith, M.E., 1926. An investigation of the development of the sentence and the extent of vocabulary in young children. In: Baldwin, B.T. (Ed.), *University of Iowa Studies in Child Welfare, Vol. III, no. 5*. University of Iowa Press, Iowa City, pp. 1–92.
- Smith, M.E., 1935. A study of the speech of eight bilingual children of the same family. *Child Dev* 6, 19–25.
- Tamis-LeMonda, C.S., Baumwell, L., Cabrera, N.J., 2013. Fathers' role in children's language development. In: Cabrera, N.J., Tamis-LeMonda, C.S. (Eds.), *Handbook of father involvement: Multidisciplinary perspectives*. Routledge, New York, NY, pp. 135–150.

- Tamis-LeMonda, C.S., Bornstein, M.H., 1991. Individual variation, correspondence, stability, and change in mother and toddler play. *Infant Behav. Dev.* 14 (2), 143–162.
- Tasimi, A., 2020. Connecting the dots on the origins of social knowledge. *Perspectives Psychol Sci* 15 (2), 397–410. <https://doi.org/10.1177/1745691619885861>.
- Toivainen, T., Papageorgiou, K.A., Tosto, M.G., Kovas, Y., 2017. Sex differences in non-verbal and verbal abilities in childhood and adolescence. *Intelligence* 64, 81–88. <https://doi.org/10.1016/j.intell.2017.07.007>.
- Tomasello, M., 1988. The role of joint attentional processes in early language development. *Lang. Sci.* 10 (1), 69–88.
- Tomasello, M., 1992. The social bases of language acquisition. *Social development* 1 (1), 67–87.
- Tomasello, M., 2003. The key is social cognition. In: Gentner, D., Goldin-Meadow, S. (Eds.), *Language in mind: Advances in the study of language and thought*. The MIT Press, Cambridge, MA, pp. 47–57.
- Tomasello, M., 2008. *Origins of human communication*. The MIT Press, Cambridge, MA.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., Moll, H., 2005. Understanding and sharing intentions: The origins of cultural cognition. *Behav. Brain Sci.* 28 (5), 675–691.
- Tomasello, M., Conti-Ramsden, G., Ewert, B., 1990. Young children's conversations with their mothers and fathers: Differences in breakdown and repair. *J. Child Lang.* 17 (1), 115–130.
- Tomasello, M., Farrar, M.J., 1986. Joint attention and early language. *Child Dev* 57 (6), 1454–1463.
- Trevarthen, C., Aitken, K.J., 2001. Infant intersubjectivity: Research, theory, and clinical applications. *J. Child Psychol. Psychiatr. Allied Disc.* 42 (1), 3–48.
- Uno, M., 2017. *Developing Question Constructions in Japanese as a First Language: The Roles of Type of Referent and Parental Input*. Unpublished doctoral dissertation. Georgetown University.
- Vabel, A.M., Cohen, A.K., Leonard, S.A., Glymour, M.M., Duarte, C.d.P., Yen, I.H., 2018. Do the health benefits of education vary by sociodemographic subgroup? Differential returns to education and implications for health inequities. *Annals Epidemiol* 28 (11), 757–766. <https://doi.org/10.1016/j.annepidem.2018.08.014> e5.
- VanDam, M., 2018. Cougar Corpus. doi:10.21415/T5WT25. Available at URL <https://homebank.talkbank.org/access/Password/Cougar.html> </Dataset>.
- VanDam, M. [computer software]. 2020. LENA\_Matlab\_audio. Available at URL [https://github.com/HomeBankCode/LENA\\_Matlab\\_audio](https://github.com/HomeBankCode/LENA_Matlab_audio).
- VanDam, M., De Palma, P., 2019. A modular, extensible approach to massive ecologically valid behavioral data. *Behav. Res. Meth.* 51 (4), 1754–1765.
- VanDam, M., and De Palma, P., 2014. Fundamental frequency of child-directed speech using automatic speech recognition. 2014 Joint 7th International Conference on Soft Computing and Intelligent Systems (SCIS) and 15th International Symposium on Advanced Intelligent Systems (ISIS) (pp. 1349-1353). doi: 10.1109/SCIS-ISIS.2014.7044876.
- VanDam, M., Jessup, C., Tully, T., 2016. Mothers' and Fathers' differential talk to daughters and sons with hearing loss. *Proc. Meet. Acoust.* 29 (1), 060012.
- VanDam, M., Silbert, N.H., 2016. Fidelity of automatic speech processing for adult and child talker classifications. *PLoS ONE* 11 (8), e0160588.
- Vigil, D.C., Hodges, J., Klee, T., 2005. Quantity and quality of parental language input to late-talking toddlers during play. *Child Language Teaching and Therapy* 21 (2), 107–122.
- Vygotsky, L., 1978. Interaction between learning and development. *Readings on the development of children* 23 (3), 34–41.
- Vygotsky, L.S., 1986. *Thought and language* (A. Kozulin, trans.). MIT Press, Cambridge, MA.
- Warlaumont, A.S., Richards, J.A., Gilkerson, J., Oller, D.K., 2014. A social feedback loop for speech development and its reduction in autism. *Psych Sci* 25 (7), 1314–1324. <https://doi.org/10.1177/0956797614531023>.
- Warren, S.F., Gilkerson, J., Richards, J.A., Oller, D.K., Xu, D., Yapanel, U., Gray, S., 2010. What automated vocal analysis reveals about the vocal production and language learning environment of young children with autism. *J. Autism Dev. Disord.* 40 (5), 555–569.
- Wells, G., 1979. Describing children's linguistic development at home and at school. *Brit. Educ. Res.* 5 (1), 75–98.
- Wells, G., 1980. Apprenticeship in meaning. In: Nelson, K.E. (Ed.), *Children's Language Volume 2*. Gardner Press, New York, pp. 45–126.
- Wertsch, J.V., 1985. *Vygotsky and the social formation of mind*. Harvard University Press, Cambridge, MA.
- Whiting, B., Edwards, C.P., 1988. A cross-cultural analysis of sex differences in the behavior of children aged 3 through 11. In: Handel, G. (Ed.), *Childhood socialization*. Hawthorne, New York, NY, pp. 281–297.
- Wiener, M., Devoe, S., Rubinow, S., Geller, J., 1972. Nonverbal behavior and nonverbal communication. *Psychol Rev* 79 (3), 185–214.
- Wright, D.B., Carlucci, M.E., 2011. The response order effect: People believe the first person who remembers an event. *Psychon. Bull. Rev.* 18 (4), 805–812.
- Xu, D., Yapanel, U., Gray, S., 2009. Reliability of the LENA Language Environment Analysis System in young children's natural home environment. Boulder. LENA Foundation, CO. Retrieved from: [https://www.lena.org/wp-content/uploads/2016/07/LTR-05-2\\_Reliability.pdf](https://www.lena.org/wp-content/uploads/2016/07/LTR-05-2_Reliability.pdf).
- Yang, F., Heeman, P.A., 2010. Initiative conflicts in task-oriented dialogue. *Computer Speech and Lang* 24 (2), 175–189.
- Yu, Y., Bonawitz, E., Shafto, P., 2019. Pedagogical questions in parent-child conversations. *Child Dev* 90 (1), 147–161.
- Zheng, L., Liu, W., Long, Y., Zhai, Y., Bai, X., Zhou, S., Li, K., Zhang, H., Liu, L., Guo, T., Lu, C., 2020. Affiliative bonding between teachers and students through interpersonal synchronisation in brain activity. *Social Cog. Affect. Neurosci.* 15 (1), 97–109. <https://doi.org/10.1093/scan/nsaa016>.