The Role of Siblings on Infant Language Exposure in Daylong Audio Recordings

Maily Galindo (mailyg2@illinois.edu)
Department of Psychology, 603 E Daniel St
Champaign, IL 61801 USA

Jessica L. Montag
Department of Psychology, 603 E Daniel St
Champaign, IL 61801 USA

Abstract

The language children hear every day is important for language learning. Different sets of speakers may lead to different kinds of speech: Speech directed at a child or speech directed to another adult or a sibling. We aimed to quantify how the presence of a sibling, specifically, affects dynamics of speech in a home. Using day long audio recordings from homes in the United States, we measured the amount of target child-directed and overheard speech in the home of infants in families with and without older siblings. Infants with an older sibling experienced significantly more overheard speech and significantly less speech directed to them, than infants without an older sibling. However, families with and without older siblings did not differ in the total amount of child-directed input, directed to either sibling. These findings suggest a more complicated relationship between overheard speech and language learning.

Keywords: overheard speech; child directed speech; siblings; infants

Introduction

A great deal of research has focused on gaining a better understanding between what is heard or available to a young child (i.e., language input) and their subsequent language learning trajectories. One source of avenue by which researchers have utilized to gain this understanding has been through the analyses of naturalistic recordings of children's auditory input. The ultimate goal of measuring children's naturalistic language input is to better understand the mechanisms by which children learn from their environment by evaluating the plausibility of hypotheses about what aspects of the language environment might be most relevant for learning.

One debate in the field of language acquisition is the relative role of child-directed versus overheard speech in language learning. Child-directed speech is speech that is produced for, and directed at a particular child while overheard speech is not intended for that particular child, but available for the child to overhear.

Many findings overwhelmingly suggest that child-directed speech is associated with language learning outcomes while overheard speech is unrelated to these outcomes. Shneidman and others (2013) revealed very little overheard speech in proportion to child-directed speech amongst US families. Other research has observed a wide range of overheard and child-directed speech use in Spanish-speaking US families (Weisleder & Fernald, 2013). However, in both datasets only child-directed input significantly predicted children's vocabulary scores (Weisleder & Fernald, 2013; Shneidman, Arroyo, Levine, & Goldin-Meadow, 2013). No significant relations between overheard speech and children's language outcomes have been observed through these data.

The finding that child-directed speech is more useful for language learning is consistent with other effects that find that speech specifically tailored to a child may lead to more learning. For example, researchers observe more learning from language in which a caregiver is intentionally trying to communicate information to a child (Baldwin, 1993), when a child interacts with a live human rather than a video (Kuhl, Tsao & Liu, 2003; Anderson & Pempek, 2005; O'Doherty, Troseth, Shimpi, Goldenberg, Akhtar, & Saylor, 2011), or when a caregiver is sensitive to where a child's attention is directed (Yu Suanda & Smith, 2019; Adamson, Bakeman, Suma, & Robbins, 2019). More broadly, input that is sensitive and warm, or *responsive*, is positively associated with development across various domains, including language (Tamis-LeMonda, Kuchirko, & Song, 2014; Madigan, Prime, Graham, Rodrigues, Anderson, Khoury, & Jenkins, 2019). Taken together, these findings suggest that caregiver and child attention are important factors in learning, perhaps accounting for the advantage observed for child-directed speech.

Despite correlational findings that suggests a minimal role for overhead speech, there are two reasons to be skeptical. First, experimental studies show that by 18 months of age, children learn words from overheard speech just as well as from child-directed speech. (Akhtar, Jipson, & Callanan, 2001; Shneidman, Buresh, Shimpi, Knight-Schwarz, & Woodward, 2009; Foushee, Srinivasan, & Xu, 2021). This novel word learning persists even in more "naturalistic" conditions such as when the novel object is not explicitly labeled or when the child is given other toys as a form of distraction (Akhtar, 2005). Young children show robust word learning through both child-directed and overheard speech in laboratory studies.

Second, the proportion of child-available input that is child-directed varies wildly across cultures, and it would seem unexpected that human language learning abilities are not more flexible considering this variability. Globally, there are communities in which it is uncommon for adults to speak directly to young children until they are old enough to

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

produce speech themselves (Shneidman & Goldin-Meadow, 2012; De León, 1998). For example, small percentages of child-directed speech have been found in communities such as the Tseltal and Yucatec Mayan villages (i.e., ~15-20% of child-directed input; Shneidman & Goldin-Meadow 2012; Casillas, Brown, & Levinson, 2020). Other communities like those in Bolivia have been observed to produce even smaller amounts of child-directed input (i.e., ~9%; Cristia, Dupoux, Gurven, & Stieglitz, 2019). Contrastingly, villages in France and some in Southern Africa produce very high amounts of child-directed input (87%-94%) that even outpass that of which has been shown for US samples (65%; Shneidman et.al., 2013), except for one Southern African village (i.e., ~27%; Loukatou, Scaff, Demuth, Cristia, & Havron, 2022).

Although there does seem to be a drastic difference in child-directed speech input in other communities, young children, specifically in Mayan villages, produce some canonical babbles, single words, and multiword utterances at comparable ages of that of children in the US (Shneidman & Goldin-Meadow 2012; Cristia et.al., 2019; Casillas et.al., 2020; Oshima-Takane, Goodz, & Derevensky, 1996). However, research from these communities on children's language outcomes still do not to find correlational relationships between overheard speech input and later language outcomes (Shneidman & Goldin-Meadow, 2012; Casillas et.al., 2020). Child-directed speech consistently emerges as a critical factor in supporting early language learning. However, given the highly variable nature of the language input and the comparable developmental trajectories, there may be unanswered questions about the nature of the child-directed and overheard speech that children hear, which may help us interpret these seemingly contrasting results.

An important issue that arises in the cross-cultural data is the potential role of the family and social structure in driving the sources of language that children hear. Many of the samples that exhibit lower rates of child-directed speech tend to occur in social contexts in which infants are surrounded by many other children and adults. The presence of more children and adults may be associated with lower rates of child-directed speech (Shneidman & Goldin-Meadow, 2012; Cristia et.al., 2019), suggesting one reason for different rates of speech across cultures. Thus, there may be logical consequences of different home or family structures.

The potential role of a child's social environment in driving the proportion of child-directed speech in a child's input leads to an opportunity to evaluate hypotheses regarding the potential utility of overheard speech to language learning. Siblings are also an interesting factor as it relates to the social environment and language learning. The presence of older siblings in the home has been associated with less child-directed speech (Hoff-Ginsberg, 1998; Oshima-Takame & Robbins, 2003; Laing & Bergelson, 2024; Jones & Adamson 1987). If having a sibling indeed leads to different amounts of child-directed speech, children with siblings should show different language learning trajectories compared to those without siblings. That is, a hypothesis that suggests that children primarily learn language from speech that is child-directed, should predict sizable birth order effects in language learning trajectories. In the absence of large birth order effects, we can conclude that either 1) the presence of an older sibling or siblings is not associated with less child-directed speech or 2) overheard speech may be more useful for learning outcomes than is often assumed.

Investigations of birth order effects from large, national, datasets, show there are statistically detectable but very small effects of siblings on language skills. Having an older sibling is negatively associated with language skills (Berglund, Eriksson, & Westerlund, 2005; Havron, Ramus, Heude, Forhan, Cristia, & Peyre, 2019; Peyre, Bernard, Hoertel, Forhan, Charles, & De Agostini, 2016; Jakiela, Ozier, Fernald, & Knauer 2020; Zambrana, Vollrath, Jacobsson, Sengpiel, & Ystrom, 2012) but these effects are typically a fraction of the size of gender effects on early language skills (Berglund et al., 2005; Zambrana et. al., 2012). Though we see consistent evidence for negative effects of older siblings on language outcomes, these effects show often miniscule effect sizes and interact with many other aspects of family life such that interpretation of sibling effects is complicated.

Perhaps the absence of substantial sibling effects means that there are in fact only small differences in the amount of child-directed speech that firstborn and later-born children receive. However, if there are indeed large differences in birth order with respect to amount of child-directed speech, we may conclude that other types of speech, including overheard speech are important sources of input for language learning given that later born children still exhibit typical language developmental norms. To better evaluate these potential explanations, we need to gather more data about how child-directed and overheard speech may vary with birth order.

We investigated differences in proportions of child-directed speech in families with and without siblings. We found that children with an older sibling received about half the amount of target-child-directed speech as children without siblings, but overall amount of child-directed speech (directed at either sibling) was similar. We describe how individuals in the home have implications for who produces utterances and who they are addressing. We also discuss implications for describing the many varieties of overhead speech in homes and the potential roles of child-directed and overhead speech in language learning.

Methods

Participants

Data was obtained from the online TalkBank-HomeBank, online repository (Fausey & Mendoza, 2018; FauseyTrio). Audio recordings were from 26 families in the Eugene and Springfield areas the state of Oregon. Each recording was at least 8 hours (range: 8- 19 hrs., M= 13.43 hrs., SD= 2.08). Target child (i.e., child who wore the audio recording device) ages ranged from 6 to 12 months of age. In half of the families, an older child also resided in the home (n= 13) and

in half of the families the target child was the only child in the home. All families spoke English as the main language. For each target child in the with-siblings group, there was an age matched target child in the without-siblings group, yielding similar mean age and age distributions in each group (with siblings: M= 9.63 mo., SD= 1.40; 10 girls, 3 boys; without-siblings: M= 9.16 mo., SD= 1.87; 5 girls, 8 boys). Each family had three audio recordings completed within one week, with the exception of two families who only had two recordings. The total amount was 76 daylong audio recordings.

Most families in the with-siblings group had one older child in the home (4 boys; 7 girls), except for two families that had two older children (2 boys; 1 boy, 1 girl).

Data collection and coding

Annotating all utterances in a daylong audio recording would be prohibitively time consuming. Our solution was to sample utterances from the daylong recording and annotate a subset that could be used to estimate the content of the entire day. Previous literature suggests that sampling, more, smaller intervals result in more accurate estimations of an entire audio recording (Marasli & Montag, 2023; Ramirez & Hippe, 2024; Scaff, Casillas, Stieglitz, & Cristia, 2024). Informed by this literature, from each daylong recording we sampled 240 30-second intervals time for a total of 2 hours for each recording.

Procedures for each daylong recording followed a consistent protocol. First, all silence times longer than 5 minutes were noted manually and taken out of the total time used for sampling. Thus, total hours of "speaking time" (audio without silence times) across all recordings averaged out to about 7 hours (SD= 2.22 hrs.). From there, 240 30-second random intervals were sampled (with replacement) from each recording using a sampling algorithm implemented in Python. The algorithm randomly selected a time in the audio file and 30 seconds from that staring time was selected as the sampled interval. All but two participating families had a total of 6 hours, or 720 30-s segments sampled across all three daylong recordings all of which were manually coded. The other two families had only two daylong recordings from which a total of 4 hours per family was sampled.

We manually annotated the speaker and addressees of all utterances spoken within each 30-s interval. Utterances were defined as a string of speech spoken without pause. A pause was defined as a silence longer than one second.

For each speaker (except the target child), we noted the total utterances and to whom each utterance addressed. All vocalizations, except for laughs, sighs, or sneezes were included in the total counts. Imitations of a vocalization (e.g., parent imitates baby by saying "baba") were included. If the coder could not discern when an utterance started or ended or if the speech was too faint, then it was not included. All speakers in the home who spoke more than a 5 or more utterances for the entire file were included as a speaker and were coded as FA (female adult), MA (male adult), FC (female child) or MC (male child), and FAE (female adult electronic) or MAE (male adult electronic). Electronic speakers were generally speakers whose voices were heard over a phone. Electronic speakers do not include electronic media, such as speech from a toy or television which we coded separately.

In addition to the speaker, each utterance was coded for its addressee. We used multiple sources of information to infer an utterance's addressee from the audio recording. First, we used contextual information such as what was said, the use of someone's name, the topic of conversation, and who responded. Second, we used recognizable acoustic markers of child-directed speech such as high-pitched voices, sing-song-like intonation, simplified speech, and diminutives (e.g., "doggy" vs. "dog"). Addressees included *target child*, *other child* (or children), *adult*(s), *both* (i.e., child and adult), *other* (i.e., talking to oneself or pets), and *unsure*.

Inter-coder reliability was assessed on 4 different files for all 240 intervals. The average inter-coder reliability between coders was .95 indicating high agreement on utterance counts within each interval and for each addressee under each speaker in the recording. Reliability was computed through a percent agreement calculation in which the proportion of number of agreed items and total number of items per speaker and addressee (e.g., female speaker 1, target child; female speaker 1, adult and so forth) were taken between two coders. Electronic media coding was not found to be reliable (<.80) among the coders and thus was not included in any statistical analyses.

For simplicity, when we discuss speakers of utterances, we designate the two adults who produced the most utterances as Adult 1 and Adult 2. These adults accounted for about 90% of the total adult speech data in both groups. This allows us capture broad differences across families in adult versus child produced and adult versus child directed speech without getting lost in speaker-identity nuances. Adult One included 74 moms and 2 grandmothers. Adult Two included Mom (n= 1), Dad (n= 65), Aunt (n= 3), and Grandma (n= 6). All raw speaker and addressee counts will be available on OSF upon publication.

Results

The analyses were conducted on the aggregated counts of utterances across multiple variables for all 240 30-s intervals per recording using R (version 4.3.1). Averages per group were computed by taking averages within families and once again across all families. Two-tailed statistical tests were used with $\alpha = 0.05$. P-values were adjusted for multiple testing across 15 variables using Bonferroni's correction.

By Addressee

Our primary research question was differences in language directed towards an infant with the presence or absence of an older sibling. We compared the amount of speech that was target child-directed or directed to *other child*(ren), across families with siblings and without siblings.

The mean number utterances addressed to each individual in a daylong recording, averaged across all families and the associated standard deviations, and absolute differences are noted in Table 1. Important differences across families with and without an older sibling emerged. An independent sample t-test showed that families with an older sibling directed significantly fewer utterances to the target child compared to families without an older sibling (411.33 vs. 745; $t(24)= -4.35$; adj. $p < .001$). However, families with and without siblings produced equal amounts of all child-directed speech, that is, the sum of *target child* directed, and *older child* directed speech (TC + OC; 833.38 vs. 749.38; $t(24)= 0.83$; adj. $p > .99$). The decrease in child-directed speech for children with siblings seems to be due to large amounts of speech addressed instead to the other child because the total speech directed to all children in the home was similar in families with and without siblings. No significant differences were found for amount of speech addressed to other individuals: *adults*, *both*, *other*, and *unsure*.

Table 1: Mean Utterance Counts per Addressee

Addressees	With Siblings		Without Siblings		Absolute Difference	
	Mean	SD	Mean	SD	Mean	*Adj.* p-value
Target child	411.3	172.8	745.0	216.3	-333.7	<.001
Other child	439.1	166.7	4.4	11.6	434.7	<.001
TC + OC	*833.4*	*295.0*	*749.4*	*212.6*	*84.0*	*>.99*
Adult	602.5	174.6	425.4	159.7	177.1	.189
Both A + C	7.4	6.2	4.2	4.7	3.2	>.99
Other	79.2	50.8	38.9	22.7	40.4	.276
Unsure	6.0	5.0	3.6	3.1	2.4	>.99

Figures 1A and 1B show the distributions of the proportion of captured speech across families that is target-child directed (1A) and any-child directed (1B) speech in families with and without a sibling. These figures represent the density, which indicates the relative frequency of utterance counts within each group across the observed range, normalized so that the total area under each curve equals one. For target-child directed speech, we see two distinct but partially overlapping distributions. However, when collapsing across any child-directed speech, then the distributions are overlapping.

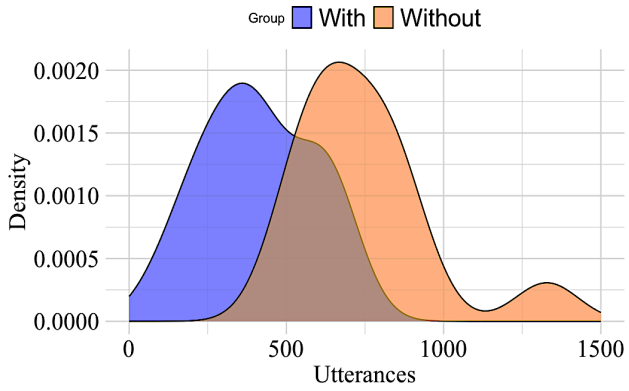

Figure 1A: Target Child Directed Utterances

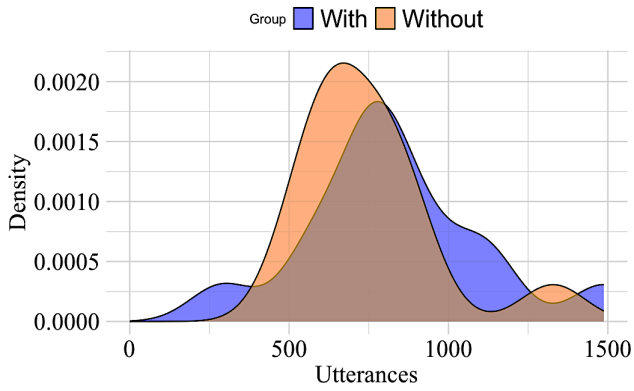

Figure 1B: Any Child Directed Utterances

Our second goal was to more generally measure difference in the amount and types of overheard speech available for infants with and without siblings. First, we found no significant difference in the total amount of speech spoken across the two groups, directed at any individual ($p > 0.05$; see Table 2). However, as we expected to see, there was a significant difference in the amount of overheard speech available to the infant ($p < .001$). Overheard speech consisted of 74% of total input for infants with older siblings compared to 39% of total input for those without siblings.

Table 2: Mean Utterance Counts for Overheard and Total Speech

Variables	With Siblings		Without Siblings		Absolute Difference	
	Mean	SD	Mean	SD	Mean	*Adj.* p-value
Total overheard	1134.2	268.6	476.4	165.0	640.7	<.001
Total speech	1528.5	344.3	1221.4	277.5	307.1	.298

By Speaker

Figure 2 highlights that children with siblings are exposed to more overheard speech than children without siblings. To better understand these differences in the distribution of speech, we gathered data on the proportions of speech spoken by a particular speaker and to whom that speech was addressed to. The proportions shown in Figure 2 represent the amount of speech directed by each speaker to a specific addressee, calculated as a proportion of the speaker's total speech. These counts were first averaged within individual families and then further averaged across all families in their respective group. Note, the *Other* speakers account only for 6% and 9%, respectively, of the total speech.

The adult speech in homes with and without siblings shows noteworthy differences. In families without siblings, over half of the adult total speech was directed to the target child (65% and 51%). The rest of the speech consisted mostly of adult directed speech (30% and 46%), typically caregivers talking to each other. In contrast, in families with siblings, adults address smaller proportions of speech to other adults in the home (19% and 28%), even smaller than the amount of speech directed to older siblings (41% and 39%). In our sample, for infants without an older sibling 39% of speech is overheard and all of it is adult-directed. In contrast for infants with an older sibling, 75% of speech is overheard and 53% of that overheard speech is adult-directed. In families with siblings, 61% of adult directed speech is produced by the sibling and 37% is produced by other adults.

Interestingly, in our sample, sibling speech was rarely directed toward the target infant. About 75% of the older sibling speech was addressed to an adult in the home and about 13% to the *other* addressee category (across most audio recordings, this addressee included mostly speech spoken to oneself). Very little of the older child's speech was addressed to their infant sibling (6%). Note, the means for the older child speaker were not included in the without-sibling table because the average spoken speech for that speaker in the without-siblings group was less than 1% of total speech produced.

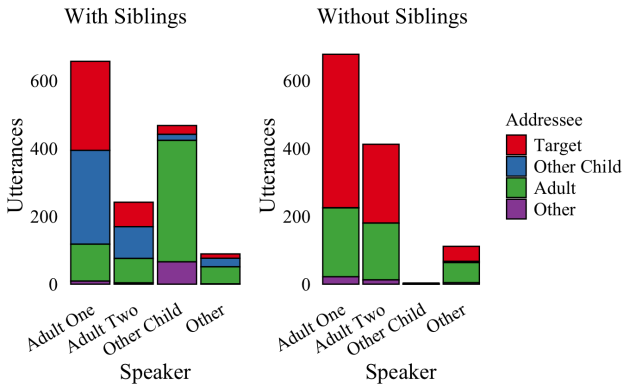

Figure 2: Proportions of Speech by Speaker per Addressee

Discussion

Using daylong audio recordings, this study provides new insights into an infant's language learning environment. The findings reveal distinct differences in speech dynamics between families with and without older siblings. In families with older children, infants received significantly less directed speech, as adult speech was divided between the infant and the older child. Consequently, these families demonstrated higher proportions of overheard speech, particularly from adult-to-adult and child-to-adult interactions. In contrast, families without older siblings showed higher proportions of target-child-directed speech, with adults focusing more of their speech on the infant. These results align with prior research highlighting the influence of sibling presence on the distribution of language input, providing further evidence for the shift in language dynamics when multiple children are present in the household even in industrialized societies like the US.

Notably, our study revealed no difference in the overall input an infant was experiencing in the home. Families with or without an older child, were still producing the same amount of speech. Instead, the differences observed were related to the quantity of overheard speech and target-child-directed speech. Specifically, infants in homes with siblings were exposed to significantly more overheard speech, resulting in reduced target-child-directed input and increased exposure to speech directed at an older child and adults. In other words, main findings reveal infants in families without older children to be exposed to significantly more directed input compared to families with multiple children. This is a replication of previous studies that have found positive relations between overheard speech and number of speakers in the home specifically, the presence of an older sibling (Shneidman et.al., 2013; Casillas et al., 2020).

Our main finding of substantial differences in the amounts of target-child-directed speech addressed to infants with and without siblings contrasts with the lack of differences reported in some prior studies (Laing & Bergelson, 2024; Shneidman et al., 2013). Laing and Bergelson (2024) only observed a significant decrease in infant-directed input for families with two or more older siblings. This discrepancy may stem from differences in methodology. Our study utilized daylong audio recordings with a sampling approach that allowed us to capture a more representative picture of speech patterns across entire days, compared to studies relying on shorter observation windows.

Additionally, by manually annotating all speech and distinguishing between infant-directed and older-child-directed input, our findings reveal that the reduction in infant-directed speech in sibling households diminishes when considering the totality of child-directed speech, which includes both infant-directed and older-child-directed speech In homes with multiple children, child-directed speech is distributed between the siblings. Specifically, infants with older siblings in our sample received approximately half the amount of directed language exposure compared to infants without older siblings. This finding supports the theory that

as the number of children in the home increases, parental attention and resources are divided (Havron et al., 2019; Hoff, 2006). Similarly, Jones and Adamson (1987) suggested that the presence of a sibling reduces the amount of directed speech each child receives, as caregivers maintain consistent levels of speech production regardless of the number of children in the household.

To our knowledge, no effect this large has yet to be borne out in the literature within an industrialized society like the US. Our findings suggest that there may in fact be substantial siblings effects on child-directed and overheard speech, aligning US populations more closely with previous cross-cultural work where directed input has been observed to be significantly lower (Cristia et.al., 2019; Shneidman & Goldin-Meadow, 2012; Casillas et.al., 2020). An important issue that arises in these cross-cultural data is the potential role of family and social structure in shaping the sources of language that children hear. These cross-cultural differences may reflect logical consequences of varying family structures, where the presence of more children and adults often results in reduced child-directed speech (Shneidman & Goldin-Meadow, 2012; Cristia et al., 2019). Similarly, our findings suggest that sibling effects in US homes may parallel some of these dynamics, with caregiver speech divided across multiple children.

The overall quantity of language input in US homes remains consistent across families with and without multiple children. The nature of this input varies significantly depending on the presence of siblings. In homes without siblings, infants hear less overheard speech overall, and the speech they do overhear tends to be primarily adult-directed. In contrast, infants in homes with older siblings are exposed to a greater amount of overheard speech, much of which is child-directed. The set of speakers around a child may have implications for the language that child hears, and possibly even how the child attends to and learns from that language.

This distinction likely impacts how infants process and learn from overheard speech, as the content and complexity of language differs depending on the speaker. Infants with siblings may be influenced by simpler, child-directed input spoken to other children in the home. These differences in exposure could cascade over time, shaping what infants pay attention and thus their subsequent broader language acquisition

Theories have argued child directed speech to be important for language learning possibly due to its simplicity and its ability to capture a young child's attention (Cartmill et.al., 2013). In fact, in most overheard word learning paradigms in which young children learn a novel word, the overheard speech used in the experiments are simple and follow that of what would be characterized as child directed speech (Akhtar, 2005; Akhtar et. al., 2001; Ma et.al., 2011). Thus, infants could be utilizing overheard speech to gain knowledge if it follows the simplicity of child-directed speech as well as including conceptually engaging topics (e.g., relating to food, toys, play, etc.) like that of what is observed in families with older children.

Decades of research have debated the role of attention in language acquisition, particularly how it intersects with CDS. Child-directed interactions often engage joint attention, where both the caregiver and child focus on the same object or activity, which is thought to enhance language learning (Ma et al., 2011; Tamis-LeMonda et al., 2014). Studies have consistently shown that children who frequently engage in joint attention with caregivers tend to have larger vocabularies (Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998; Tomasello & Farrar, 1986; Tomasello & Todd, 1983). However, even in the absence of joint attention, children's sustained attention during an interaction—whether directed at them or merely observed—has been linked to successful word learning (Shneidman et al., 2009; Macroy-Higgins & Montemarano, 2016). This suggests that attention itself, rather than joint attention specifically, may be a critical factor. Infants exposed to high levels of overheard speech, especially those in homes with older siblings, might engage in sustained attention, enabling them to learn from overheard interactions in everyday life.

We hypothesize that children do not necessarily require full joint attention to learn from language but instead need to be paying attention to the interaction, even as passive observers. This theory can allow for certain types of overheard speech to play a more significant role in language development than previously assumed. For example, overheard speech that incorporates elements of child-directed speech, such as simplicity or engaging content, may still effectively support learning. Understanding what makes speech engaging for children and prompts them to tune in will be essential for future research, offering valuable insights into the mechanisms of early language acquisition.

To investigate these hypotheses, future research using naturalistic data will have to address the potential attentional processes that children and others use as it relates to word learning through overheard speech. Furthermore, our sample of families for each group is relatively small that all reside in a similar area. Although, overall, we had a sizable amount of data. Yet, with no additional demographic information, it is likely that some families included a stay-at-home parent or caregiver given that many audio recordings included a mom that was present throughout an entire day. This may mean that the sample is not fully representative of the diversity in U.S. family structures, backgrounds, or socioeconomic statuses. Finally, because we did not have information on the siblings' ages, we are not able to consider that as a factor for speech input. Sibling age gaps can play a critical role in not only the quantity of speech input but also the qualities that language holds.

References

Adamson, L. B., Bakeman, R., Suma, K., & Robins, D. L. (2019). An expanded view of joint attention: Skill, engagement, and language in typical development and autism. *Child Development*, *90*(1), e1-e18.

Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science*, *8*(2),199-209.

Akhtar, N., Jipson, J., & Callanan, M. A. (2001). Learning words through overhearing. *Child Development*, *72*(2), 416-430.

Anderson, D. R., & Pempek, T. A. (2005). Television and very young children. *American Behavioral Scientist, 48*(5), 505-522.

Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Development*, *67*(6), 3135-3153.

Berglund, E. V. A., Eriksson, M., & Westerlund, M. (2005). Communicative skills in relation to gender, birth order, childcare and socioeconomic status in 18-month-old children. *Scandinavian Journal of Psychology, 46*(6), 485-491

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, i-174.

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

Cristia, A., Dupoux, E., Gurven, M., & Stieglitz, J. (2019). Child-directed speech is infrequent in a forager-farmer population: A time allocation study. *Child Development*, *90*(3), 759-773.

De León, L. (1998). The emergent participant: Interactive patterns in the socialization of Tzotzil (Mayan) infants. *Journal of Linguistic Anthropology*, *8*(2), 131-161.

Foushee, R., Srinivasan, M., & Xu, F. (2021). Self-directed learning by preschoolers in a naturalistic overhearing context. *Cognition*, *206*, 104415.

Havron, N., Ramus, F., Heude, B., Forhan, A., Cristia, A., Peyre, H., & EDEN Mother-Child Cohort Study Group. (2019). The effect of older siblings on language development as a function of age difference and sex. *Psychological Science, 30*(9), 1333-1343.

Hoff-Ginsberg, E. (1998). The relation of birth order and socioeconomic status to children's language experience and language development. *Applied Psycholinguistics*, *19*(4), 603-629.

Jakiela, P., Ozier, O., Fernald, L. C., & Knauer, H. (2020). Big sisters.

Jones, C. P., & Adamson, L. B. (1987). Language use in mother-child and mother-child-sibling interactions. *Child Development*, 356-366.

Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, 100(*15), 9096-9101.

Laing, C., & Bergelson, E. (2024). Analyzing the effect of sibling number on input and output in the first 18 months. *Infancy*, *29*(2), 175-195.

Loukatou, G., Scaff, C., Demuth, K., Cristia, A., & Havron, N. (2022). Child-directed and overheard input from different speakers in two distinct cultures. *Journal of Child Language, 49*(6), 1173-1192.

Macroy-Higgins, M., & Montemarano, E. A. (2016). Attention and word learning in toddlers who are late talkers. *Journal of Child Language*, *43*(5), 1020-1037.

Madigan, S., Prime, H., Graham, S. A., Rodrigues, M., Anderson, N., Khoury, J., & Jenkins, J. M. (2019). Parenting behavior and child language: A meta-analysis. *Pediatrics*, *144*(4).

Marasli, Z., & Montag, J. L. (2023). Optimizing random sampling of daylong audio. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 45, No. 45).

O'Doherty, K., Troseth, G. L., Shimpi, P. M., Goldenberg, E., Akhtar, N., & Saylor, M. M. (2011). Third-party social interaction and word learning from video. *Child Development*, *82*(3), 902-915.

Oshima-Takane, Y., Goodz, E., & Derevensky, J. L. (1996). Birth order effects on early language development: Do secondborn children learn from overheard speech?. *Child Development*, *67*(2), 621-634.

Oshima-Takane, Y., & Robbins, M. (2003). Linguistic environment of secondborn children. *First Language*, *23*(1), 21-40.

Peyre, H., Bernard, J. Y., Hoertel, N., Forhan, A., Charles, M. A., De Agostini, M., ... & EDEN Mother-Child Cohort Study Group. (2016). Differential effects of factors influencing cognitive development at the age of 5-to-6 years. *Cognitive Development, 40*, 152-162.

Ramírez, N. F., & Hippe, D. S. (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.

Scaff, C., Casillas, M., Stieglitz, J., & Cristia, A. (2024). Characterization of children's verbal input in a forager-farmer population using long-form audio recordings and diverse input definitions. *Infancy*, *29*(2), 196-215.

Shneidman, L. A., Arroyo, M. E., Levine, S. C., & Goldin-Meadow, S. (2013). What counts as effective input for word learning?. *Journal of Child Language*, *40*(3), 672-686.

Shneidman, L. A., Buresh, J. S., Shimpi, P. M., Knight-Schwarz, J., & Woodward, A. L. (2009). Social experience, social attention and word learning in an overhearing paradigm. *Language Learning and Development*, *5*(4), 266-281.

Shneidman, L. A., & Goldin-Meadow, S. (2012). Language input and acquisition in a Mayan village: How important is directed speech?. *Developmental Science*, *15*(5), 659-673.

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness?. *Current Directions in Psychological Science*, *23*(2), 121-126.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, 1454-1463.

Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, *4*(12), 197-211.

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

Yu, C., Suanda, S. H., & Smith, L. B. (2019). Infant sustained attention but not joint attention to objects at 9 months predicts vocabulary at 12 and 15 months. *Developmental Science*, *22*(1), e12735.

Zambrana, I. M., Vollrath, M. E., Jacobsson, B., Sengpiel, V., & Ystrom, E. (2021). Preterm birth and risk for language delays before school entry: A sibling-control study. *Development and Psychopathology, 33*(1), 47-52.