Language Input to Blind Infants/Toddlers

Erin Campbell[1], Lillianna Righter[1], Genia Lukin[1], & Elika Bergelson[1]

[1] Department of Psychology & Neuroscience, Duke University, Durham, NC

Conflicts of Interest: The authors have no conflicts of interest to report. **Funding**: This work was supported by the National Science Foundation CAREER grant (BCS-1844710) to EB and Graduate Research Fellowship (2019274952) to EC.

Language Input to Blind Infants/Toddlers

Introduction

The early language skills of blind children are highly variable (**campbell_inprep?**), with some blind children demonstrating age-appropriate vocabulary from the earliest stages of language learning (Landau & Gleitman, 1985; **bigelow_early_1987?**), while others experience large and persistent language delays (**CITE?**). Canonically, blind adults become competent speakers of their language and are even reported to have faster language processing skills than their sighted peers (**roder_event-related_2000?**; **roder_semantic_2003?**). The causes of this variability and the later ability to "catch up" remain poorly understood. In particular, the higher incidence of severe language delays in blind children yields questions about the process of language development in the absence of visual perception: what makes the language learning problem different and apparently more difficult for the blind child? There are multiple possible contributors, including characteristics of the child (e.g., visual acuity, comorbid conditions, gender) as well as characteristics of the environment (e.g., access to early intervention services; school setting; caretakers tailoring interactions to their child's sensory access). Here, we explore the characteristics of the language environment of blind children as it compares to the language environment of their sighted peers. In doing so, we begin to narrow down the role that visual input plays in language development, among all other factors.

Among both typically-developing children and children with developmental differences, language input can predict variability in language outcomes (Anderson, Graham, Prime, Jenkins, & Madigan, 2021, 2021; Gilkerson et al., 2018; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Rowe, 2008, 2012). There are many ways to operationalize language input, that tend to be grouped into **quantity of language input** and **input characteristics** (often discussed as **quality of language input**, c.f. MacLeod and Demers (2023)). Quantity of language input can be broadly construed as the number of words or utterances a child is exposed to. At a coarse level, children who are exposed to more speech (or sign, **watkins_deaf_1998?**) tend to have better language outcomes (Anderson et al., 2021; Gilkerson et al., 2018; Huttenlocher et al., 1991; Rowe, 2008). However, if only the *amount* of language exposure mattered, then infants should be able to sit in front of the television all day and become fluent language users. Yet young children struggle to learn language from video alone (e.g., Roseberry, Hirsh-Pasek, & Golinkoff, 2014 May-Jun).

The specific characteristics of that language input are perhaps even more important (Hirsh-Pasek et al., 2015; Rowe, 2012), although it is somewhat trickier to turn the qualitative characteristics of language input into operationalizable properties. In this analysis, we move away from describing these linguistic characteristics as “quality” measures[^1]. Rowe and Snow (Rowe & Snow, 2020) divide this space into three dimensions of language input: interactive features (e.g., parent responsiveness, speech directed *to* child vs. overheard; conversational turn-taking), linguistic features (e.g., lexical diversity, grammatical complexity), and conceptual features (e.g., topic diversity). These environmental features at various stages interact with the child’s own cognitive, linguistic, and conceptual abilities.

[^1] In the past, the directionality of these terms has favored the types of language used by white and abled groups as immutable universal standards, thereby framing racialized and disabled peoples’ language as deficit and “low quality” by nature. Describing a singular source of input variation as “high quality” ignores the sociocultural variation of talk styles, and the presence of many rich sources of information that children can learn from (MacLeod & Demers, 2023).

An important social feature of the language environment is the amount of interactivity in parent-child communication. Prior literature reports that back-and-forth communicative exchanges (also known as conversational turns) between caregivers and children predict better language learning across infancy (Donnellan, Bannard, McGillion, Slocombe, & Matthews, 2020; Goldstein & Schwade, 2008) and toddlerhood (Hirsh-Pasek et al., 2015; Romeo et al., 2018), indicating that parents' active response to their children's actions and utterances supports their learning. Adults' attunement to children's non-linguistic cues of attention and interest, like pointing or eye gaze, also contributes to interactivity. In infancy, words heard in contexts where the adult and child share joint attention are more likely to be learned (Tomasello & Farrar, 1986; **lucca2017?**). Parents' interaction with their child and the world around them ties together the linguistic and conceptual characteristics of the language input, to which we turn next.

Two commonly-analyzed linguistic features are lexical diversity (often measured as type/token ratio) and syntactic complexity. In accounts of the development of sighted children, lexical diversity of language input seems to exert different effects as children get older. In early infancy, children who are exposed to more repetitions (and therefore less lexical diversity) at 7 months have higher vocabulary at age 2 (Newman, Rowe, & Bernstein Ratner, 2016). This relationship later flips: toddlers who are exposed to greater diversity of words in their language input tend to have larger vocabulary scores (Anderson et al., 2021; Hsu, Hadley, & Rispoli, 2017; Huttenlocher et al., 2010; Rowe, 2012; Weizman & Snow, 2001). Lexical diversity is intertwined with input quantity: parents who talk more also tend to provide more lexical diversity (Hoff & Naigles, 2002 Mar-Apr). Likewise, the diversity of syntactic constructions in parental language input is associated both with children's vocabulary growth and structure diversity in their own productions (De Villiers, 1985; Hadley et al., 2017; Hoff, 2003 Sep-Oct; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002; Huttenlocher et al., 2010; Naigles & Hoff-Ginsberg, 1998).

The conceptual dimension of language input aims to capture the extent to which the language signal maps onto present objects and co-occurring events. As with the other dimensions, the influence of the conceptual content of language input may shift across developmental time: as children develop, their ability to represent abstract, displaced, decontextualized referents improves [CITE]. For example, young infants are more likely to learn a new word when the referent is perceptually salient, dominating their field of view (Yu & Smith, 2012). Parents responding to a child's point and labeling the object of interest might boost learning in that instance (Lucca & Wilbourn, 2018). By contrast, *displaced* language use– that is, talking about past, future, or hypothetical events, or people and items that are not currently present in the environment– may be beneficial at later stages of development (**rowe_decontextualized_2013?**). Indeed, greater decontextualized language use in speech to toddlers predicts kindergarten vocabulary (Rowe, 2012), children's own decontextualized language use (Demir, Rowe, Heller, Goldin-Meadow, & Levine, 2015), and academic achievement in adolescence (Uccelli, Demir-Lira, Rowe, Levine, & Goldin-Meadow, 2019). Decontextualized language may support language learning because it provides an opportunity to discuss a broader range of topics and reflects typical adult language usage, which is often abstract (**CITE?**). It also provides the opportunity for more lexical and syntactic diversity.

From this review, it appears that sighted children learn about the world and language simultaneously from many sources, including sensory perception, linguistic input, and conceptual and social knowledge. For blind children, however, language input may constitute a greater proportion of the available clues for learning than for sighted children; in the absence of visual input, language is an important source of information about the world (E. E. Campbell & Bergelson, 2022). Syntactic structure provides cues to word meaning that may be lost without visual cues, such as the relationship between two entities that aren't within reach (**gleitman1990?**). In our review so far, we have presented a pattern wherein the features of the input that are most helpful for language learning change over the course of children's development: early on, many of these cues require visual access,such as parental gaze, shared visual attention, pointing to remote object and the presence of salient objects in the visual field. Only later in development do the handholds to language learning become more abstract. This may be part of the reason why language delays are common in blind toddlers, but often resolved in older childhood [CITE]. If direct sensory access is the key to unlocking the meaning of early words, it may take longer to gain enough environmental experience to make early language learning strides– that is, it may take longer in infancy to build a "semantic seed" (**babineau2021?**; **babineau2022?**). By hypothesis, once this initial seed of linguistic knowledge is acquired, blind children are able to use more abstract and linguistic features as cues, and learning proceeds rapidly (E. E. Campbell & Bergelson, 2022). Nevertheless, we cannot assume that access to visual experience is the *only* difference in the language learning experiences for blind and sighted children. The language input itself may very well differ for blind children relative to sighted children, for a variety of reasons.

First, speakers regularly tailor input to communicate efficiently with the listener (**grice1975?**). Parents are sensitive to their child's developmental level and tune language input accordingly (Snow, 1972; Vygotsky & Cole, 1978). Child-directed speech is one example–whereby parents speak to young children with exaggerated prosody, slower speech rate, and increased vowel clarity (Bernstein Ratner, 1984; Fernald, 1989), which is in some cases helpful to the young language learner (Thiessen, Hill, & Saffran, 2005). Parents show increased alignment (a tendency to re-use the conversation partner's expressions) for younger children, which decreases as children get older (Yurovsky, Doyle, & Frank, 2016). When interacting with infants and toddlers, parents repeat words more often than when interacting with older children or adults (Snow, 1972). Communicative tailoring is also common in language input to children with disabilities, who tend to receive simplified, more directive language input, and less interactive input compared to typically-developing children (Yoshinaga-Itano, Sedey, Mason, Wiggin, & Chung, 2020; **dirks2019?**).

In addition to tailoring communication to children's developmental level, speakers also adjust their conversation in accordance with the conversation partner's sensory access (Gergle, Kraut, & Fussell, 2004; Grigoroglou, Edu, & Papafragou, 2016). In a noisy environment, speakers will adapt the acoustic-phonetic features of their speech with the intent to make it easier for their interlocutor to understand them (**hazan_acoustic-phonetic_2011?**), which demonstrates sensitivity to even temporary sensory conditions of their conversation partner. When describing scenes, speakers aim to provide the information their listeners lack but avoid redundant visual description (Ostarek, Paridon, & Montero-Melis, 2019; **grice1975?**). During in-lab tasks with sighted participants, participants tailor their descriptions and requests by verbally providing visually-absent cues when an object is occluded to their partner (Hawkins, Gweon, & Goodman, 2021; Rubio-Fernandez, 2019; **jaraettinger2021?**). These results suggest that adults and even infants (Chiesa, Galati, & Schmidt, 2015; Ganea et al., 2018; Senju et al., 2013) can flexibly adapt communication to the visual and auditory abilities of their partner.

Curiously though, these patterns are not borne out in the existing literature on interactions between blind infants and their sighted parents. We might expect parents to verbally compensate for missing visual input, resulting in parents providing more description of the child's environment. Instead, caregivers of blind children seem to restrict conversation to things that the blind child is currently engaged with, rather than attempt to redirect their attention to other stimuli (Andersen, Dunlea, & Kekelis, 1993; J. Campbell, 2003; Kekelis & Andersen, 1984). In naturalistic settings, parents of blind children use *fewer* declaratives *more* imperatives than parents of sighted children, suggesting that children might be receiving less description than sighted children (Kekelis & Andersen, 1984; Landau & Gleitman, 1985). That said, we see some evidence for parents adapting to their child's visual abilities. Tadić, Pring, and Dale (2013 Nov-Dec) and colleagues find that in a structured book reading task, parents of blind children provide more descriptive utterances than parents of sighted children. Further, parents of blind children provide more tactile cues to initiate interactions or establish joint attention (Gunilla M. Preisler, 1991; Urwin, 1983), which may serve the same social role as shared gaze in sighted children. These mixed results suggest that parents of blind children might alter language input in some domains but not others.

Better understanding how sensory perception and linguistic input interact to influence blind children's language outcomes is of great clinical and scientific importance. Based on our own interactions with participants' families in the present study, parents are looking for evidence-based guidance to help them support their children's language development. If properties of language input influence the likelihood of language delays among blind infants and toddlers (**campbell_inprep?**), capturing this variation may reveal a more nuanced picture of how infants use the input to learn language. By contrast, if there is no relationship between language input properties and children's language outcomes, then trying to modify language input can be one less worry for caregivers. In the present study, we examine daylong recordings of the naturalistic language environments of blind and sighted children in order to characterize the input to each group. We first measure input quantity (adult word count) and analyze several characteristics that may be information-rich learning cues, including interactivity(conversational turn counts, proportion of child-directed speech), conceptual features (temporal displacement, sensory modality), and linguistic complexity (type/token ratio and mean length of utterance). We then link these properties of language input to language outcomes and explore whether the effects vary as a function of children's perceptual ability.

Methods

Participants

15 blind infants and their families participated in this study. Blind participants were recruited through opthamologist referral, preschools, early intervention programs, social media, and word of mouth. To be eligible for this study, participants had to be 6–30 months old, have no additional disabilities (developmental delays; intellectual disabilities, or hearing loss), and be exposed to $\geq$ 75% English at home. Given the wide age range of the study, to control for age, each blind participant was matched to a sighted partcipicant, based on age ($\pm$ 6 weeks), gender, maternal education ($\pm$ one education level: less than high school

diploma, high school diploma, some college / Associate's, Bachelor's, graduate school), and number of siblings (± 1 sibling). When more than one match was available, we prioritized matching the blind participants as closely as possible on each characteristics in the preceding order. Caregivers were asked to complete a demographics survey and the MacArthur-Bates Communicative Development Inventory (CDI, Fenson et al., 1994) within one week of the home language recording. See XXX for sample characteristics.

Recording Procedure

Eligible families were asked to complete two surveys and complete a daylong home language recording. For the recording portion of the study, caregivers of participating infants received a LENA wearable audio recorder (Ganek & Eriks-Brophy, 2016) and vest. They were instructed to place the recorder in the vest on the day of their scheduled recording and put the vest on their child from the time they woke up until the recorder automatically shut off after 16 hours (setting vest nearby during bath, nap, and car times). They were also instructed how to pause the recording at any time, but asked to keep these pauses to a minimum. Actual recording length ranged from 8 hours 17 minutes to 15 hours 59 minutes (14 hours 49 minutes).

Processing

Audio recordings were first processed by LENA proprietary software, creating algorithmic measures such as conversational turn counts. Each recording was then run through an in-house automated sampler that selected 15- non-overlapping 5-minute segments, randomly distributed across the duration of the recording. The process output a codeable ELAN file (.eaf, (**ELAN?**)). Each segment consists of 2 core minutes of annotated time, with 2 minutes of listenable context marked out preceding the annotation clip and 1 minute of additional context following the annotation clip. Each file therefore contains 30 minutes of coded recording time and 75 minutes of total time listened. Because these segments were sampled randomly, and not on a high-volubility measure such as conversational turns or adult speech density, the amount of time with codeable speech input varied for each recording. Indeed, across participants (FIND A WAY TO DO MATH WITH # SEGMENTS THAT ARE SILENT) of the 2-minute coding segments contained no speech at all.

Once the randomly selected segments were annotated, we also chose to annotate 15 additional segments specifically for their high levels of speech. To select these segments of dense talk, we first conducted an automated analysis of the audio file using the voice type classifier for child-centered daylong recordings (**lavechin2020?**) which identified all human speech in the recording. The entire recording was then broken into 2-minute chunks marked out at zero-second timestamps (e.g. 00:02:00.000 to 00:04:00.000). Each of these chunks was then ranked highest to lowest by the total duration of speech contained within the boundaries. For our high volubility sample, we chose the highest-ranked 15 segments of each recording, excluding those that overlapped with already-coded random segments.

Annotation

Trained annotators listened through each 2-minute segment plus its surrounding context and coded it using the Analyzing Child Language Experiences around the World (ACLEW) Daylong Audio Recording of Children's Linguistic Environments (DARCLE) annotation scheme (Soderstrom et al., 2021). Prior to annotating lab data, annotators are trained on previously coded samples of child recordings and are required to reach 95% overall agreement with the gold standard version of the file for three different age ranges: 0-7 months, 8-18 months, and 19-36 months. For more information about this annotation scheme and the larger project, please see the ACLEW homepage (https://sites.google.com/view/aclewdid/home). Following the first pass, all files were by a highly-trained "superchecker" to ensure the consistency of annotations.

This annotation scheme is designed to capture both utterances by the target child and speech in the child's environment, including adults, other children, and pre-recorded electronic speech (e.g. toys, television, the radio). Annotators segment the duration of each utterance on a separate coding tier for each unique speaker (exceptions: all electronic speech is coded on the same tier, and some speakers who appear briefly in these files were not easily distinguishable from others by annotators naive to their identities, so they may be concatenated on the same tier). Speech by people other than the target child is transcribed using an adapted version of CHAT transcription style (MacWhinney, 2019), dubbed minCHAT for the ACLEW project (Soderstrom et al., 2021). Because the majority of target children in the project are pre-lexical or phonetically immature, utterances produced by the target child are not transcribed.

Each utterance is coded for additional linguistic properties from a set of pre-determined categories. Target child utterances are coded for vocal maturity, lexical status, and multi-word status. Vocal maturity classifies utterances into the following categories: laughing; crying; canonical syllables that contain a consonant-like and vowel-like sound component, including both babbling and identifiable words; non-canonical syllables, which do not contain both consonant and vowel portions, or which do not transition between them in a speech-like way; and unsure, when the vocalization type is unclear. Each vocalization that contains canonical syllables is then coded for lexical status (does it contain an identifiable lexical item?). Finally, each utterance with a lexical item is coded for multi-word status (does it contain more than one unique word type?).

Environmental speech from everyone else is coded for the addressee of each utterance: speech directed to a child, whether or not it is directed to the target child; adult-directed speech; speech directed to both an adult and a child; speech directed to pets or other animals; speech with an unclear addressee; or speech directed towards a recipient that doesn't fit into another category (e.g. voice control of Siri or Alexa, prayer to a metaphysical entity).

Results

Measuring Properties of Language Input

We first seek to assess whether language input to blind children is categorically different from the language input to sighted children, along the dimensions of quantity, interactiveness, linguistic properties, and conceptual properties. For continuous variables, we test for group differences using xxx, and for categorical variables we test for differences with xxx.,We use non-parametric versions of these tests when a Shapiro-Wilks test indicates that the variable is not normally distributed. Because this analysis involves multiple tests of the null hypothesis (*that there is no difference in the language input to blind vs. sighted kids*), we use the conservative Bonferroni correction to set our threshold for significance (p = 0.05 / 8 tests = 0.01).

Language Input Quantity.

```
##
##  Shapiro-Wilk normality test
##
## data:  LENA_counts$AWC
## W = 0.93872, p-value = 0.04635
```

We first compare the quantity of language input to blind and sighted children using two measures of the number of words in their environment: LENA's automated Adult Word Count and word token count from our manual annotations. Shapiro-Wilks tests indicated that both of these variables were normally distributed (ps > .05).

Turning first to LENA's automated measure, a two-sample t-test shows that despite wide variability in the number of words children hear (Range: 6233–31745 $\text{words}_{\text{blind}}$, 6027–25500 $\text{words}_{\text{sighted}}$), blind and sighted children do not differ in language input quantity ($t(22.56) = -1.04$, $p = .311$). If we instead measure this using word counts from the transcriptions of the audio recordings, we find parallel results: blind and sighted children do not differ in language input quantity ($t(27.00) = 0.08$, $p = .939$); see Figure 1.

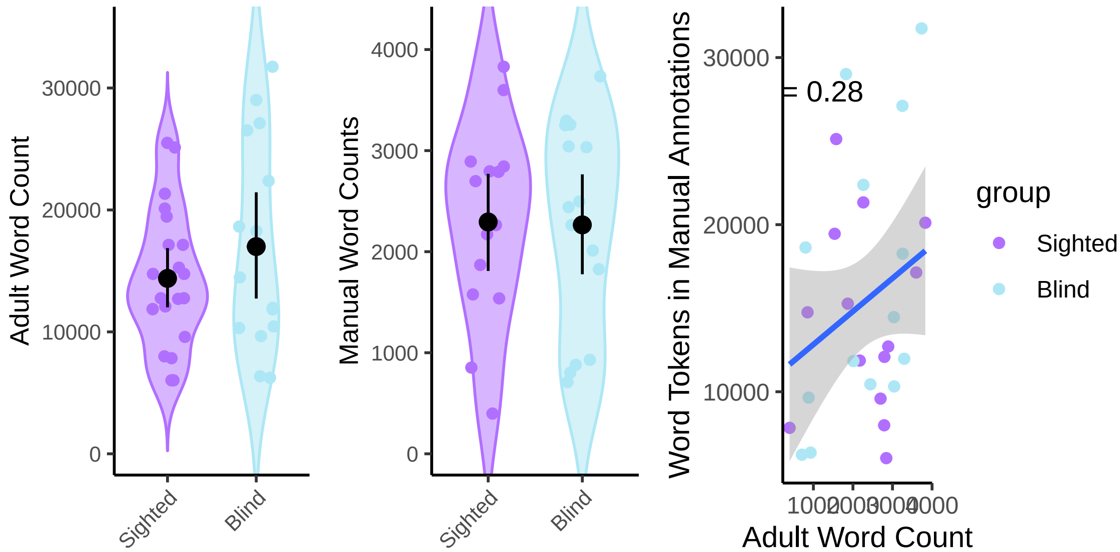


Figure 1. Comparing LENA-generated adult word counts (left) and transcription-based word counts in the input of blind and sighted children. Each dot represents the estimated number of words in one child's recording.

Interactiveness. We compared the proportions of child-directed speech (CDS) between the blind children and their sighted matches. Each proportion was calculated as the number of utterances produced by someone *other* than the target child (non-CHI utterances) tagged with a child addressee out of the total number of non-CHI utterances for each sensory group. A two-sample test for equality of proportions revealed no significant difference in the overall proportions of CDS to blind children and CDS to sighted children.

We next compare the number of conversational turn counts for blind and sighted children, using LENA's automated Conversational Turn Count measure. This measure is not normally distributed (W = 0.92, $p = .924$). Despite wide variability in conversational turns (210–1436 $_{\text{blind}}$, 112–1348 $_{\text{sighted}}$), we find no evidence for group-level differences between blind and sighted children (W = 140, $p = .585$).

Linguistic Features.

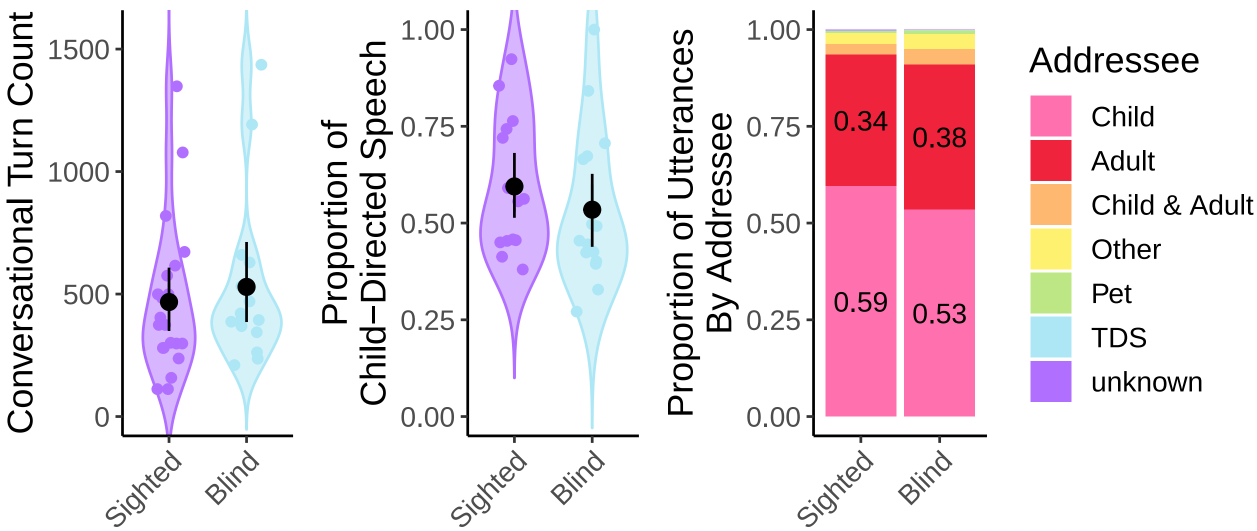
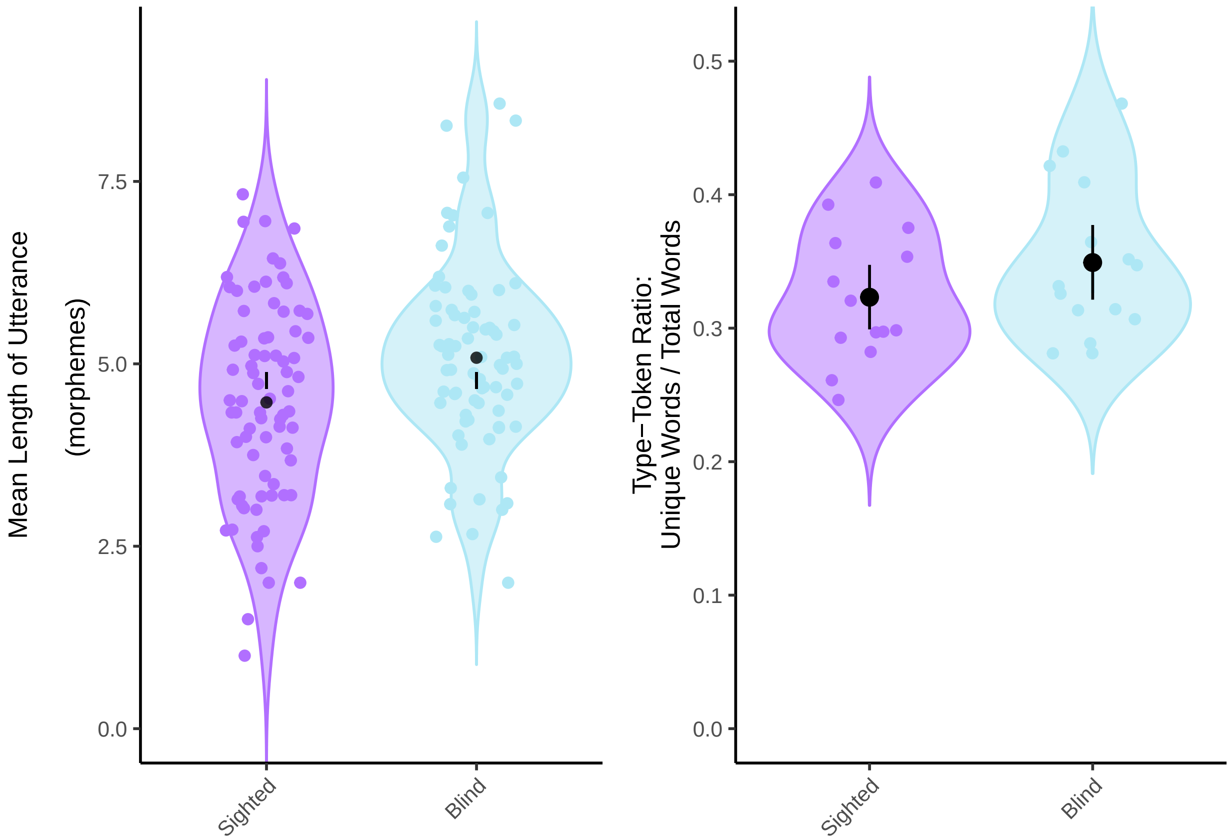


Figure 2. Comparing LENA-generated conversational turn counts (left) and proportion of utterances in child-directed speech (center). Each dot represents one child's recording. The full breakdown by addressee is shown in the rightmost panel.

```
##
##  Shapiro-Wilk normality test
##
## data:  manual_word_TTR$TTR
## W = 0.95771, p-value = 0.2883
```

```
##
##  Welch Two Sample t-test
##
## data:  manual_word_TTR$TTR by manual_word_TTR$group
## t = -1.2906, df = 26.749, p-value = 0.2079
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.06732386  0.01534716
## sample estimates:
```

```
## mean in group TD mean in group VI
##         0.3231817         0.3491700
```



For linguistic features, we first measure the proportion of unique words divided by the number of total words in the input, or type-token ratio, from the manual annotations. Because this variable met the normality assumption, we performed a two-sample t-test. Results indicated that there was no significant difference in the type-token ratio between the two groups (t(26.75) = -1.29, p = .208). This suggests that, on average, the type-token ratio is similar for blind (M: 0.35) and sighted (M: 0.32) children (see Figure **??**). These results provide evidence that the variety of words in the input is not affected by children's vision.

We also analyzed the syntactic complexity of children's language input, approximated as utterance length in morphemes. Each utterance by a non-CHI speaker was tokenized into morphemes using the 'morphemepiece' R package [CITE]. We then calculated the mean length of utternace (MLU) per speaker in each audio recording, and then compared the

MLU of environmental speech to blind children (M(SD) = 5.08 (1.29)) to that of sighted children (M(SD) = (M(SD) = 4.47 (1.39)); this variable was normally distributed (W = 0.92, p = .924). A two-sample t-test revealed that the MLU was slightly but significantly higher in speech to blind children than to their sighted peers (t(147.71) = -2.80, p = .006).

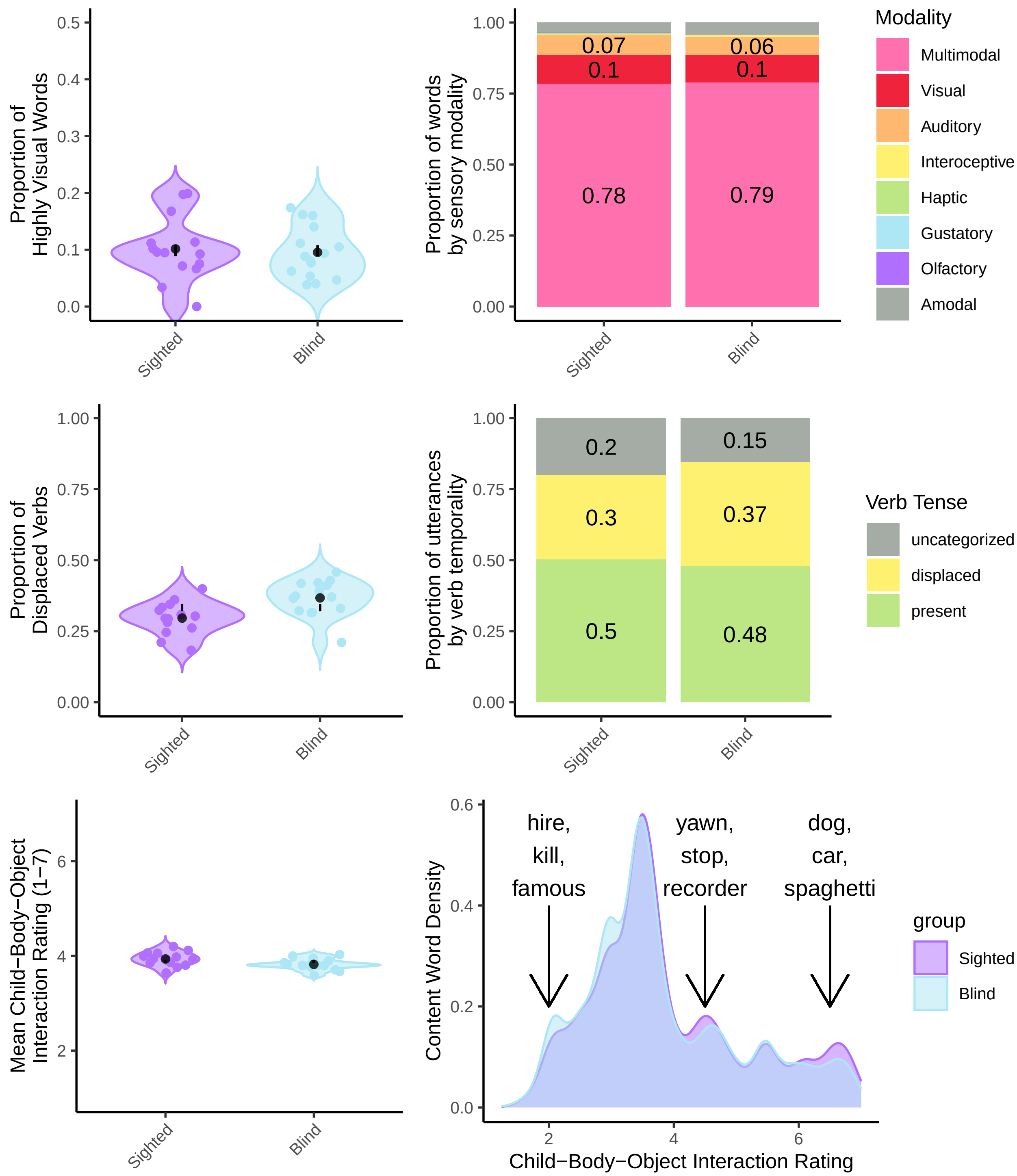
Conceptual Features. Our analysis of the conceptual features aims to measure whether the extent to which language input centers around the *"here and now"*: objects/events that are currently present/occurring vs. displaced objects/events. Prior work has quantified such *here-and-now*ness by counting object presence co-occurring with a related noun label [CITE]. The audio format of our data and the coding scheme we use make it difficult to ascertain object presence, so instead of object displacement, in this analysis, we approximate *here-and-now*ness using lexical and syntactic properties of the input. We do this by comparing 1) What proportion of words are temporally displaced?; 2) To what extent can children physically engage in / interact with words' referents?; and 3) What proportion of words have referents that can only be experienced through vision?

The last conceptual feature we examined is the displacement of events discussed in children's linguistic environment, via properties of the verbs in their input. Notably, we are attempting to highlight semantic features of the language environment; however, given the constraints of large-scale textual analysis, we are categorizing utterances based on a combination of closely related syntactic and morphological features of verbs, since these contain time-relevant information. We recognize that these linguistic features do not perfectly align with the temporal structure of the world. We assigned each utterance a **temporality** value: utterances tagged *displaced* describe events that take place in the past, future, or irrealis space, while utterances tagged *present* describe current, ongoing events. A small amount of utterances (n = 1306) were left *uncategorized* because they were fragments or because the automated parser failed to tag any of the relevant features. To do this, we used the udpipe package [CITE] to tag the transcriptions with parts of speech and other lexical features, such as tense, number agreement, or case inflection. To be marked as present, a verb either had to be marked with both present tense and indicative mood, or appear in the gerund form with no marked tense (e.g. *you talking to Papa?*). Features that could mark an utterance as displaced included past tense, presence of a modal, presence of *if*, or presence of *gonna/going to*, *have to*, *wanna/want to*, or *gotta/got to*, since these typically indicate belief states and desires, rather than real-time events. In the case of utterances with multiple verbs, we selected the features from the first verb or auxiliary, as a proxy for hierarchical dominance.

We compare the proportion of temporally displaced verbs using a Wilcoxon rank-sum test, given that a Shapiro-Wilks test indicates that the proportion of displaced verbs does not follow a normal distribution (W = 0.98, p = .977). We find that blind children hear proportionally more displaced verbs than blind children (W = 36.50, p = .003).

Next, we measure whether Child-Body-Object Interaction (CBOI) rating (**muraki2022?**). These norms were generated by asking parents of six-year-olds to rate the extent to which children physically interact with words' referents, from 1 (*things that a typical child does not easily physically interact with*) to 7 (*things a typical child would easily physically interact with*). We first use the udpipe part-of-speech tags to filter to content words (adjectives, adverbs, nouns, and verbs). Words without a CBOI rating (N = XXX/XXX) were removed. We then compared the distribution of CBOI ratings in word tokens in blind children's input to that in sighted children's input using a two-sample Kilgomorov-Smirnov test. We find that these distributions significantly differ (D = 0.98, $p <$ .001) this difference survives Bonferroni correction. Descriptively, low CBOI words were more common in language input to blind children, and high CBOI words were more common in language input to sighted children.

Lastly, we measure whether the language input to blind children contains a different proportion of words referring to visual objects/actions/properties. This is perhaps the dimension that people tend to have the strongest a priori hyptheses about: *Perhaps parents* *speak less about visual concepts to blind children because they're less relevant to the children's experiences* or alternatively *Perhaps parents speak* more* *about visual concepts, in order to compensate for experiences they perceive their children as missing.* We categorize the perceptual modalities of words' referents using the Lancaster Sensorimotor Norms, ratings from typically-sighted adults about the extent to which a word evokes a visual/tactile/auditory/etc. experience (**lynott2019?**). Words with higher ratings in a given modality are more strongly associated with perceptual experience in that modality. A word's dominant perceptual modality is the modality which received the highest mean rating. We tweak this categorization in two ways: words which received low ratings (< 3.5) across all modalities were re-categorized as *amodal*, and words whose ratings were distributed across modalities were re-categorized as *multimodal*. Using this system, each of the content words in children's input (adjectives, adverbs, nouns, and verbs) were categorized into their primary perceptual modality. For each child, we extracted the proportion of "visual" words in their language environment; this variable was normally distributed (W = 0.96, $p = .962$). We found no differences across groups in the proportion of highly visual words ($t(25.11) = 0.32$, $p = .755$).



Linking Language Input to Children's Expressive Language

As a last analysis, we construct models connecting children's own communicative productions to properties of the language input, as measured above, in order to explore whether different properties of language input predict child language outcomes in blind vs. sighted children. We measure child language in two complementary ways: LENA-estimated child vocalization percentile expressive vocabulary percentile on the Communicative Development Inventory (Fenson et al., 1994; CDI, **bates1994?**).

Measuring Child Language. The CDI is a parent-report measure of children's receptive and expressive vocabulary that has been used to measure XXX in many cultures and developmental contexts (**frankCITE?**), including with blind children (**campbell_submitted?**). For each of the $398_{\text{Words \& Gestures}}$-$680_{\text{Words \& Sentences}}$ words on the CDI, parents report whether their child understands and/or produces the word. We use the word production measure only, to avoid uncertainty about how parents make the "understands word" judgement and whether that differs across groups. Expressive vocabulary percentiles were then extracted from XXX. Given that the American English CDI is only normed for 8-30 months [CITE], participants younger than 8 months ($n = 4$) and their sighted matches were excluded from the model predicting vocabulary percentile.

In order to include all participants in this analysis, we also build a model predicting children's automatic vocal analysis (AVA) standard score, a measure that uses LENA's speech recognition to categorize and quantify the speech sounds in child vocalizations (Richards, Gilkerson, Paul, & Xu, 2009). The distribution of these speech sounds is used to predict children's vocal maturity. AVA has been normed and validated for sighted, typically-developing children from two- to 48-months-old (Richards et al., 2017), correlates strongly with other measures of expressive language, demonstrates decent test-retest reliability, and has also been used in clinical settings with children with autism, children with language delays, and Deaf/Hard-of-Hearing children (Oller et al., 2010; VanDam et al., 2015; **warren2009?**).

Predicting Expressive Language. We constructed two linear models, one predicting vocabulary percentile and one predicting AVA standard score, using the same model fitting procedure. First, we re-aggregated the language input variables such that each

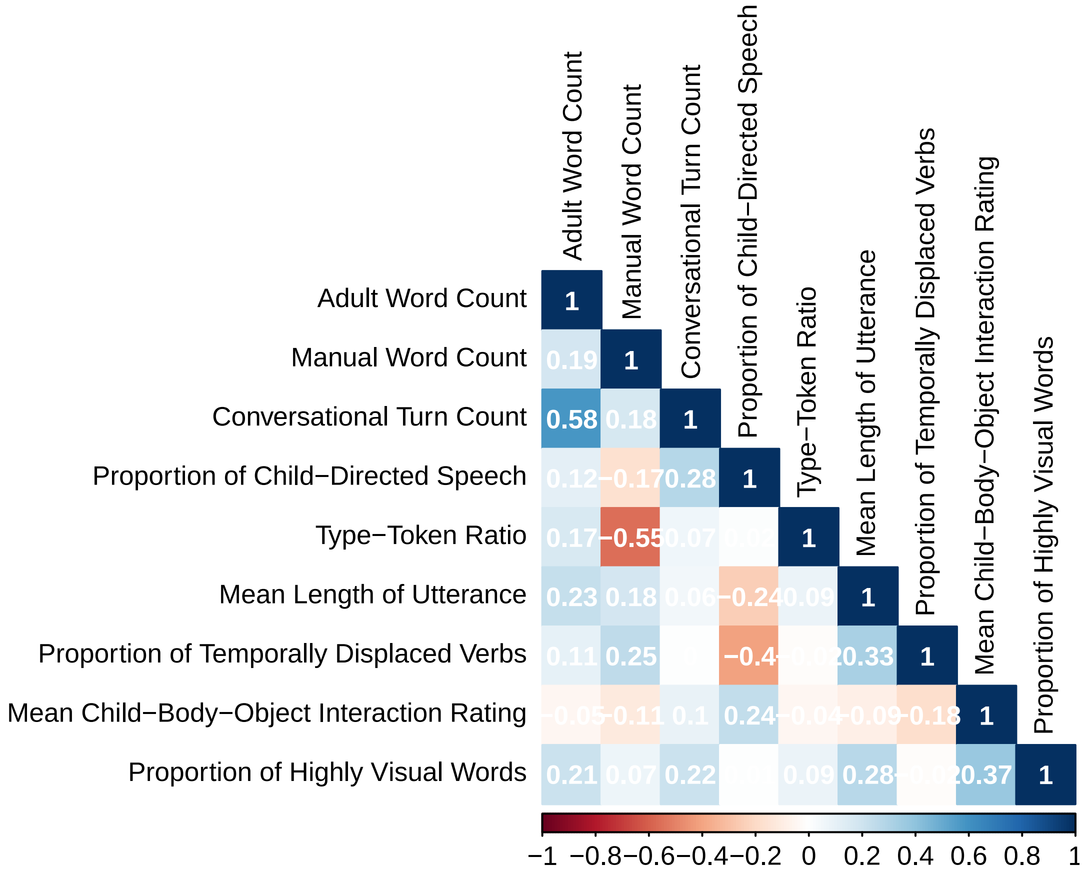


Figure 3. Kendall's Tau correlations between subject-level language input variables. Significant correlations are marked with asterisks.

child had a single value for each predictor; this required calculating MLU over child rather than over speaker and giving each child's input a mean child-body-object interaction rating. Next, we checked for associations among our language input predictor variables by running zero-order Kendall's Tau correlations; see Figure **??**. For predictor pairs were significantly correlated with each other ($p < .05$), we removed one of the variables. Seven out of the thirty-six predictor combinations were significantly associated with each other: Adult Word Count / Conversational Turn Count (r = 0.54, $p < .001$), Manual Word Count / Type-Token Ratio (r = -0.55, $p < .001$), Mean Length of Utterance / Proportion of Child-Directed Speech (r = -0.24, p = .072), Proportion of Child-Directed Speech / Proportion of Temporally Displaced Verbs (r = -0.40, p = .003), Mean Length of Utterance

/ Proportion of Temporally Displaced Verbs ($r = 0.33$, $p = .014$), Mean Length of Utterance / Proportion of Highly Visual Words ($r = 0.28$, $p = .038$), Mean Child-Body-Object Interaction Rating / Proportion of Highly Visual Words ($r = 0.37$, $p = .005$). To avoid collinearity, we removed conversational turn count, manual word count, proportion of child-directed speech, proportion of temporally displaced verbs, and proportion of highly visual words from our set of predictor variables. Thus, our final set of predictor variables was: adult word count, type-token ratio, mean length of utterance, and mean child-body-object interaction rating for content words.

Discussion - ignore me for now

Among our measures of input quantity and various characteristics, three differed between blind and sighted children: the distribution of highly child-interactable words, the displacement of events, and the mean length of utterance. Along these lines, blind children's input is more similar to adult-directed speech than sighted children's, with slightly longer utterances, lower child-body interaction ratings, and more displaced speech, although all differences between groups are small in magnitude. Displacement and MLU both predict vocabulary outcomes as measured by the CDI. Critically, however, there is no interaction of these effects with children's perceptual abilities.

One explanation for the minimal differences between blind and sighted children's language environments is parents' ability to assess their children's engagement and cognitive level, and thereby tailor their speech accordingly. Sighted parents may be unfamiliar with blind children's signals of interest (Perez-Pereira & Conti-Ramsden, 1999), and as a result, may respond less often to infants' vocalizations and bids for communication (Rowland, 1984), instead defaulting to more adultlike language. This lends credence to the idea that sensory access is most useful in the earliest stages of word learning, and without the scaffolding of simpler language and visual reference, blind infants are delayed in their ability to take advantage of more complex linguistic information.

The small differences in linguistic input might also be accentuated by differences in nonverbal communication between blind infants and their sighted caregivers. Young children born with visual impairment may differ in their nonverbal communication cues. For example, (G. M. Preisler, 1995) found that 6–9-month-old blind infants communicated using leaning, eyebrow raising, and lip movements. Caregivers who responded to these nonverbal cues as conversational turns had higher rates of interaction with the child, higher rates of appropriate response, and increased positive affect. By contrast, caregivers who did not recognize these signals as communicative had lower rates of response and increased negative affect.

References

Andersen, E. S., Dunlea, A., & Kekelis, L. (1993). The impact of input: Language acquisition in the visually impaired. *First Language*, *13*(37), 23–49. https://doi.org/10.1177/014272379301303703

Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). Linking Quality and Quantity of Parental Linguistic Input to Child Language Skills: A Meta-Analysis. *Child Development*, *92*(2), 484–501. https://doi.org/10.1111/cdev.13508

Bernstein Ratner, N. (1984). Patterns of vowel modification in mother–child speech. *Journal of Child Language*, *11*, 557–578.

Campbell, E. E., & Bergelson, E. (2022). Making sense of sensory language: Acquisition of sensory knowledge by individuals with congenital sensory impairments. *Neuropsychologia*, *174*, 108320. https://doi.org/10.1016/j.neuropsychologia.2022.108320

Campbell, J. (2003). Maternal Directives to Young Children who are Blind. *Journal of Visual Impairment & Blindness*, *97*(6), 355–365. https://doi.org/10.1177/0145482X0309700604

Chiesa, S., Galati, D., & Schmidt, S. (2015). Communicative interactions between visually impaired mothers and their sighted children: Analysis of gaze, facial expressions, voice and physical contacts. *Child: Care, Health and Development*, *41*(6), 1040–1046. https://doi.org/10.1111/cch.12274

De Villiers, J. (1985). Learning how to use verbs: Lexical coding and the influence of the input*. *Journal of Child Language*, *12*(3), 587–595. https://doi.org/10.1017/S0305000900006668

Demir, Ö. E., Rowe, M. L., Heller, G., Goldin-Meadow, S., & Levine, S. C. (2015). Vocabulary, syntax, and narrative development in typically developing children and children with early unilateral brain injury: Early parental talk about the "there-and-then" matters. *Developmental Psychology*, *51*(2), 161–175. https://doi.org/10.1037/a0038476

Donnellan, E., Bannard, C., McGillion, M. L., Slocombe, K. E., & Matthews, D. (2020).

Infants' intentionally communicative vocalizations elicit responses from caregivers and are the best predictors of the transition to language: A longitudinal investigation of infants' vocalizations, gestures and word production. *Developmental Science*, *23*(1), e12843. https://doi.org/10.1111/desc.12843

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., ... Stiles, J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development*, *59*(5), i. https://doi.org/10.2307/1166093

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, *60*(6), 1497–1510.

Ganea, N., Hudry, K., Vernetti, A., Tucker, L., Charman, T., Johnson, M. H., & Senju, A. (2018). Development of adaptive communication skills in infants of blind parents. *Developmental Psychology*, *54*(12), 2265–2273. https://doi.org/10.1037/dev0000564

Ganek, H., & Eriks-Brophy, A. (2016). The Language ENvironment Analysis (LENA) system: A literature review. In *Proceedings of the joint workshop on NLP for Computer Assisted Language Learning and NLP for Language Acquisition* (pp. 24–32). Umeå, Sweden: LiU Electronic Press.

Gergle, D., Kraut, R. E., & Fussell, S. R. (2004). Language Efficiency and Visual Technology: Minimizing Collaborative Effort with Visual Information. *Journal of Language and Social Psychology*, *23*(4), 491–517. https://doi.org/10.1177/0261927X04269589

Gilkerson, J., Richards, J. A., Warren, S. F., Oller, D. K., Russo, R., & Vohr, B. (2018). Language Experience in the Second Year of Life and Language Outcomes in Late Childhood. *Pediatrics*, *142*(4), e20174276. https://doi.org/10.1542/peds.2017-4276

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, *19*(5), 515–523. https://doi.org/10.1111/j.1467-9280.2008.02117.x

Grigoroglou, M., Edu, U., & Papafragou, A. (2016). Are children flexible speakers? Effects of typicality and listener needs in children's event descriptions. *Cognitive Science*, 6.

Hadley, P. A., Rispoli, M., Holt, J. K., Papastratakos, T., Hsu, N., Kubalanza, M., & McKenna, M. M. (2017). Input Subject Diversity Enhances Early Grammatical Growth: Evidence from a Parent-Implemented Intervention. *Language Learning and Development: The Official Journal of the Society for Language Development*, *13*(1), 54–79. https://doi.org/10.1080/15475441.2016.1193020

Hawkins, R. D., Gweon, H., & Goodman, N. D. (2021). The Division of Labor in Communication: Speakers Help Listeners Account for Asymmetries in Visual Perspective. *Cognitive Science*, *45*(3), e12926. https://doi.org/10.1111/cogs.12926

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., … Suma, K. (2015). The Contribution of Early Communication Quality to Low-Income Children's Language Success. *Psychological Science*, *26*(7), 1071–1083. https://doi.org/10.1177/0956797615581493

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

Hoff, E., & Naigles, L. (2002 Mar-Apr). How children use input to acquire a lexicon. *Child Development*, *73*(2), 418–433. https://doi.org/10.1111/1467-8624.00415

Hsu, N., Hadley, P. A., & Rispoli, M. (2017). Diversity matters: Parent input predicts toddler verb production. *Journal of Child Language*, *44*(1), 63–86. https://doi.org/10.1017/S0305000915000690

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*, 236–248. https://doi.org/10.1037/0012-1649.27.2.236

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, *45*(3), 337–374. https://doi.org/10.1016/s0010-0285(02)00500-5

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343–365. https://doi.org/10.1016/j.cogpsych.2010.08.002

Kekelis, L. S., & Andersen, E. S. (1984). Family Communication Styles and Language Development. *Journal of Visual Impairment & Blindness*, *78*(2), 54–65. https://doi.org/10.1177/0145482X8407800202

Landau, B., & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child* (pp. xi, 250). Cambridge, MA, US: Harvard University Press.

Lucca, K., & Wilbourn, M. P. (2018). Communicating to Learn: Infants' Pointing Gestures Result in Optimal Learning. *Child Development*, *89*(3), 941–960. https://doi.org/10.1111/cdev.12707

MacLeod, A. A. N., & Demers, C. (2023). Transmitting white monolingual Anglo-American norms: A concept analysis of "quality of language" in parent-child interactions. *Applied Psycholinguistics*, 1–29. https://doi.org/10.1017/S014271642300005X

MacWhinney, B. (2019). CHAT Manual. https://doi.org/10.21415/3MHN-0Z89

Naigles, L. R., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language*, *25*(1), 95–120. https://doi.org/10.1017/S0305000997003358

Newman, R. S., Rowe, M. L., & Bernstein Ratner, N. (2016). Input and uptake at 7 months predicts toddler vocabulary: The role of child-directed speech and infant processing skills in language development. *Journal of Child Language*, *43*(5), 1158–1173. https://doi.org/10.1017/S0305000915000446

Oller, D. K., Niyogi, P., Gray, S., Richards, J. A., Gilkerson, J., Xu, D., … Warren, S. F. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(30), 13354–13359. https://doi.org/10.1073/pnas.1003882107

Ostarek, M., Paridon, J. van, & Montero-Melis, G. (2019). Sighted people's language is not helpful for blind individuals' acquisition of typical animal colors. *Proceedings of the National Academy of Sciences*, *116*(44), 21972–21973. https://doi.org/10.1073/pnas.1912302116

Perez-Pereira, M., & Conti-Ramsden, G. (1999). *Language Development and Social Interaction in Blind Children.* London: Psychology Press. https://doi.org/10.4324/9780203776087

Preisler, Gunilla M. (1991). Early patterns of interaction between blind infants and their sighted mothers. *Child: Care, Health and Development*, *17*(2), 65–90. https://doi.org/10.1111/j.1365-2214.1991.tb00680.x

Preisler, G. M. (1995). The development of communication in blind and in deaf infants–similarities and differences. *Child: Care, Health and Development*, *21*(2), 79–110. https://doi.org/10.1111/j.1365-2214.1995.tb00412.x

Richards, J. A., Gilkerson, J., Paul, T., & Xu, D. (2009). *The LENA Automatic Vocalization Assessment* (pp. 1–20). Boulder, CO: The LENA Foundation.

Richards, J. A., Xu, D., Gilkerson, J., Yapanel, U., Gray, S., & Paul, T. (2017). Automated Assessment of Child Vocalization Development Using LENA. *Journal of Speech, Language, and Hearing Research*, *60*(7), 2047–2063. https://doi.org/10.1044/2017_JSLHR-L-16-0157

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. E. (2018). Beyond the 30-Million-Word Gap: Children's Conversational Exposure Is Associated With Language-Related Brain Function. *Psychological Science*, *29*(5), 700–710. https://doi.org/10.1177/0956797617742725

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014 May-Jun). Skype me! Socially contingent interactions help toddlers learn language. *Child Development*, *85*(3), 956–970. https://doi.org/10.1111/cdev.12166

Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill*. *Journal of Child Language*, *35*(1), 185–205. https://doi.org/10.1017/S0305000907008343

Rowe, M. L. (2012). A Longitudinal Investigation of the Role of Quantity and Quality of Child-Directed Speech in Vocabulary Development. *Child Development*, *83*(5), 1762–1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x

Rowe, M. L., & Snow, C. E. (2020). Analyzing input quality along three dimensions: Interactive, linguistic, and conceptual. *Journal of Child Language*, *47*(1), 5–21. https://doi.org/10.1017/S0305000919000655

Rowland, C. (1984). Preverbal Communication of Blind Infants and Their Mothers. *Journal of Visual Impairment & Blindness*, *78*(7), 297–302. https://doi.org/10.1177/0145482X8407800701

Rubio-Fernandez, P. (2019). Overinformative Speakers Are Cooperative: Revisiting the Gricean Maxim of Quantity. *Cognitive Science*, *43*(11), e12797. https://doi.org/10.1111/cogs.12797

Senju, A., Tucker, L., Pasco, G., Hudry, K., Elsabbagh, M., Charman, T., & Johnson, M. H. (2013). The importance of the eyes: Communication skills in infants of blind parents. *Proceedings. Biological Sciences*, *280*(1760), 20130436. https://doi.org/10.1098/rspb.2013.0436

Snow, C. E. (1972). Mothers' Speech to Children Learning Language on JSTOR. *Child Development*, *43*, 549–565.

Soderstrom, M., Casillas, M., Bergelson, E., Rosemberg, C., Alam, F., Warlaumont, A. S., & Bunce, J. (2021). Developing a Cross-Cultural Annotation System and MetaCorpus for Studying Infants' Real World Language Experience. *Collabra: Psychology*, *7*(1), 23445. https://doi.org/10.1525/collabra.23445

Tadić, V., Pring, L., & Dale, N. (2013 Nov-Dec). Story discourse and use of mental state language between mothers and school-aged children with and without visual impairment. *International Journal of Language & Communication Disorders*, *48*(6), 679–688. https://doi.org/10.1111/1460-6984.12040

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-Directed Speech Facilitates Word Segmentation. *Infancy: The Official Journal of the International Society on Infant Studies*, *7*(1), 53–71. https://doi.org/10.1207/s15327078in0701_5

Tomasello, M., & Farrar, M. J. (1986). Joint Attention and Early Language. *Child Development*, *57*(6), 1454. https://doi.org/10.2307/1130423

Uccelli, P., Demir-Lira, Ö. E., Rowe, M. L., Levine, S., & Goldin-Meadow, S. (2019). Children's Early Decontextualized Talk Predicts Academic Language Proficiency in Midadolescence. *Child Development*, *90*(5), 1650–1663. https://doi.org/10.1111/cdev.13034

Urwin, C. (1983). Dialogue and cognitive functioning in the early language development of three blind children: Normal and deficient, 142–161.

VanDam, M., Oller, D. K., Ambrose, S. E., Gray, S., Richards, J. A., Xu, D., … Moeller, M. P. (2015). Automated Vocal Analysis of Children with Hearing Loss and Their Typical and Atypical Peers. *Ear and Hearing*, *36*(4), e146–e152. https://doi.org/10.1097/AUD.0000000000000138

Vygotsky, L. S., & Cole, M. (1978). *Mind in Society: Development of Higher Psychological Processes*. Harvard University Press.

Weizman, Z. O., & Snow, C. E. (2001). Lexical input as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology*, *37*(2), 265–279. https://doi.org/10.1037/0012-1649.37.2.265

Yoshinaga-Itano, C., Sedey, A. L., Mason, C. A., Wiggin, M., & Chung, W. (2020). Early Intervention, Parent Talk, and Pragmatic Language in Children With Hearing Loss. *Pediatrics*, *146*(Supplement_3), S270–S277. https://doi.org/10.1542/peds.2020-0242F

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262. https://doi.org/10.1016/j.cognition.2012.06.016

Yurovsky, D., Doyle, G., & Frank, M. C. (2016). Linguistic input is tuned to children's developmental level. *Cognitive Science*, 6.