Comparing Language Input in Homes of Blind and Sighted Children: Insights from Daylong Recordings

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.

Comparing Language Input in Homes of Blind and Sighted Children: Insights from Daylong Recordings

Abstract

Purpose: This study compared language input to young blind children and their sighted peers in naturalistic home settings.

Methods: Using the LENA audio recorder, naturalistic speech in the home was captured and analyzed for various dimensions of language input, including quantitative, interactive, linguistic, and conceptual features.

Results: Our data showed far more similarity than difference across groups, with all differences being small in magnitude. Both groups received similar speech quantity, interactiveness, and lexical diversity. Fine-grained analysis revealed that blind children's language environments contained longer utterances, more temporal displacement, and content words that are harder for children to interact with, suggesting greater similarity to adult-directed speech.

Conclusions: The findings challenge the notion that blind children's language input places them at a disadvantage and suggest that blind children receive rich and complex language input that can support their language development.

Introduction

The early language skills of blind children are highly variable (E. E. Campbell, Casillas, & Bergelson, submitted), with some blind children demonstrating age-appropriate vocabulary from the earliest stages of language learning (Ann Bigelow, 1987; E. E. Campbell et al., submitted; Landau & Gleitman, 1985), while others experience large and persistent language delays (E. E. Campbell et al., submitted). By adulthood, blind individuals are fluent speakers of their language and are even reported to have faster auditory and lexical processing skills than sighted adults (Röder, Demuth, Streb, & Rösler, 2003; Röder, Rösler, & Neville, 2000). The causes of this variability and the later ability to "catch up" remain poorly understood. In particular, the incidence of severe language delays in blind children (E. E. Campbell et al., submitted) yields questions about the process of language development in the absence of visual perception: what could make the language learning problem different and initially more difficult for the blind child? There are multiple possible contributors to the variability of the language development within the group of blind children, 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). Across groups, we expect the primary systematic difference to be children's sensory ability. However, the child's cognitive and perceptual abilities interplay with their language input to facilitate language development, and we should explore the possibility of systematic differences in both addends. 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.

Why would input matter?

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 & 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, Pittman, & Walden, 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 influential (Hirsh-Pasek et al., 2015; Rowe, 2012), although it is somewhat trickier to turn the qualitative characteristics of language input into operationalizable properties. In the present study, 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.

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,

[1] In the field thus far, the directionality of the term “quality” 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 from which children can learn (MacLeod & Demers, 2023).

words heard in contexts where the adult and child share joint attention are more likely to be learned (Lucca & Wilbourn, 2018; Tomasello & Farrar, 1986). 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, 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 provide more unique words (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 objects and events in the world (Rowe & Snow, 2020). The pieces of the conceptual content of language input that are most informative may shift across developmental time: as children develop, their ability to represent abstract, displaced, decontextualized referents improves (Bergelson & Swingley, 2013; Kramer, Hill, & Cohen, 1975; Luchkina, Xu, Sobel, & Morgan, 2020) (though object permanence and related skills may be delayed in blind children, S. J. Rogers and Puchalski (1988)). For example, infants are more likely to learn a new word when the referent is perceptually salient, dominating their field of view (Yu & Smith, 2012; Yurovsky, Smith, & Yu, 2013). 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, 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 in turn, can offer 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 in particular provides cues to word meaning that may be lost without visual cues, such as the relationship between two entities that aren't within reach (Gleitman, 1990). 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 (Landau & Gleitman, 1985). If direct sensory access to referents provides an initial "brute force" mechanism for mapping words onto meanings, it may take longer for blind children to acquire the first few words. By hypothesis, once this initial seed of lexical knowledge is acquired, blind children and sighted children alike are able to use more abstract and linguistic features as cues, and learning can proceed more rapidly thereafter (Babineau, de Carvalho, Trueswell, & Christophe, 2021;

Babineau, Havron, Dautriche, de Carvalho, & Christophe, 2022; 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.

Why would the input differ?

First, speakers regularly tailor input to communicate efficiently with the listener (Grice, 1975). 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 (Dirks, Stevens, Kok, Frijns, & Rieffe, 2020; Yoshinaga-Itano, Sedey, Mason, Wiggin, & Chung, 2020).

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 & Baker, 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 (Grice, 1975; Ostarek, Paridon, & Montero-Melis, 2019). 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; Jara-Ettinger & Rubio-Fernandez, 2021; Rubio-Fernandez, 2019). These results suggest that adults and even infants (Chiesa, Galati, & Schmidt, 2015; N. 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; though c.f., **moore1994?**). In naturalistic settings, parents of blind children use *fewer* declaratives and *more* imperatives than parents of sighted children, suggesting that children might be receiving less description than sighted children (Kekelis & Andersen, 1984; Landau & Gleitman, 1985). On the other hand, some parents may adapt to their children's visual abilities in specific contexts. 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 (Preisler, 1991; Urwin, 1983, 1984), 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. Blind children's own use of decontextualized language appears to develop later than sighted children's (A. Bigelow, 1990; Urwin, 1984).

The Present Study

Reaching a better understanding of how sensory perception and linguistic input interact to influence blind children's language outcomes is of great scientific, clinical, and education importance. If properties of language input influence the likelihood of language delays among blind infants and toddlers (E. E. Campbell et al., submitted), 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. Using both automated measures and manual transcription of these recordings, we measure input quantity (adult word count) and analyze several characteristics that have been previously suggested as 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).

Methods

Participants

29 blind infants and their families participated in this study. Blind participants were recruited through ophthalmologist 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 participant, 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 ($\pm$ 1 sibling). When more than one match was available, we prioritized matching the

blind participants as closely as possible on each characteristic 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 Table @ref(tab: participant-characteristics) for sample characteristics.

Recording Procedure

Eligible families were asked to complete two surveys as well as a daylong home language recording. For the recording portion of the study, caregivers of participating infants received a LENA wearable audio recorder and vest (Ganek & Eriks-Brophy, 2016; Gilkerson & Richards, 2008). 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 on 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 (Mean: 15 hours 16 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 outputs a codeable ELAN file (.eaf, Brugman & Russel, 2009). Each segment consists of 2 core minutes of annotated time, with 2 minutes of listenable context 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 roughly 27% of the random 2-minute coding segments contained no speech at all. For questions of *how much does a phenomenon occur*, random sampling schemes can help avoid overestimating speech in the input, but for questions of input *content*, randomly selected samples may be too sparse (Pisani, Gautheron, & Cristia, 2021).

Therefore, we also chose to annotate 5 additional segments specifically for their high density 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 (Lavechin, Bousbib, Bredin, Dupoux, & Cristia, 2021) 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 5 segments of each recording, excluding those that overlapped with already-coded random segments. These high volubility segments allowed us to characterize features of the language as proportions of the linguistic input children receive and to compare our findings more closely to studies classifying the input during structured play sessions, which paint a denser and differently-proportioned makeup of the language input (Bergelson, Amatuni, Dailey, Koorathota, & Tor, 2019). In sum, 30 minutes of randomly sampled input and 10 minutes of input produced 40 minutes of annotated recording time per child.

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 reviewed 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. Speech by people other than the target child is transcribed using an adapted version of the 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, utterances produced by the target child are not yet transcribed. Environmental speech is then coded for the addressee of each utterance: speech directed to a 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).

Extracting Measures of Language Input

To go from our dimensions of interest (quantity, interactiveness, linguistic, conceptual), to quantifiable properties, we used a combination of automated measures [generated by the proprietary LENA algorithm; Xu, Yapanel, and Gray (2009)] and manual measures (generated from the transcriptions made by our trained annotators). These manual annotations can be analyzed for the random segments, the high-volume segments, or both. The decision on which segments to analyze was made according to the goal of the analysis: quantity and interactiveness analyses were conducted on the random samples only, to try to capture a more representative estimate. Linguistic and conceptual analyses were conducted on all available annotations in order to maximize the amount of speech over which we could calculate them. These measures are summarized in Table 1.

Quantity.

Adult Word Count.

To derive this count, first the LENA algorithm segments the recording into clips of varying length. These segments are then classified as female adult speech, male adult speech, target child, other child, overlapping vocalization/noise, electronic noise, noise, silence, or uncertain, each of which is further categorized into "near" or "far". Only segments that are classified as nearby male or female adult speech are included in the Adult Word Count estimation; Segments that the LENA algorithm identifies as "far", "child", or "overlapping", do not contribute to this count (Xu et al., 2009). Validation work suggests that this automated count correlates strongly with word counts derived from manual annotations [r = .71 – .92; Lehet, Arjmandi, Houston, and Dilley (2021)], but Lehet et al. (2021) and colleagues find that the amount of error may vary substantially across families. Compared to short samples that they had manually transcribed and counted, LENA's AWC estimate ranged from undercounting words by 17% to overcounting words by 208% (Lehet et al., 2021). Perhaps reassuringly however, meta-analytic work finds that AWC is associated with children's language outcomes across developmental contexts [Wang, Williams, Dilley, and Houston (2020); e.g., autism, hearing loss]. Because the recordings varied in length (8 hours 17 minutes to 15 hours 59 minutes), we normalized AWC by dividing by recording length [2].

Manual Word Count.

We also compare a manual count of speech in the children's environment. Manual word count is simply the number of intelligible words in our transcriptions of each child's recording. Speech that was too far or muffled to be intelligible, as well as speech from the

[2] To make this comparable to the manual word count estimates, which are derived from the 30 minutes of random sample annotation, we calculate AWC per half hour.

target child and electronic speech (TV, radio, toys) are excluded from this count. In order to try to get a representative estimate of the amount of talk in a children's environment, we use the random samples only for this measure.

By using Adult Word Count and Manual Word Count, we hope to capture complementary estimates of the amount of speech children are exposed to. AWC is less accurate, but commonly used, and provides an estimate of the speech across the whole day. MWC, because it comes from human annotations, is the gold-standard for accurate speech estimates, but is only derived from 30 minutes of the recording.

Interactivity.

Conversational Turn Count.

One commonly used and easily-extracted metric of communicative interaction (e.g., Ganek & Eriks-Brophy, 2018; Magimairaj, Nagaraj, Caballero, Munoz, & White, 2022) is conversational turn count (or CTC), an automated measure generated by LENA (Xu et al., 2009). Like AWC, a recent meta-analysis finds that CTC is associated with children's language outcomes (Wang et al., 2020). After tagging vocalizations for speaker identity, LENA algorithm looks for alternations between adult and target child speech in close temporal proximity. The algorithm counts any temporally close (within 5 seconds) switch between adult and target child vocalizations, which can erroneously include non-contingent interactions (e.g., mom talking to dad while the infant babbles to herself nearby), and therefore inflate the count especially for younger ages and in houses with multiple children (Ferjan Ramírez, Hippe, & Kuhl, 2021). Still, this measure correlates moderately well with manually-coded conversational turns (Busch, Sangen, Vanpoucke, & van Wieringen, 2018; Ganek & Eriks-Brophy, 2018), and because participants in our sample are matched on both age and number of siblings, CTC overestimation should not be biased towards either groups. Conversational turn count is calculated over the entire recording, but to normalize for recording length, we divided this by recording length.

Proportion of Child-Directed Speech.

Our other measure of interactivity is the proportion of utterances that are child-directed, derived from the manual annotations. Each proportion was calculated as the number of utterances (produced by someone *other* than the target child) tagged with a child addressee out of the total number of utterances. To try to get a representative measure of child-directed speech in the environment overall, we use the random samples only for this calculation.

Linguistic Features.

Type-Token Ratio.

As in previous work (Montag, Jones, & Smith, 2018; Pancsofar & Vernon-Feagans, 2006; e.g., Templin, 1957), we calculated type-token ratio by dividing the number of unique words by the total number of words. Because the type-token ratio changes as a function of the size of the language sample (Montag et al., 2018; Richards, 1987), we first standardized the sample length by cutting children's input (from the manual annotations) in each recording into 100-word bins. We then calculated the type-token ratio within each of these bins by dividing the number of unique words in each bin by the number of total words (~100). For each child, type-token ratio is the average of the type-token ratios for each of the bins in their input.

MLU.

We also analyzed the syntactic complexity of children's language input, approximated as mean utterance length in morphemes. Both type-token ratio and mean length of utterance in speech to infants remain consistent for individual caretakers, in and out of lab settings (Stevenson, Leavitt, Roach, Chapman, & Miller, 1986). Each utterance was tokenized into morphemes using the 'morphemepiece' R package (Bratt, Harmon, & Learning, 2022). We then calculated the mean length of utterance (number of morphemes) per speaker in each audio recording. We manually checked utterance length in a random subset of 10% of the utterances, and reached XXX agreement with the udpipe approach, indicating high reliability.

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/people 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 (P. A. Ganea & Saylor, 2013; Harris, Jones, Brookes, & Grant, 1986; e.g., Osina, Saylor, & Ganea, 2013; **moore1994?**). The audio format of our data 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?

Proportion temporally displaced verbs.

We examined 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 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. This coding scheme roughly aligns with both the temporal displacement and future hypothetical categories in Grimminger, Rohlfing, Lüke, Liszkowski, and Ritterfeld (2020; Hudson, 2002; see also: Lucariello and Nelson, 1987). To do this, we used the udpipe package (Wijffels, 2023) 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. A small number of utterances in our corpus were left *uncategorized* (n = 1306/7867) because they were fragments or because the automated parser failed to tag any of the relevant features. We manually checked verb temporality in a random subset of 10% of the utterances, and reached XXX agreement with the NLP approach, indicating moderate reliability.

CBOI distribution.

Next, we measured whether the distribution of Child-Body-Object Interaction (CBOI) rating differed across groups (Muraki, Siddiqui, & Pexman, 2022). 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*). These ratings are another measure of the amount of sensorimotor information wrapped up in language input to children, which may make certain words easier to learn and process (Muraki et al., 2022). 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 = 5188/28821) were removed.

Proportion highly visual words.

In addition to these two more traditional measures of decontextualized language, we include one measure that is uniquely decontextualized for the blind children relative to their sighted matches: the proportion of words in the input with referents that are highly and exclusively visual. 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 (Lynott, Connell, Brysbaert, Brand, & Carney, 2020). 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 exclusively "visual" words in their language environment.

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, interactivity, linguistic properties, and conceptual properties. We test for group differences using paired, t-tests or the non-parametric Wilcoxon signed rank tests, when a Shapiro-Wilks test indicates that the variable is not normally distributed. Because this analysis involves multiple tests against 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. 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: 195–992 words$_{\text{blind}}$, 238–804 words$_{\text{sighted}}$), blind and sighted children do not differ in language input quantity ($t(28) =$ -1.19, $p = .245$). 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.

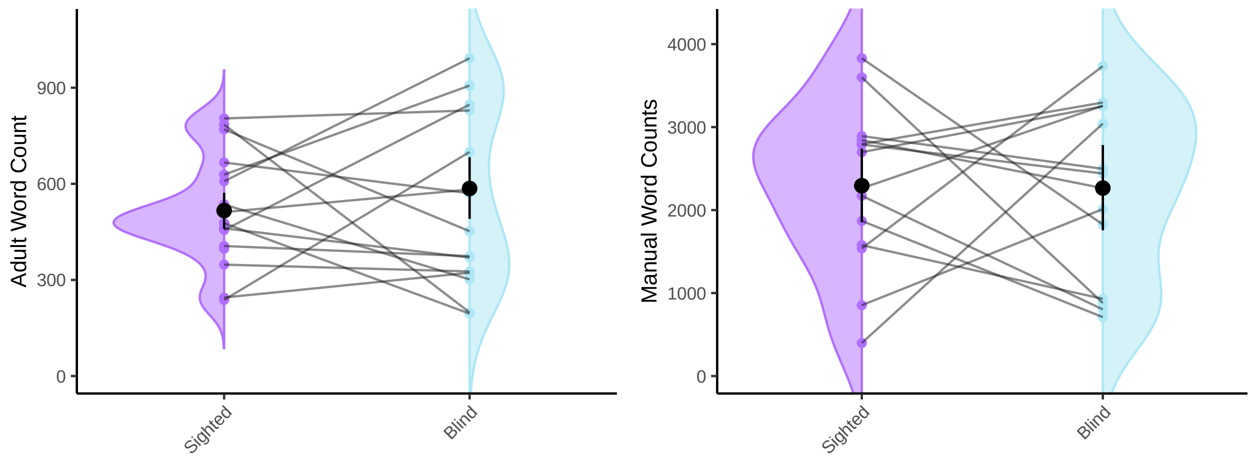

Figure 1. Comparing LENA-generated adult word counts (left) and transcription-based word counts in the input of blind and sighted children. Each dot represents the estimated number of words in one child's recording.

Interactivity. We compared the proportions of child-directed speech (CDS) between the blind children and their sighted matches. A two-sample test for equality of proportions revealed no significant difference in the overall proportions of child-directed speech to blind children versus to sighted children (W = 0.86, $p = .399$).

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 normally distributed (W = 0.88, $p = .924$). Despite wide variability in conversational turns (7–45 $_{\text{blind}}$, 4–42 $_{\text{sighted}}$), we find no evidence for group-level differences between blind and sighted children ($t = 295$, $p = .096$).

Linguistic Features. For linguistic features, we first measure, or type-token ratio, from the manual annotations. Because this variable met the normality assumption, we

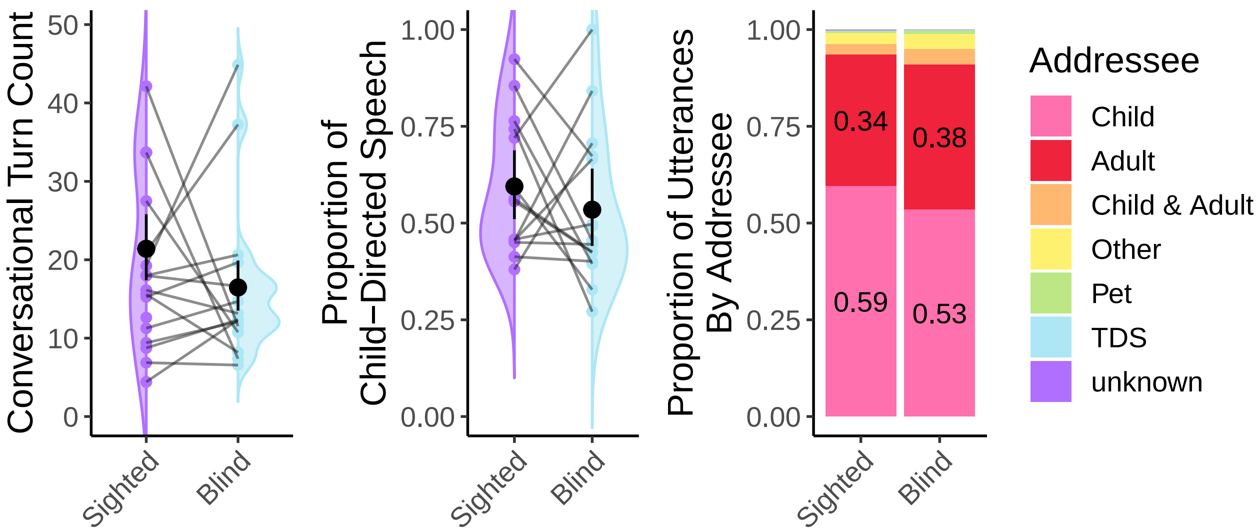

Figure 2. Comparing LENA-generated conversational turn counts (left) and proportion of utterances in child-directed speech (center). Each dot represents one child's recording. The full breakdown by addressee is shown in the rightmost panel.

performed a two-sample t-test. Results indicated that there was no significant difference in the type-token ratio between the two groups ($t(18.46)$ = -2.36, p = .030). This suggests that, on average, the type-token ratio is similar for blind (M: 0.70) and sighted (M: 0.67) children (see Figure 3). These results provide evidence that the variety of words in the input is not affected by children's vision.

We 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); see Figure 3).

Conceptual Features. 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 sighted children

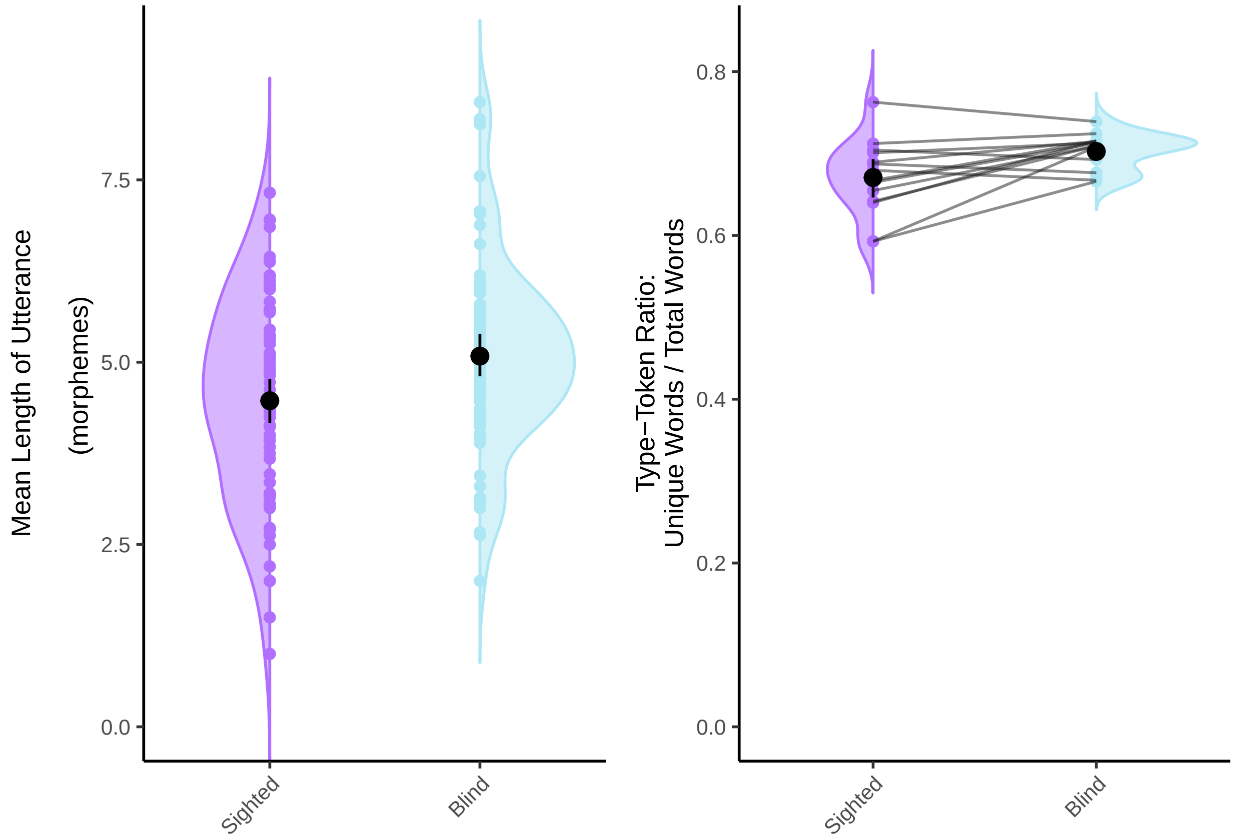

Figure 3. Comparing linguistic features: Mean length of utterance (left); each dot represents one speaker. Type-token ratio (right). Each dot represents one child's recording.

($W = 36.50$, $p = .003$).

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 variable was normally distributed (W = 0.96, $p = .962$). We found no differences across groups in the proportion of highly visual words ($t(25.11) = 0.32$, $p = .755$).

Discussion

This study measured language input to young blind children and their sighted peers, using the LENA audio recorder to capture naturalistic speech in the home. We found that across many dimensions of language input, parents largely talk similarly to blind and sighted children, with a few nuanced differences, that we discuss further below.

Quantity

Across both of these measures of language input quantity, one estimated from the full sixteen hour recording (Adult Word Count) and one precisely measured from a 30-minute window of that day (Manual Word Count), blind and sighted children were exposed to similar amounts of speech in the home. Quantity was highly variable *within* groups, but we found no evidence for *between* group differences in input quantity.

Interactivity

We quantified interactivity in two ways: through the LENA-estimated conversational turn count, and through the proportion of child-directed speech in our manual annotations. Again, we found no differences across groups in the amount of parent-child interaction. This finding runs counter to previous research; other studies report *less* interaction in dyads where the child is blind (Pérez-Pereira & Conti-Ramsden, 2001; Rowland, 1984; Andersen et al., 1993; Grumi et al., 2021; Kekelis & Andersen, 1984; Preisler, 1991;, **moore1994?**). Using a non-visual sampling method (i.e., our audio recordings) might provide a different, more naturalistic perspective on parent-child interactions, particularly in this population. For one thing, many of these studies (e.g., Kekelis & Andersen, 1984; Pérez-Pereira & Conti-Ramsden, 2001; Preisler, 1991; **moore1994?**) involve video recordings in the child's home, with the researcher present. Like other young children, blind children distinguish between familiar individuals and strangers, and react with trepidation to the presence of a stranger (Fraiberg, 1975; McRae, 2002); for blind children, this reaction may involve "quieting", wherein children cease speaking or vocalizing when they hear a new voice in the home (Fraiberg, 1975; McRae, 2002). By having a researcher present during the recordings[3] , prior research may have artificially suppressed blind children's initiation of interactions. Even naturalistic observer-free video-recordings appear to inflate aspects of parental input, relative to daylong recordings (Bergelson et al., 2019). In these cases, the video camera acts as an observer itself, making participants aware of its presence, limiting participants' mobility, and therefore shrinking the pragmatic scope of possible interactions. Together, these factors could explain why past parent-child interaction research finds that blind children initiate less (Andersen et al., 1993; Dote-Kwan, 1995; Kekelis & Andersen, 1984; Tröster & Brambring, 1992; **moore1994?**), that parents do most of the talking (Andersen et al., 1993; Kekelis & Andersen, 1984), and that there is overall less interaction (Nagayoshi et al., 2017; Sally J. Rogers & Puchalski, 1984; Rowland, 1984; Tröster & Brambring, 1992).

Additionally, a common focus in earlier interaction literature is to measure visual cues of interaction, such as shared gaze or attentiveness to facial expressions (Baird, Mayfield, & Baker, 1997; Preisler, 1991; Sally J. Rogers & Puchalski, 1984). We can't help but wonder: are visual markers of social interaction the right yardstick to measure blind children against? In line with MacLeod and Demers (2023), perhaps the field should move away from sighted indicators of interaction "quality", and instead try to to situate blind children's interactions within their own developmental niche, one that may be better captured with auditory- or tactile-focused coding schemes.

Linguistic Features

Along the linguistic dimension, we measured type-token ratio and mean length of utterance. Type-token ratio was similar across groups, and in line with type-token ratio in

[3] Fraiberg (1975) writes "these fear and avoidance behaviors appear even though the observer, a twice-monthly visitor, is not, strictly speaking, a stranger." (pg. 323).

other child-centered corpora (e.g., Newman, Rowe, & Bernstein Ratner, 2016). However, we found slightly but significantly higher MLU in blind children's language environment. The MLU finding runs counter to common advice: Parents of children with disabilities (Chernyak, n.d.; including parents of blind children! e.g., FamilyConnect, n.d.) are often advised to use shorter, simpler sentences with their children, in order to promote children's understanding. We find instead that the language environments of blind children contain *longer* utterances, which could suggest that consciously modifying your linguistic behavior is difficult for parents. In any case, this advice is not supported by the literature: evidence suggests that longer, more complex utterances are associated with better child language outcomes in both typically-developing children (Hoff & Naigles, 2002 Mar-Apr) and children with cognitive differences (Sandbank & Yoder, 2016). Infants older than 12 months also show no difference in listening preference modulated by utterance length or repetition (Segal & Newman, 2015), which indicates that they do not use complexity as a basis for discrimination in what types of input to learn from.

Conceptual Features

The conceptual features of language input feel slipperiest to operationalize. For this analysis, we chose to capture *here-and-now*-ness by measuring the proportion of temporally displaced verbs, the distribution of high vs. low child-body-object interaction ratings for content words, and the proportion of highly visual words. Relative to other aspects of language input, the conceptual dimension seemed to vary most across groups: though blind and sighted participants were exposed to a similar proportion of highly visual words, blind children heard more displaced verbs and their content words were distributed slightly more to the not-interactable side of the child-body-object interaction ratings.

The extent to which blind children's language input is centered on the *here-and-now* has been contested in the literature (Andersen et al., 1993; J. Campbell, 2003; Kekelis & Andersen, 1984; Urwin, 1984; **moore1994?**). Our sample, which contains more blind participants than prior research alongside a carefully peer-matched sighted sample, finds that blind children's input is *less* focused on the here-and-now. One possible explanation is that because children have less access to immediate visual cues, caregivers might instead refer to past or future events. To illustrate, while riding on a train, instead of describing the scenery passing outside the window, parents may choose to talk about what happened earlier in the day or their plans upon home. Without further information about the social and perceptual context, it is difficult to determine the communicative function of these results.

To that end, our exploratory analysis points to potential group differences in the *context* of conceptual information. Blind children's proportion of temporally displaced verbs was inversely correlated with their mean child-body-object interaction rating, whereas sighted children showed the reverse relationship. Could this suggest that when sighted children hear about words that are perceivable or manipulable, it tends to be in the context of co-present objects / events, but when blind children hear about things that can be interacted with, it tends to be related to past/future events? Additionally, while we found that overall, blind and sighted children hear a similar proportion of highly visual words (blue, mirror, rainbow, see), blind children (but not sighted children) who receive *more* child-directed speech seem to receive *less* of this highly visual language, which may point to caregivers tailoring their speech. Our present analyses can only hint at potential relationships between these variables at the child level, but as more dense annotation becomes available, we can explore the social and environmental context of conceptual information as it unfolds across discourse.

Patterns in Language Input

Before synthesizing any of these differences, we wish to highlight again how much variability there is *within* groups and how much consistency there is *between* groups. One could imagine a world in which the language environments of blind and sighted children are radically different from each other. Our data do not support that hypothesis. Rather, we find far more similarity across groups than differences, and all differences were small in magnitude. This is worth emphasizing and re-emphasizing: across developmental contexts, including, as we show here, visual experience, children's language input is resoundingly similar (Bergelson et al., 2022).

When we zoom into more fine-grained aspects of the input, we found that blind children's language environments contained longer utterances, more temporal displacement, and content words that are harder for children to interact with. Together, these features seem to suggest that blind children's input is more similar to speech directed towards older children or adults (Rowe, 2012; Snow, 1972) than sighted children's. This does not seem attributable to differences in addressee: our annotations indicate that there is a similar proportion of child-.vs.adult-directed speech across the two groups.

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. On the other hand, we found between-group differences in how these measures relate to each other. In speech to sighted children, there is a small positive relationship between the amount of child-directed speech and the quantity of highly visual words, but in speech to blind children the opposite is true: parents who use more child-directed speech use *less* highly visual language, which suggests that at least some caregivers are tailoring their language to their child's sensory access when speaking to their child specifically.

Yet, despite all this variation both within and between groups, both blind and sighted children grow up to be competent speakers (Röder et al., 2003; Röder et al., 2000). Future work should explore the relationship between these input measures and the children's own language outcomes; however, given the high variability of all of these variables, we do not expect parental input to be rigidly deterministic of successful language acquisition.

Connecting to Language Outcomes

Returning to the larger equation of language development, blind and sighted infants differ in their access to perceptual input, and we have shown that language input is different along only a few axes: conceptual features, where language and the perceptual world interact, and complexity, with blind children hearing slightly longer and more adult-like utterances, on average. Initial vocabulary delays in blind children may then primarily be a result of the conflict between their lack of visual access and the majority-visual cues to early "brute-force" word learning (e.g., shared gaze, pointing, visual perception of referents). It could be precisely this linguistic input complexity which aids blind children in acquiring semantic knowledge in the absence of visual input. Testing this prediction awaits further research.

Conclusion

In summary, our study compared language input in homes of 15 blind and 15 sighted infants/toddlers. We found that both groups received similar quantities of adult speech and had similar levels of interaction. However, blind children were exposed to longer utterances and more decontextualized language, suggesting that they are being exposed to a rich and complex linguistic environment that differs from the language input of sighted children. Our study does not imply that parents should change their communication styles, but rather highlights the importance of recognizing and appreciating the unique language experiences of blind children. Future research could investigate how these input differences impact the language development and cognitive abilities of blind children, and how we can better support their language and learning needs.

Results To Do: * Filter to random sample for manual word count, proportion CDS * Switch all statistical tests to paired tests

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

Babineau, M., de Carvalho, A., Trueswell, J., & Christophe, A. (2021). Familiar words can serve as a semantic seed for syntactic bootstrapping. *Developmental Science*, *24*(1), e13010. https://doi.org/10.1111/desc.13010

Babineau, M., Havron, N., Dautriche, I., de Carvalho, A., & Christophe, A. (2022). Learning to predict and predicting to learn: Before and beyond the syntactic bootstrapper. *Language Acquisition*, *0*(0), 1–24. https://doi.org/10.1080/10489223.2022.2078211

Baird, S. M., Mayfield, P., & Baker, P. (1997). Mothers' Interpretations of the Behavior of Their Infants with Visual and Other Impairments during Interactions. *Journal of Visual Impairment & Blindness*, *91*(5), 467–483. https://doi.org/10.1177/0145482X9709100507

Bergelson, E., Amatuni, A., Dailey, S., Koorathota, S., & Tor, S. (2019). Day by day, hour by hour: Naturalistic language input to infants. *Developmental Science*, *22*(1), e12715. https://doi.org/10.1111/desc.12715

Bergelson, E., Soderstrom, M., Schwarz, I.-C., Rowland, C., Ramirez-Esparza, N., Hamrick, L., … Casillas, M. (2022). Everyday language input and production in 1001 children from 6 continents.

Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. *Cognition*, *127*(3), 391–397. https://doi.org/10.1016/j.cognition.2013.02.011

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

Bigelow, Ann. (1987). Early words of blind children. *Journal of Child Language*, *14*(1), 47–56. https://doi.org/10.1017/S0305000900012721

Bigelow, A. (1990). Relationship between the Development of Language and Thought in Young Blind Children. *Journal of Visual Impairment & Blindness*, *84*(8), 414–419. https://doi.org/10.1177/0145482X9008400805

Bratt, J., Harmon, J., & Learning, B. F. &. W. P. G. L. D. M. (2022). Morphemepiece: Morpheme Tokenization.

Brugman, H., & Russel, A. (2009). Annotating Multimedia / Multi-modal resources with ELAN. *Proceedings of the Fourth International Conference on Language Resources and Evaluation*.

Busch, T., Sangen, A., Vanpoucke, F., & van Wieringen, A. (2018). Correlation and agreement between Language ENvironment Analysis (LENA™) and manual transcription for Dutch natural language recordings. *Behavior Research Methods*, *50*(5), 1921–1932. https://doi.org/10.3758/s13428-017-0960-0

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, E. E., Casillas, R., & Bergelson, E. (submitted). The Role of Vision in the Acquisition of Words: Vocabulary Development in Blind Toddlers. https://doi.org/10.17605/OSF.IO/UW6ZM

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

Chernyak, P. (n.d.). 3 Ways to Teach Your Blind or Visually Impaired Child to Talk. *WikiHow*. https://www.wikihow.life/Teach-Your-Blind-or-Visually-Impaired-Child-to-Talk.

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

Dirks, E., Stevens, A., Kok, S., Frijns, J., & Rieffe, C. (2020). Talk with me! Parental linguistic input to toddlers with moderate hearing loss. *Journal of Child Language*, *47*(1), 186–204. https://doi.org/10.1017/S0305000919000667

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

Dote-Kwan, J. (1995). Impact of Mothers' Interactions on the Development of Their Young Visually Impaired Children. *Journal of Visual Impairment & Blindness*, *89*(1), 46–58. https://doi.org/10.1177/0145482X9508900109

FamilyConnect. (n.d.). Understanding the Stages of Language Development for Babies Who Are Blind. *FamilyConnect*.

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

Ferjan Ramírez, N., Hippe, D. S., & Kuhl, P. K. (2021). Comparing Automatic and Manual Measures of Parent–Infant Conversational Turns: A Word of Caution. *Child Development*, *92*(2), 672–681. https://doi.org/10.1111/cdev.13495

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

Fraiberg, S. (1975). The development of human attachments in infants blind from birth. *Merrill-Palmer Quarterly*, *21*, 315–334.

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

Ganea, P. A., & Saylor, M. M. (2013). Talking about the near and dear: Infants' comprehension of displaced speech. *Developmental Psychology*, *49*(7), 1299–1307. https://doi.org/10.1037/a0030086

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.

Ganek, H., & Eriks-Brophy, A. (2018). Language ENvironment analysis (LENA) system investigation of day long recordings in children: A literature review. *Journal of Communication Disorders*, *72*, 77–85. https://doi.org/10.1016/j.jcomdis.2017.12.005

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. (2008). *The LENA Natural Language Study*. Boulder, CO: LENA Foundation.

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

Gleitman, L. (1990). The Structural Sources of Verb Meanings. *Language Acquisition*, *1*(1), 3–55. Retrieved from https://www.jstor.org/stable/20011341

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

Grice, H. P. (1975). Logic and Conversation. In *Syntax and semantics*. New York San Francisco London: Academic press, Harcourt Brace Jovanovich.

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.

Grimminger, A., Rohlfing, K. J., Lüke, C., Liszkowski, U., & Ritterfeld, U. (2020). Decontextualized talk in caregivers' input to 12-month-old children during structured interaction. *Journal of Child Language*, *47*(2), 418–434. https://doi.org/10.1017/S0305000919000710

Grumi, S., Cappagli, G., Aprile, G., Mascherpa, E., Gori, M., Provenzi, L., & Signorini, S. (2021). Togetherness, beyond the eyes: A systematic review on the interaction between visually impaired children and their parents. *Infant Behavior and Development*, *64*, 101590. https://doi.org/10.1016/j.infbeh.2021.101590

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

Harris, M., Jones, D., Brookes, S., & Grant, J. (1986). Relations between the non-verbal context of maternal speech and rate of language development. *British Journal of Developmental Psychology*, *4*(3), 261–268. https://doi.org/10.1111/j.2044-835X.1986.tb01017.x

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

Hazan, V., & Baker, R. (2011). Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. *The Journal of the Acoustical Society of America*, *130*(4), 2139–2152. https://doi.org/10.1121/1.3623753

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

Hudson, J. A. (2002). "Do You Know What We're Going to Do This Summer?": Mothers' Talk to Preschool Children About Future Events. *Journal of Cognition and Development*, *3*(1), 49–71. https://doi.org/10.1207/S15327647JCD0301_4

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

Jara-Ettinger, J., & Rubio-Fernandez, P. (2021). The social basis of referential communication: Speakers construct physical reference based on listeners' expected visual search. *Psychological Review*, No Pagination Specified–No Pagination Specified. https://doi.org/10.1037/rev0000345

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

Kramer, J. A., Hill, K. T., & Cohen, L. B. (1975). Infants' Development of Object Permanence: A Refined Methodology and New Evidence for Piaget's Hypothesized Ordinality. *Child Development*, *46*(1), 149–155. https://doi.org/10.2307/1128843

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

Lavechin, M., Bousbib, R., Bredin, H., Dupoux, E., & Cristia, A. (2021). An open-source voice type classifier for child-centered daylong recordings. arXiv. https://doi.org/10.48550/arXiv.2005.12656

Lehet, M., Arjmandi, M. K., Houston, D., & Dilley, L. (2021). Circumspection in using automated measures: Talker gender and addressee affect error rates for adult speech detection in the Language ENvironment Analysis (LENA) system. *Behavior Research Methods*, *53*(1), 113–138. https://doi.org/10.3758/s13428-020-01419-y

Lucariello, J., & Nelson, K. (1987). Remembering and planning talk between mothers and children. *Discourse Processes*, *10*(3), 219–235. https://doi.org/10.1080/01638538709544673

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

Luchkina, E., Xu, F., Sobel, D., & Morgan, J. (2020). *Sixteen-month-olds comprehend unanchored absent reference* (Preprint). Open Science Framework. https://doi.org/10.31219/osf.io/5tc6d

Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (2020). The Lancaster Sensorimotor Norms: Multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods*, *52*(3), 1271–1291. https://doi.org/10.3758/s13428-019-01316-z

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

Magimairaj, B., Nagaraj, N., Caballero, A., Munoz, K., & White, K. (2022). A Systematic Review of the Effects of LENA-based Feedback on Parent-Child Language Interactions in Families with Young Children. *Journal of Early Hearing Detection and Intervention*, *7*(3), 47–60. https://doi.org/10.26077/6c72-973b

McRae, K. A. (2002). *Attachment in blind infants : A systematic investigation using Ainsworth's Strange Situation.* (PhD thesis). University of Toronto, Toronto, Canada.

Montag, J. L., Jones, M. N., & Smith, L. B. (2018). Quantity and Diversity: Simulating Early Word Learning Environments. *Cognitive Science*, *42 Suppl 2*(Suppl 2), 375–412. https://doi.org/10.1111/cogs.12592

Muraki, E. J., Siddiqui, I. A., & Pexman, P. M. (2022). Quantifying children's sensorimotor experience: Child body-object interaction ratings for 3359 English words. *Behavior Research Methods*, *54*(6), 2864–2877. https://doi.org/10.3758/s13428-022-01798-4

Nagayoshi, M., Hirose, T., Toju, K., Suzuki, S., Okamitsu, M., Teramoto, T., … Takeo, N. (2017). Related visual impairment to mother-infant interaction and development in infants with bilateral retinoblastoma. *European Journal of Oncology Nursing: The*

Official Journal of European Oncology Nursing Society, *28*, 28–34. https://doi.org/10.1016/j.ejon.2017.02.002

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

Osina, M. A., Saylor, M. M., & Ganea, P. A. (2013). When familiar is not better: 12-month-old infants respond to talk about absent objects. *Developmental Psychology*, *49*, 138–145. https://doi.org/10.1037/a0027903

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

Pancsofar, N., & Vernon-Feagans, L. (2006). Mother and father language input to young children: Contributions to later language development. *Journal of Applied Developmental Psychology*, *27*(6), 571–587. https://doi.org/10.1016/j.appdev.2006.08.003

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

Pérez-Pereira, M., & Conti-Ramsden, G. (2001). The use of Directives in Verbal Interactions between Blind Children and their Mothers. *Journal of Visual Impairment & Blindness*, *95*(3), 133–149. https://doi.org/10.1177/0145482x0109500302

Pisani, S., Gautheron, L., & Cristia, A. (2021). *Long-form recordings: From A to Z.*

Preisler, G. 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

Richards, B. (1987). Type/Token Ratios: What do they really tell us? *Journal of Child Language*, *14*(2), 201. https://doi.org/doi:10.1017/s0305000900012885

Röder, B., Demuth, L., Streb, J., & Rösler, F. (2003). Semantic and morpho-syntactic priming in auditory word recognition in congenitally blind adults. *Language and Cognitive Processes*, *18*(1), 1–20. https://doi.org/10.1080/01690960143000407

Röder, B., Rösler, F., & Neville, H. J. (2000). Event-related potentials during auditory language processing in congenitally blind and sighted people. *Neuropsychologia*, *38*(11), 1482–1502. https://doi.org/10.1016/S0028-3932(00)00057-9

Rogers, Sally J., & Puchalski, C. B. (1984). Social Characteristics of Visually Impaired Infants' Play. *Topics in Early Childhood Special Education*, *3*(4), 52–56. https://doi.org/10.1177/027112148400300409

Rogers, S. J., & Puchalski, C. B. (1988). Development of Object Permanence in Visually Impaired Infants. *Journal of Visual Impairment & Blindness*, *82*(4), 137–142. https://doi.org/10.1177/0145482X8808200407

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. (2013). Decontextualized Language Input and Preschoolers' Vocabulary Development. *Seminars in Speech and Language*, *34*(4), 260–266. https://doi.org/10.1055/s-0033-1353444

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

Sandbank, M., & Yoder, P. (2016). The Association Between Parental Mean Length of Utterance and Language Outcomes in Children With Disabilities: A Correlational Meta-Analysis. *American Journal of Speech-Language Pathology*, *25*(2), 240–251. https://doi.org/10.1044/2015_AJSLP-15-0003

Segal, J., & Newman, R. S. (2015). Infant Preferences for Structural and Prosodic Properties of Infant-Directed Speech in the Second Year of Life. *Infancy*, *20*(3), 339–351. https://doi.org/10.1111/infa.12077

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

Stevenson, M. B., Leavitt, L. A., Roach, M. A., Chapman, R. S., & Miller, J. F. (1986). Mothers' speech to their 1-year-old infants in home and laboratory settings. *Journal of Psycholinguistic Research*, *15*(5), 451–461. https://doi.org/10.1007/BF01067725

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

Templin, M. C. (1957). *Certain language skills in children; their development and interrelationships* (pp. xviii, 183). Minneapolis, MN, US: University of Minnesota Press.

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

Tröster, H., & Brambring, M. (1992). Early social-emotional development in blind infants. *Child: Care, Health and Development*, *18*(4), 207–227. https://doi.org/10.1111/j.1365-2214.1992.tb00355.x

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.

Urwin, C. (1984). Language for absent things: Learning from visually handicapped children. *Topics in Language Disorders*, *4*(4), 24.

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

Wang, Y., Williams, R., Dilley, L., & Houston, D. M. (2020). A meta-analysis of the predictability of LENA™ automated measures for child language development. *Developmental Review*, *57*, 100921. https://doi.org/10.1016/j.dr.2020.100921

Watkins, S., Pittman, P., & Walden, B. (1998). The Deaf Mentor Experimental Project for young children who are deaf and their families. *American Annals of the Deaf*, *143*(1), 29–34. https://doi.org/10.1353/aad.2012.0098

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

Wijffels, J. (2023). UDPipe.

Xu, D., Yapanel, U., & Gray, S. (2009). *Reliability of the LENA Language Environment Analysis System in Young Children's Natural Home Environment* (pp. 1–16). Boulder, CO: The LENA Foundation.

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.

Yurovsky, D., Smith, L. B., & Yu, C. (2013). Statistical word learning at scale: The baby's view is better. *Developmental Science*, *16*(6), 959–966. https://doi.org/10.1111/desc.12036

Variable	Blind	Sighted	Overall
Age in months			
Mean (SD)	15.77 (8.20)	16.15 (8.15)	15.96 (8.04)
Min, Max	6.41, 30.38	6.18, 31.76	6.18, 31.76
Gender			
(Col %)			
F	7 (43.75%)	7 (43.75%)	14 (43.75%)
M	9 (56.25%)	9 (56.25%)	18 (56.25%)
Maternal education level			
(Col %)			
Some college	0 (0.00%)	0 (0.00%)	0 (0.00%)
Associate's degree	3 (23.08%)	1 (7.69%)	4 (15.38%)
Bachelor's degree	1 (7.69%)	2 (15.38%)	3 (11.54%)
Master's degree	5 (38.46%)	9 (69.23%)	14 (53.85%)
Missing	4 (30.77%)	1 (7.69%)	5 (19.23%)
Maternal education level	0 (0.00%)	0 (0.00%)	0 (0.00%)
Number of older siblings			
Mean (SD)	0.50 (0.82)	1.09 (1.04)	0.74 (0.94)
Min, Max	0.00, 2.00	0.00, 3.00	0.00, 3.00

Table 1

Language input variables extracted from recordings.

Variable	Coding	Portion of Recording	Description
Adult Word Count / half hour (AWC)	Automated	Whole day	Estimated number of words in recording categorized as nearby adult speech by LENA algorithm
Manual Word Count (WC)	Manual	Random	Number of word tokens from speakers other than target child
Conversational Turn Count / half hour (CTC)	Automated	Whole day	Count of temporally close switches between adult and target-child vocalizations, divided by recording length
Proportion of Child-Directed Speech (Prop. CDS)	Manual	Random	Number of utterances tagged with child addressee out of total number of utterances, from speakers other than target child
Type-Token Ratio	Manual	Random + High Volume	Average of the type-token ratios (number of unique words divided by number of total words) for each of the 100-word bins in their sample
Mean Length of Utterance	Manual + NLP parsing	Random + High Volume	Average number of morphemes per utterance (by speaker)
Proportion of Temporally Displaced Verbs (Prop. Displaced)	Manual + NLP tagging	Random + High Volume	Proportion of verbs that refer to past, future, or hypothetical events
Child-Body-Object Interaction Ratings (CBOI)	Manual + NLP tagging	Random + High Volume	Distribution of ratings of “how much a child can interact with” each word (adjectives, adverbs, nouns, verbs)
Proportion of Highly Visual Words	Manual	Random + High Volume	Proportion of words in the input with high visual association ratings and low ratings for other perceptual modalities