Supplement 1: Validity of the Communicative Development Inventory for Populations with Sensory Impairments

The data and code used to generate this paper are stored in an OSF repository, which will be made publicly accessible upon publication. This study received approval from xxx Institutional Review Board. A big thank you to xxx for her help transcribing the blind children's utterances.

Supplement 1: Validity of the Communicative Development Inventory for Populations with Sensory Impairments

This Supplemental Information accompanies the main manuscript titled "Early Production of Imperceptible Words by Infants and Toddlers Born Deaf or Blind". It contains two sections: an overview of previous validation work for the MacArthur-Bates Communicative Development Inventory (CDI, Fenson et al., 1994), and new validation data for a sample of blind children that supplements and converges with this prior work.

Review of Previous CDI Validity Studies

The CDI boasts a long history of validation testing, beginning in the late 1980s, which demonstrates that the CDI is a highly valid and reliable measure of children's vocabulary (Marchman & Dale, 2023). The instrument has multiple versions, which vary slightly across languages, but generally include a younger child version (Words & Gestures, typically used for 8-18mo) and an older child version (Words & Sentences, typically used for 16-30mo). While each version has multiple sections, the bulk of each consists of a vocabulary checklist with hundreds of words of varying difficulty falling into varying semantic and grammatical categories. Caretakers are asked to indicate which words the child understands or says (Words & Gestures) or just says (Words and Sentences). We first summarize validation efforts for word production for the American English CDI as used in typically-hearing/sighted children, then describe in more detail the results of studies validating the CDI for our populations of interest (See Table S1). See also Jarůšková, Smolík, Chládková, Oceláková, and Paillereau (2023) for a cross-linguistic review of validation efforts of the CDI and San Roque, Norcliffe, and Majid (2024) for a subset-analysis approach consistent to ours.

For typically-hearing/sighted children, the original CDI manual contains a summary of reliability and validity testing of the American English CDI (Fenson et al., 1994, pg. 25).

The authors show test-retest reliability of 0.95 (N=$137_{\text{Words \& Gestures}}$, $216_{\text{Words \& Sentences}}$, Fenson et al., 1994). The "Words Produced" measure of the Words & Gestures form of this instrument correlates at 0.72–0.79 with words produced during an in-lab language sample (N=27, Bates et al., 1988; Beeghly, Jernberg, & Burrows, 1989). On the Words & Sentences form, this measure correlates at 0.61 with the Vocabulary subscale of the Preschool Language Scales (N=20, Fenson et al., 1994), 0.73 with scores on the Expressive One-Word Picture Vocabulary Test (N=24, Dale, 1991), and 0.95 with children's Total Vocabulary score on the Language Development Survey (N=239, Rescorla, Ratner, Jusczyk, & Jusczyk, 2005).

Thal, Desjardin, and Eisenberg (2007) validated the American English CDI for use with deaf children who use cochlear implants. In their study, parents completed the CDI for their 32–86-month-old children (who had used a cochlear implant for 3–60 months). Children provided a spontaneous language sample during a brief free play session and two short book-based tasks. Children were also tested using the Reynell Developmental Language Scales. The "Words Produced" measure correlated at 0.84–0.89 with the receptive and expressive subscales of the Reynell Developmental Language Scales and 0.64–0.68 with the number of unique words produced by the child during the spontaneous language sample; see Table S1.

Likewise, the ASL CDI 2.0 (Caselli, Lieberman, & Pyers, 2020) shows high validity among deaf children of deaf (signer) parents (22–68 months, N=29) and deaf children of hearing (non-fluent signer) parents (24–67 months, N=11). For this validation, parents completed the ASL CDI 2.0 (which is administered with videos of ASL signs, accompanied by written English labels for each sign). Children also participated in in-person expressive language tasks, during which a native-signing adult showed a picture (representing a word on the ASL CDI 2.0) to children and asked them to provide the sign for that picture. Both deaf and hearing parents were found to be highly accurate reporters of their children's vocabulary, with correlations of 0.69-0.75 between parents' report and children's picture

naming; see Table S1.

Finally, in the following section we add a validation of the American English CDI for blind children (9–48 months old), which to our knowledge, has not been undertaken previously. For this validation, parents of 40 blind children completed the CDI, and 15 children also completed a daylong home language recording, via LENA (Gilkerson et al., 2017). Alongside an automated "Child Vocalization Count" generated from the entire daylong recordings, forty minutes of the home language recordings were transcribed by trained annotators, from which word types, word tokens, and canonical speech values were extracted (detailed in the next section). Correlations between Words Produced on the CDI and these four speech production measures ranged from 0.64-0.97; see Table S1.

To summarize, parents of a diverse set of children (including blind children, deaf signing children, and deaf children with cochlear implants) can accurately report on their children's productive vocabulary with CDI Word Production measures achieving correlations of $R = 0.55$–0.97 with other measures of children's productive language abilities.

Table 1

Summary of relevant CDI validation studies for blind children, deaf signing children, and deaf children with cochlear implants. See Fenson et al. (1994) for a summary of validation work in typically-hearing/sighted populations. W&G = Words and Gestures; W&S = Words and Sentences; RDLS = Reynell Developmental Language Scales.

				Validation	
Reference	**Form**	**Population**	**N**	**Method**	**Correlation**
Caselli, Lieberman, & Pyers, 2020	ASL CDI 2.0	deaf signing children of hearing parents	11	Picture naming	0.693

Reference	Form	Population	N	Validation Method	Correlation
Caselli, Lieberman, & Pyers, 2020	ASL CDI 2.0	deaf signing children of deaf parents	29	Picture naming	0.753
Thal, Desjardin, & Eisenberg, 2007	American English W&G	deaf children with cochlear implants	24	RDLS - Expressive	0.84
Thal, Desjardin, & Eisenberg, 2007	American English W&S	deaf children with cochlear implants	24	RDLS - Expressive	0.84
Thal, Desjardin, & Eisenberg, 2007	American English W&G	deaf children with cochlear implants	24	spontaneous language sample - word tokens	0.62
Thal, Desjardin, & Eisenberg, 2007	American English W&S	deaf children with cochlear implants	24	spontaneous language sample - word tokens	0.55
Thal, Desjardin, & Eisenberg, 2007	American English W&G	deaf children with cochlear implants	24	spontaneous language sample - word types	0.68
Thal, Desjardin, & Eisenberg, 2007	American English W&S	deaf children with cochlear implants	24	spontaneous language sample - word types	0.64
see below	American English W&G / W&S	blind children	15	spontaneous language sample - word tokens	0.88

Reference	Form	Population	N	Validation Method	Correlation
see below	American English W&G / W&S	blind children	15	spontaneous language sample - word types	0.88

Validation of the CDI for Blind Children

Methods

To measure concurrent validity of the CDI for blind children, we compared CDI scores from blind children with their spontaneous language in a daylong, at-home, spontaneous speech sample, collected via LENA recorders (Gilkerson et al., 2017). All data collection procedures were approved by the xxx University Institutional Review Board. These data are part of an ongoing larger project to analyze the home language environment of blind children.

Participants. All participants exhibited severe-to-profound visual impairment (defined as no more than minimal light perception), heard >75% English at home, and had no co-occurring auditory, cognitive, or developmental disabilities. We recruited participants from the United States and Canada via pediatric ophthalmology clinics, early intervention and preschool programs for blind children, social media, and word of mouth. There were N=15 participants (6.61-30.05 months, M: 15.21) who completed a CDI and a home recording, making them eligible for inclusion in the present analysis. These are a subset of the participants reported on the main manuscript.

Measures.

CDI. The validation analyses focused on the Words Produced measure of the American English CDI, in which parents are asked to report *"Does your child produce the [word]"* for 396 or 680 words, depending on whether they completed the Words & Gestures or Words & Sentences version of the instrument.

LENA. For the home recording, children wore a small recorder, tucked into the pocket of a wearable vest, to capture their language input and production. Caregivers 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 the vest nearby during baths, naps, and car rides). We extracted four measures of child language production from these daylong recordings, as detailed below.

Child Vocalization Count (CVC). The audio recordings were first processed by the LENA proprietary software (Xu, Yapanel, & Gray, 2009), which creates algorithm-generated measures such as Child Vocalization Count (CVC) from the entire daylong recording. For CVC, a child vocalization is counted when the algorithm identifies a vocalization as made by the child wearing the recorder, i.e. when target child vocalizations of any length are surrounded by at least 300 milliseconds of silence or non-speech. CVC does not count non-communicative target child vocalizations such as burps, cries, or non-target child vocalizations in this count. We extracted the CVC as an automated measure of children's vocal productivity.

Vocal Maturity. In addition to the automated LENA output, a subset of each recording was annotated using the ELAN software (Brugman & Russel, 2009). Trained annotators applied the ACLEW annotation scheme (Soderstrom et al., 2021) to 40 minutes of each recording, 10 2-minute clips selected at random throughout the day, and 5 2-minute clips selected for having a high volume of talk based on the LENA output. Each target child vocalization tagged in this process was manually tagged as containing Noncanonical speech, Canonical speech, Laughter, or Crying; vegetative sounds (e.g., burps) were not

counted. For more information about this scheme, see the ACLEW homepage. From this tagging scheme, we extracted the proportion of child vocalizations that contained canonical speech.

Word Counts. For all target-child clips that were identified to contain a word, a pair of trained annotators manually transcribed all words in the recordings. All clips were reviewed by both annotators, and disagreements were resolved by discussion. From these transcriptions, we tallied the number of total words (tokens) and unique words (types) produced by the child throughout the recording.

Results

To evaluate concurrent validity, we measured the associations between the number of words children produced on the CDI (logged to normalize distribution), and each of our four LENA measures of language production: CVC, proportion of canonical clips, word tokens, and word types. Using Pearson's correlations, we found strong positive correlations (Rs = 0.64-0.97, all $ps < .05$; see Figure 1) between the CDI production measure and all four LENA measures. Children who were reported to have a larger vocabulary on the CDI produced more vocalizations, a higher proportion of canonical speech, more words, and more unique words in the recordings; see Figure 1.

References

Bates, E., Bretherton, I., Snyder, L., Beeghly, M., Shore, C., McNew, S., ... et al. (1988). *From first words to grammar: Individual differences and dissociable mechanisms* (pp. xii, 326). New York, NY, US: Cambridge University Press.

Beeghly, M., Jernberg, E., & Burrows, E. (1989). *Validity of the Early Language Inventory (ELI) for use with 25-month-olds.* Kansas City, MO.

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

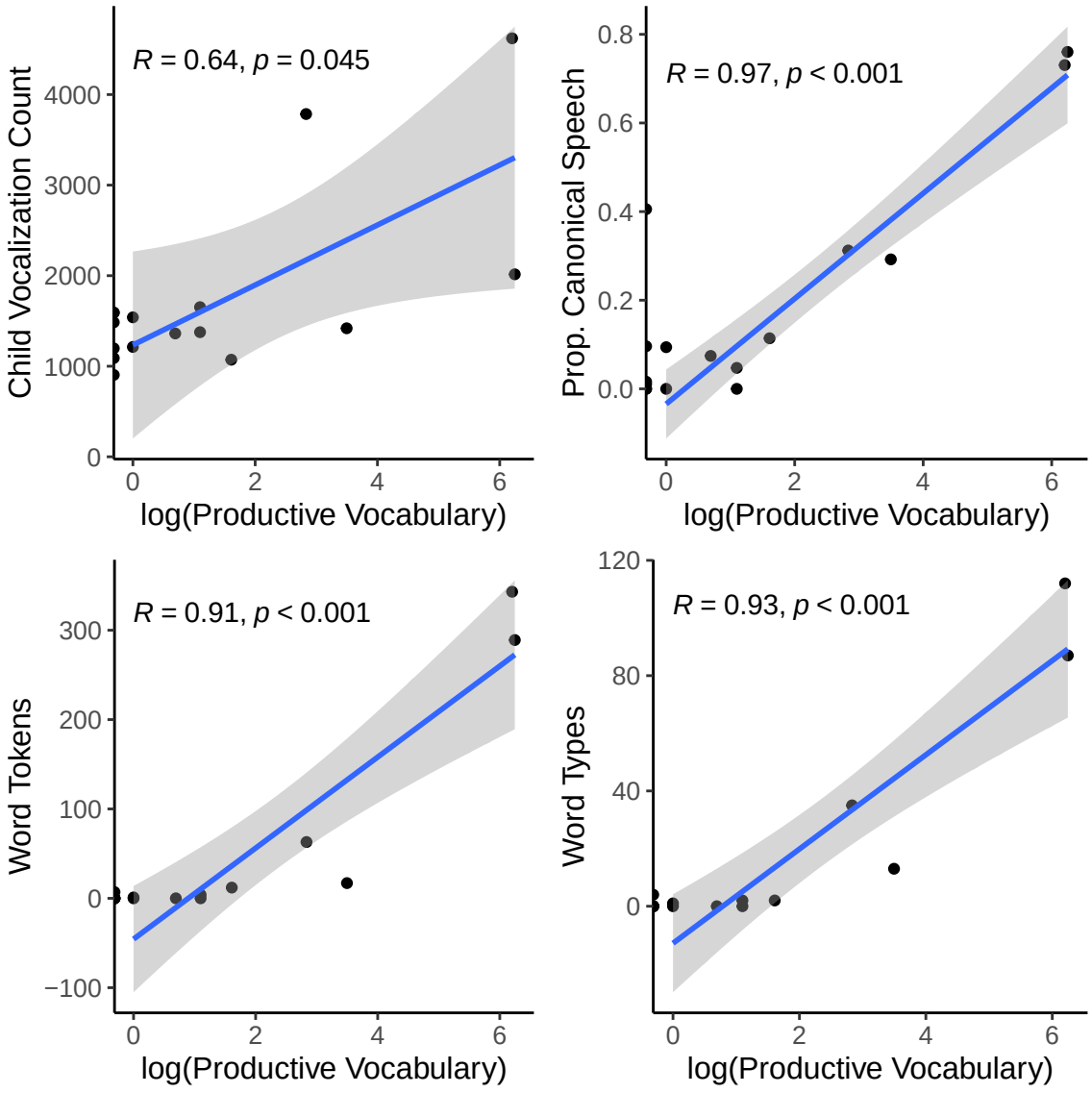

Figure 1. Correlations between CDI Words Produced and various LENA measures. Child Vocalization Count is the LENA algorithm's estimate, calculated over the entire recording. Word types, word tokens, and proportion of canonical vocalizations are derived from manual trancriptions of child speech from 40 minutes of the recordings.

Evaluation.

Caselli, N. K., Lieberman, A. M., & Pyers, J. E. (2020). The ASL-CDI 2.0: An updated, normed adaptation of the MacArthur Bates Communicative Development Inventory for American Sign Language. *Behavior Research Methods*, *52*(5), 2071–2084. https://doi.org/10.3758/s13428-020-01376-6

Dale, P. S. (1991). The Validity of a Parent Report Measure of Vocabulary and Syntax at 24 Months. *Journal of Speech, Language, and Hearing Research*, *34*(3), 565–571. https://doi.org/10.1044/jshr.3403.565

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

Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Kimbrough Oller, D., ... Paul, T. D. (2017). Mapping the Early Language Environment Using All-Day Recordings and Automated Analysis. *American Journal of Speech-Language Pathology*, *26*(2), 248–265. https://doi.org/10.1044/2016_AJSLP-15-0169

Jarůšková, L., Smolík, F., Chládková, K., Oceláková, Z., & Paillereau, N. (2023). How to Build a Communicative Development Inventory: Insights From 43 Adaptations. *Journal of Speech, Language, and Hearing Research*, *66*(6), 2095–2117. https://doi.org/10.1044/2023_JSLHR-22-00591

Marchman, V. A., & Dale, P. S. (2023). The MacArthur-Bates Communicative Development Inventories: Updates from the CDI Advisory Board. *Frontiers in Psychology*, *14*.

Rescorla, L., Ratner, N. B., Jusczyk, P., & Jusczyk, A. M. (2005). Concurrent Validity of the Language Development Survey: Associations With the MacArthur—Bates Communicative Development Inventories. *American Journal of Speech-Language Pathology*, *14*(2), 156–163. https://doi.org/10.1044/1058-0360(2005/016)

San Roque, L., Norcliffe, E., & Majid, A. (2024). Vision Verbs Emerge First in English Acquisition but Touch, not Audition, Follows Second. *Cognitive Science*, *48*(6), e13469. https://doi.org/10.1111/cogs.13469

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

Thal, D., Desjardin, J., & Eisenberg, L. S. (2007). Validity of the MacArthur–Bates Communicative Development Inventories for Measuring Language Abilities in Children With Cochlear Implants. *Article in American Journal of Speech-Language Pathology*, 54–64. https://doi.org/10.1044/1058-0360(2007/007)

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.