# Supplemental Appendix For:

## Informing Mothers about the Benefits of Conversing with Infants: Experimental Evidence from Ghana

*American Economic Journal: Economic Policy*

Pascaline Dupas, Camille Falezan, Seema Jayachandran, and Mark Walsh

# Appendix A

## A.1 Appendix Figures

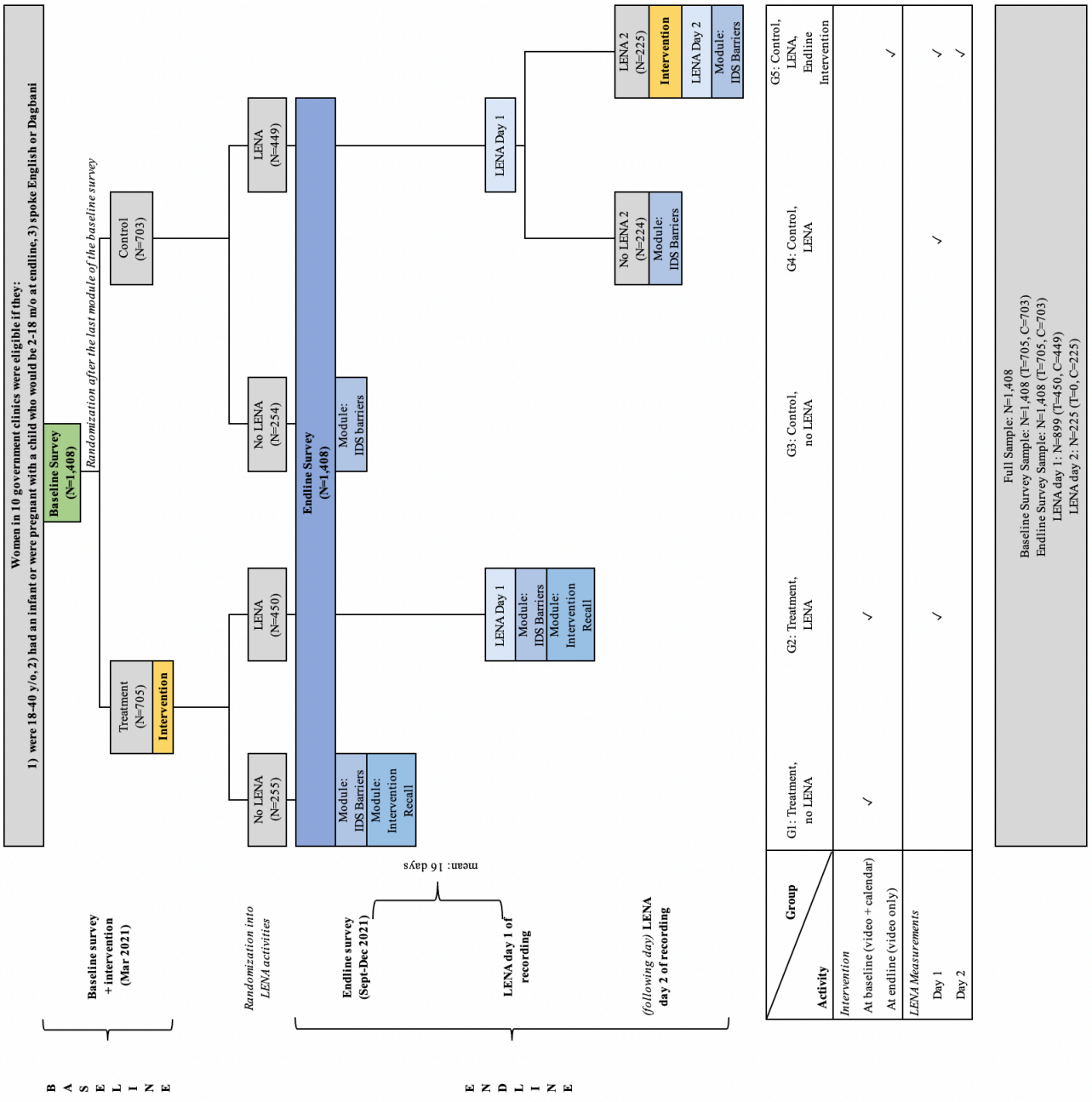
Figure A.1: Calendar for Treated Respondents


Note: 50% of the sample (N=705) watched the video and received an IDS-themed calendar at the end of the baseline survey. The calendar displays a star at the end of each week. Respondents were encouraged to fill in the stars next to each week in the calendar if they conversed with their child each day that week.

Figure A.2: Calendar for Control Respondents


Note: 50% of the sample (N=703) received a regular calendar at the end of the baseline survey as a token of gratitude for participating in the survey. Control respondents did not see the IDS-themed video.

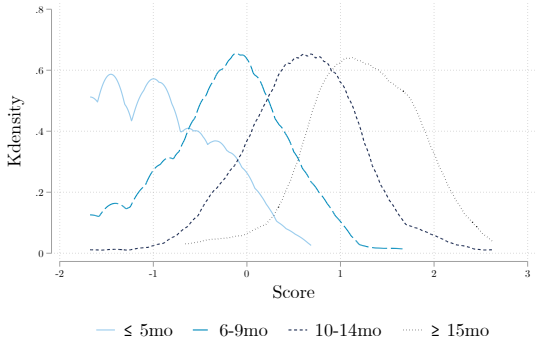
Figure A.3: Experimental Design and Timeline



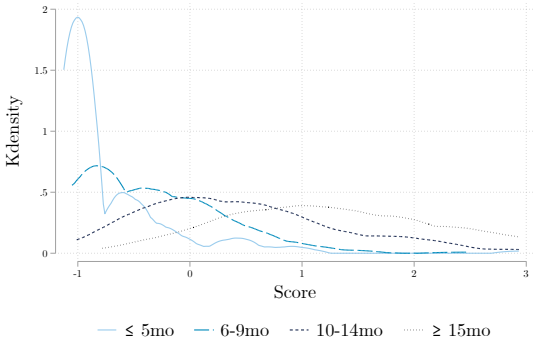

Note: See main text Section II for further details on the study design and timeline. On average, 6.4 months elapsed between the baseline and endline surveys.

Figure A.4: Distribution of Child Language and Development Scores (Control Group)

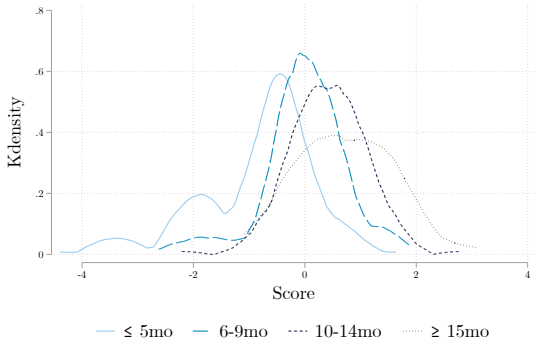
(a) Panel A: Reported Child Language Score



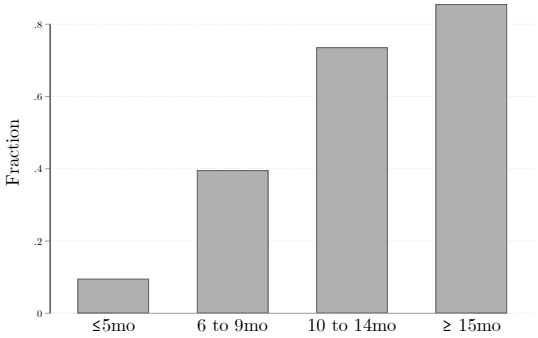

(b) Panel B: Reported Child Gestural Communication Index




(c) Panel C: Observed Child Development Index

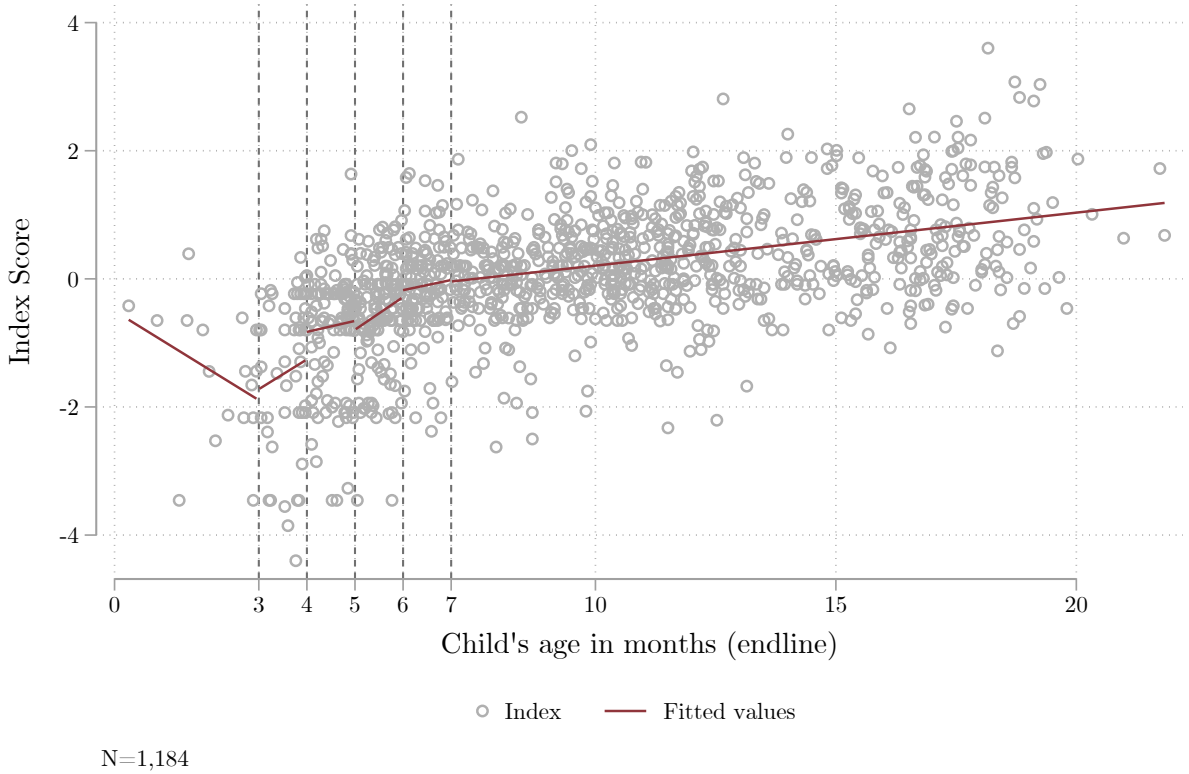



(d) Panel D: Observed Child Babbles (=1)



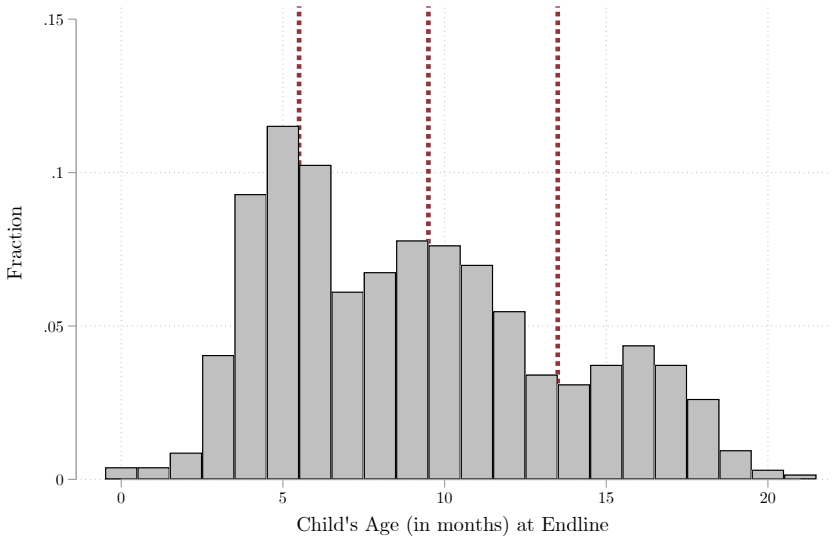

Note: Endline data. Distributions are plotted for children in the control group only. For details on outcomes, please refer to Table 4. Treatment effects for each age groups are shown in Table 5.

Figure A.5: Observed Child Development Index by Child Age




Note: Endline data. The index is an Anderson index, normalized over the control group. It is based on a selection of items adapted from the Ages and Stages Questionnaire (ASQ) and the Oxford Neurodevelopment Assessment (Ox-NDA). The assessment was administered by the surveyor to the child during the survey. See Table A.7 for the list of components included in the index.

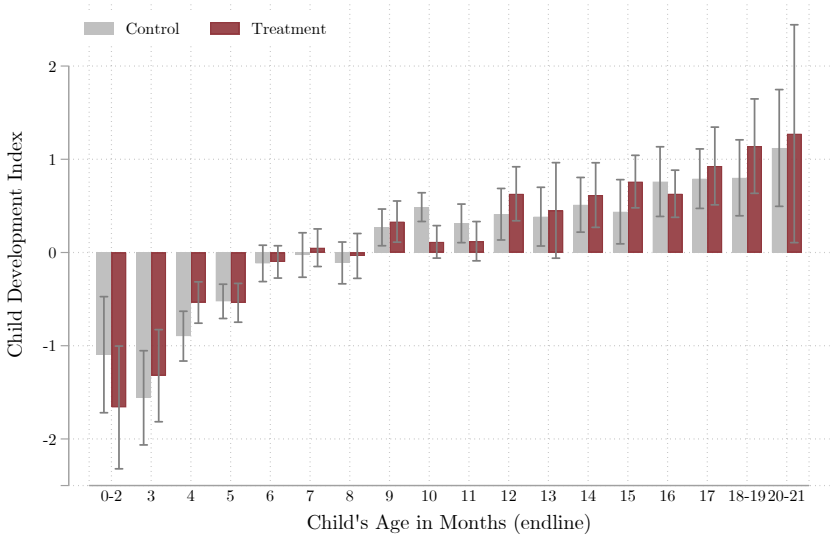
Figure A.6: Distribution of Children's Ages at Endline



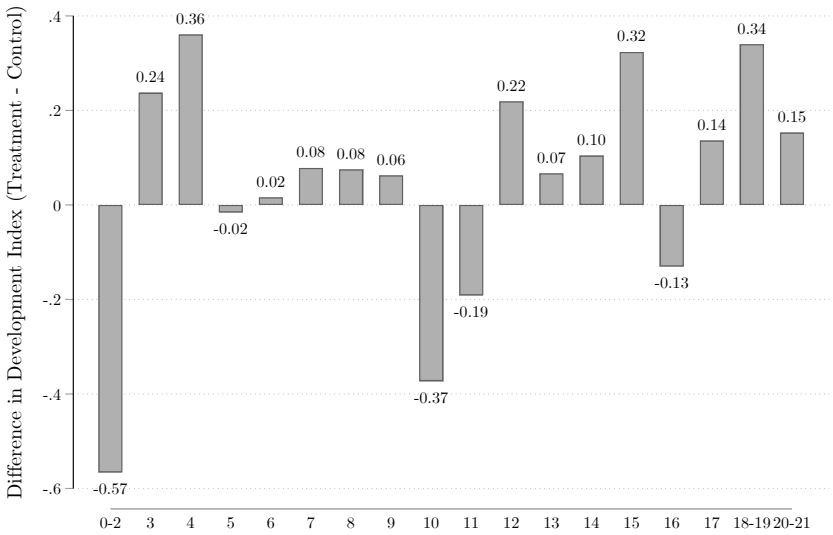

Note: Endline data. N=1,258/1,408 households participated in the endline survey. Average infant age is 9.6 months in the treatment group and 9.5 months in the control group. The dashed red lines delimit the 4 groups for which we present the disaggregated treatment effects in Table 5: 5mo or less, 6 to 9mo old, 10 to 14mo old, and 15 months or more.

Figure A.7: Observed Child Development Index by Child's Age (in Months)

(a) Panel A: Mean by Child's Age (in Months) and Group

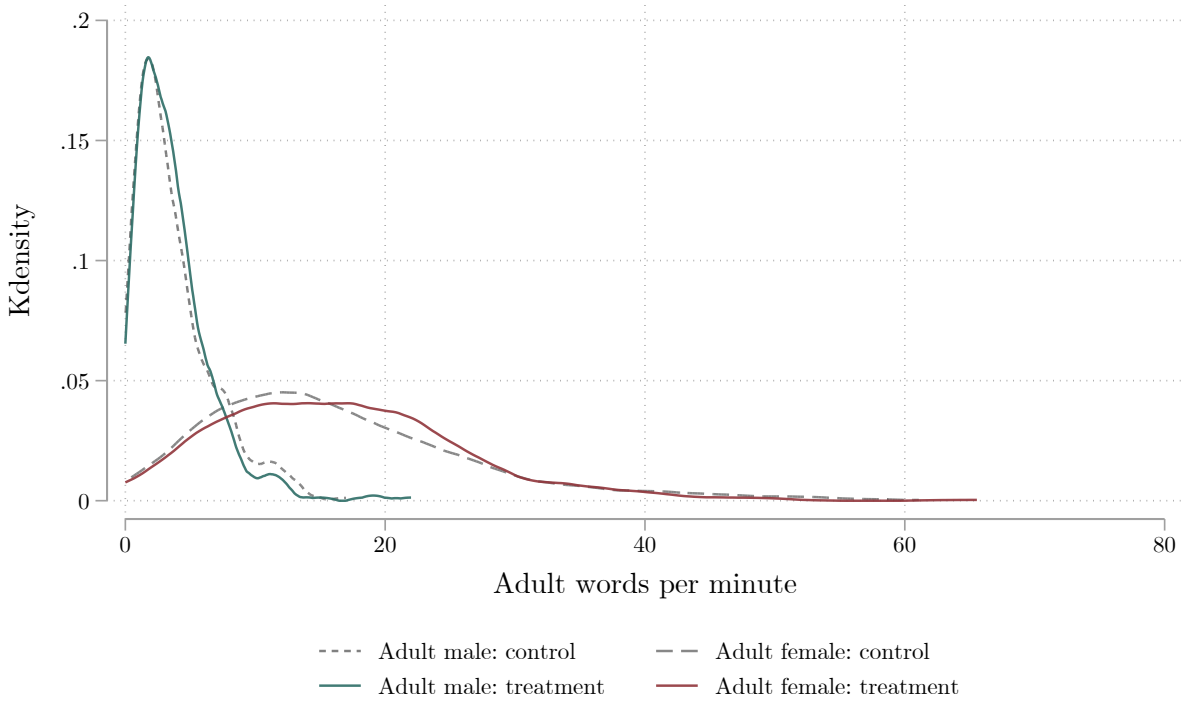



(b) Panel B: Difference in Means (Treatment - Control)




Note: Panel A: the bars show the control and treatment group means by age group with 95% confidence intervals. Panel B: the bars show the group difference in means (treatment minus control) by child's age. We pool children aged 0-2 months, 18-19 months, and 20-21 months as there are few observations in those groups (see Figure A.6 for children's age distribution at endline). The Observed child development index is an Anderson index, normalized over the control group. It is based on a selection of items from the Ages & Stages Questionnaire (ASQ) and the Oxford Neurodevelopment Assessment (Ox-NDA) (see Table A.8 for further details).

Figure A.8: LENA-measured Adult Words per Minute by Speaker Gender and Treatment

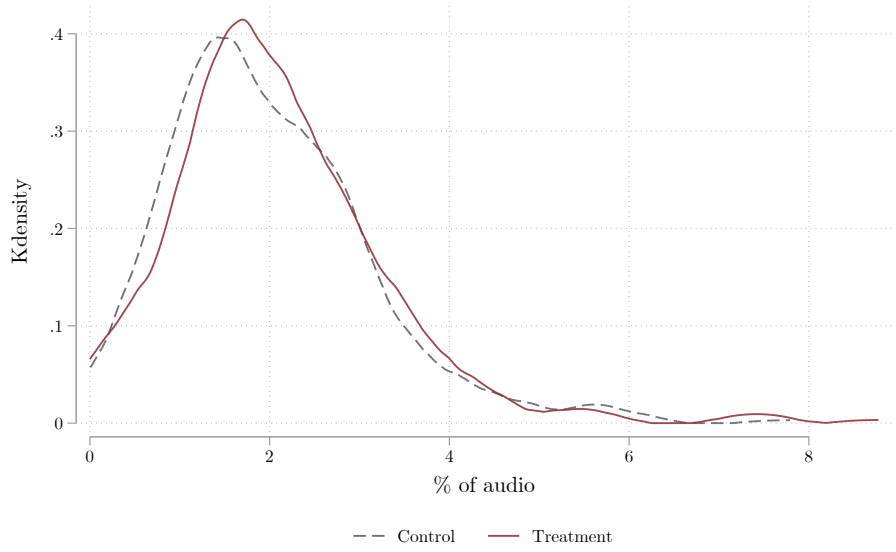



Exact p-values of the Kolmogorov–Smirnov two-sample tests for equality of distribution are 0.742 (adult male) and 0.342 (adult female)

Note: LENA day 1 recording data. N=774 recordings (1 recording does not have the breakdown of adult words by gender). Please refer to Table A.2 for details on the sample. The LENA software estimates the number of words spoken by post-pubescent males and females in the child's vicinity. Adult word per minute is the estimated total number of words spoken by adults during the recording divided by the length of the recording. For further details on the LENA outcomes, please refer to Section A.2.

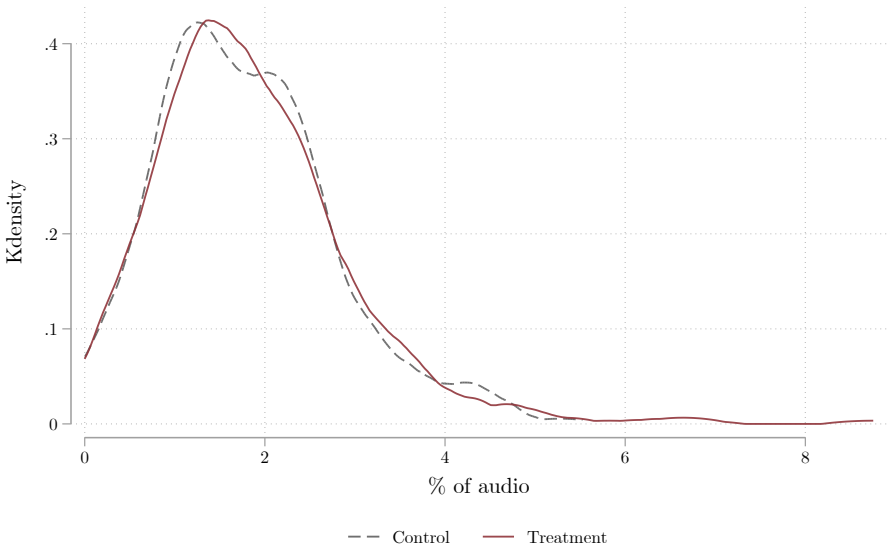
Figure A.9: LENA Measurements: Child Sounds by Treatment

(a) Panel A: Vocalizations (% of total audio time)




Exact p-value of the Kolmogorov–Smirnov two-sample test for equality of distribution is 0.169

(b) Panel B: Non-Speech Sounds (% of total audio time)




Exact p-value of the Kolmogorov–Smirnov two-sample test for equality of distribution is 0.936

Note: LENA day 1 data. N=775 recordings. Please refer to Table A.2 for details on the sample. The LENA software categorizes the focal child sound segments into 1) **vocalizations** (including words, babbles, and pre-speech communicative sounds or "protophones" such as squeals, growls, or raspberries) and 2) **non-speech** sounds (including fixed signals and vegetative sounds such as breathing or crying). For further details on the LENA outcomes, please refer to Section A.2.

## A.2 Appendix Tables

Table A.1: List of Government Health Facilities

| Name | District |
|---|---|
| Choggu RCH | Sagnarigu Municipal |
| Kalpohini Health Centre | Sagnarigu Municipal |
| Kanvilli Health Centre | Sagnarigu Municipal |
| Malshegu CHPS | Sagnarigu Municipal |
| Sagnarigu Health Centre | Sagnarigu Municipal |
| Bilpela Health Centre | Tamale Metropolitan |
| Moshie Zongo Health Centre | Tamale Metropolitan |
| Tamale Central Hospital | Tamale Metropolitan |
| Tamale SDA Hospital | Tamale Metropolitan |
| Tamale West Hospital | Tamale Metropolitan |

Note: List of health facilities in Tamale (Northern Ghana) where women were recruited when coming for prenatal or postnatal checkups.

Table A.2: Attrition and Endline Survey Status

| | All | | | | Control | | | Treatment | | | Treatment = Control |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Count | N | Mean | SD | Count | Mean | SD | Count | P-value |
| **Endline Survey** | | | | | | | | | | | |
| Dead | 0.00 | 0.03 | 1 | 1,408 | 0.00 | 0.04 | 1 | 0.00 | 0.00 | 0 | 0.317 |
| Had COVID symptoms | 0.00 | 0.04 | 2 | 1,408 | 0.00 | 0.00 | 0 | 0.00 | 0.05 | 2 | 0.157 |
| Refused to participate | 0.00 | 0.07 | 7 | 1,408 | 0.00 | 0.07 | 3 | 0.01 | 0.08 | 4 | 0.708 |
| Moved temporarily | 0.01 | 0.08 | 8 | 1,408 | 0.00 | 0.07 | 3 | 0.01 | 0.08 | 5 | 0.481 |
| Unavailable (other reason) | 0.01 | 0.12 | 20 | 1,408 | 0.02 | 0.12 | 11 | 0.01 | 0.11 | 9 | 0.648 |
| Ineligible | 0.01 | 0.12 | 21 | 1,408 | 0.01 | 0.10 | 7 | 0.02 | 0.14 | 14 | 0.125 |
| Moved permanently | 0.02 | 0.15 | 33 | 1,408 | 0.03 | 0.16 | 18 | 0.02 | 0.14 | 15 | 0.592 |
| Not found | 0.04 | 0.20 | 58 | 1,408 | 0.04 | 0.19 | 27 | 0.04 | 0.21 | 31 | 0.600 |
| Completed survey | 0.89 | 0.31 | 1,258 | 1,408 | 0.90 | 0.30 | 633 | 0.89 | 0.32 | 625 | 0.398 |
| Age of child at endline (months) | 9.58 | 4.41 | | 1,258 | 9.52 | 4.35 | | 9.64 | 4.47 | | 0.651 |
| Number of Respondents | | | | 1,258 | | | 633 | | | 625 | |
| **Child Assessment** | | | | | | | | | | | |
| Consented to child test | 1.00 | 0.07 | 1,252 | 1,258 | 1.00 | 0.06 | 631 | 0.99 | 0.08 | 621 | 0.406 |
| Child available (if consented) | 0.96 | 0.19 | 1,203 | 1,252 | 0.97 | 0.18 | 611 | 0.95 | 0.21 | 592 | 0.172 |
| **LENA Recording Day 1** | | | | | | | | | | | |
| Refusal (survey or LENA) | 0.02 | 0.15 | 22 | 899 | 0.03 | 0.17 | 14 | 0.02 | 0.13 | 8 | 0.194 |
| Not available/eligible main survey | 0.10 | 0.30 | 92 | 899 | 0.09 | 0.29 | 41 | 0.11 | 0.32 | 51 | 0.277 |
| Complete | 0.87 | 0.33 | 785 | 899 | 0.88 | 0.33 | 394 | 0.87 | 0.34 | 391 | 0.698 |
| Number of Respondents | | | | 899 | | | 449 | | | 450 | |
| If complete: kept in analysis | 0.99 | 0.11 | 775 | 785 | 0.99 | 0.10 | 390 | 0.98 | 0.12 | 385 | 0.517 |
| **LENA Recording Day 2** | | | | | | | | | | | |
| Missing/Lost | 0.00 | 0.07 | 1 | 225 | 0.00 | 0.07 | 1 | | | | |
| Refusal (survey or LENA) | 0.04 | 0.19 | 8 | 225 | 0.04 | 0.19 | 8 | | | | |
| Not available/eligible main survey | 0.11 | 0.31 | 24 | 225 | 0.11 | 0.31 | 24 | | | | |
| Complete | 0.85 | 0.35 | 192 | 225 | 0.85 | 0.35 | 192 | | | | |
| Number of Respondents | | | | 225 | | | 225 | | | | |
| If complete: kept in analysis | 0.98 | 0.14 | 188 | 192 | 0.98 | 0.14 | 188 | | | | |

Note: Endline data. Due to monetary constraints, only a sub-sample of respondents were randomized to receive a LENA device (N=900). A subsample of the control group was randomized to keep the LENA device for two days instead of only one (N=225). Before the start of the second day of recording, those respondents were shown the intervention video (see Section II and Figure A.3 for further details on the study design and timeline). In the panels "LENA Day 1" and "LENA Day 2", "If complete: kept in analysis" is a dummy equal to 1 if the audio has 9 hours (rounded up) of recording between 10 a.m. and 7 p.m., and, hence, is kept in the analysis.

Table A.3: Reported Barriers to IDS

| | (1) Pure Control | (2) Main Intervention (6-8 mo ago) | (3) Endline Intervention (day before) | (4) P-value Main Intervention = Endline Intervention | (5) P-value Control = Endline Intervention |
|---|---|---|---|---|---|
| *=1 if it could be a barrier to other families* | | | | | |
| It's hard to remember/make a habit | 0.35 | 0.35 | 0.29 | 0.130 | 0.185 |
| It's mocked/frowned upon in the community | 0.32 | 0.30 | 0.28 | 0.661 | 0.311 |
| It's not clear it makes a difference | 0.28 | 0.20 | 0.07 | <0.001 | <0.001 |
| Too busy/Not enough time | 0.08 | 0.06 | 0.01 | <0.001 | <0.001 |
| Parents are too preocuppied or unhappy | 0.01 | 0.02 | 0.01 | 0.417 | 0.884 |
| Other | 0.02 | 0.00 | 0.00 | 0.083 | 0.003 |
| Lack of patience | 0.00 | 0.01 | 0.01 | 0.915 | 0.478 |
| Laziness | 0.01 | 0.00 | 0.00 | 0.157 | 0.083 |
| Child may grow to be disrespectful | 0.00 | 0.00 | 0.00 | 0.083 | 0.318 |
| Parent's personality: shy, not talkative | 0.00 | 0.00 | 0.00 | 0.318 | 0.318 |
| Lack of reaction/responsiveness from the child | 0.00 | 0.00 | 0.00 | 0.318 | 0.318 |
| No barriers to IDS cited | 0.43 | 0.46 | 0.48 | 0.660 | 0.251 |
| *=1 if could be the main barrier to other families* | | | | | |
| It's hard to remember/make a habit | 0.16 | 0.19 | 0.18 | 0.828 | 0.448 |
| It's mocked/frowned upon in the community | 0.18 | 0.17 | 0.29 | 0.002 | 0.003 |
| It's not clear it makes a difference | 0.08 | 0.07 | 0.02 | 0.002 | <0.001 |
| Too busy/Not enough time | 0.06 | 0.04 | 0.01 | <0.001 | <0.001 |
| Parents are too preocuppied or unhappy | 0.01 | 0.01 | 0.01 | 0.915 | 0.688 |
| Other barrier (specify) | 0.13 | 0.11 | 0.03 | <0.001 | <0.001 |
| *=1 if it's a barrier to respondent and her family* | | | | | |
| It's hard to remember/make a habit | | | 0.31 | | |
| It's mocked/frowned upon in the community | | | 0.32 | | |
| It's not clear it makes a difference | | | 0.16 | | |
| Parents are too preocuppied or unhappy | | | 0.02 | | |
| No barriers to IDS cited | | | 0.37 | | |
| Observations | 424 | 615 | 191 | | |

Note: Endline data. Respondents were asked about barriers that may prevent families from talking to their babies. Questions were asked at the end of the endline survey if the household did not receive a LENA device, or after the last day of recording if the household received a LENA (see timing of the "Module: IDS Barriers" in the design chart Figure A.3). The endline intervention sample received the intervention between the 1st and 2nd day of LENA recording and the IDS barrier questions were asked after the second day of recording. Column 1 presents the means for respondents in the pure control group (who never received the intervention), column 2 those for respondents who received the main intervention, at baseline (their views incorporate their experience with IDS over the past 6 to 8 months between the intervention and the endline survey), and column 3 for respondents who received the intervention at endline (their views incorporate their experience with IDS over the past 24 hours). The last two columns report the p-values from t-tests comparing the means between the respondents who received the intervention at baseline vs at endline (column 4) and respondents who did not receive the intervention (pure control group) vs those who received it at endline. Questions in the last panel "=1 if it's a barrier to respondent and her family" were only asked to those who received the endline intervention.

Table A.4: Treatment Effects on Mother-reported Parental Beliefs

| | Treatment Effect | | | Control Group | | |
|---|---|---|---|---|---|---|
| | Coefficient | SE | P-value | Mean | SD | N |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Believes should talk to child from birth | 0.104 | 0.027 | 0.000 | 0.33 | 0.47 | 1,257 |
| ***Outcomes in the index*** | | | | | | |
| Age (in mo) when babbles/makes noise in response | -0.145 | 0.483 | 0.764 | 7.37 | 8.11 | 1,256 |
| Age (in mo) when says meaningful words | -1.638 | 0.638 | 0.010 | 19.96 | 12.22 | 1,248 |
| Age (in mo) for talking to child | -1.170 | 0.486 | 0.016 | 5.55 | 7.94 | 1,257 |
| Age (in mo) for telling stories to child | -2.530 | 0.858 | 0.003 | 18.30 | 15.90 | 1,228 |
| Age (in mo) for talking to child in full sentences | -1.507 | 1.346 | 0.263 | 26.00 | 24.98 | 1,250 |
| Importance to brain development of talking in full sentences to a child (/10) | 0.024 | 0.112 | 0.832 | 8.72 | 2.02 | 1,252 |
| *How strongly do you agree with: (1=strongly disagree to 4= strongly agree)* | | | | | | |
| Intelligence is set at birth | 0.053 | 0.071 | 0.451 | 3.08 | 1.27 | 1,250 |
| Infants learn little language in their 1st year | -0.005 | 0.053 | 0.931 | 3.54 | 0.94 | 1,255 |
| Parents shouldn't talk back to babble | 0.002 | 0.074 | 0.981 | 3.15 | 1.31 | 1,257 |
| Children learn more from overhearing than being spoken to | -0.100 | 0.060 | 0.094 | 3.49 | 1.00 | 1,249 |
| Adults can't have conversations with babies who can't talk | -0.097 | 0.074 | 0.191 | 2.03 | 1.36 | 1,253 |

Note: Each line reports the result of a different regression for which the outcome (variable indicated on the left) is regressed on a dummy equal to 1 if the household received the main intervention, at baseline (treatment group). Column 1 reports the coefficient on treatment, column 2 the standard error, and column 3 the p-value. Columns 4 and 5 report the control group mean and standard deviation. Column 6 reports the number of observations. "Believes should talk to child from birth" is not included in the mother's belief index in Table 4, but all other outcomes are. For outcomes in the panel "How strongly do you agree with the following statements:", respondents were asked to choose from a 4-point Likert scale (strongly disagree (1), somewhat disagree, somewhat agree, strongly agree (4)). As for the main text Table 4, all regressions include clinic fixed effects and controls for the child's age (in days), survey date, and surveyor gender.

Table A.5: Treatment Effects on Mother-reported Parental Behavior (Verbal Inputs)

| | Treatment Effect | | | Control Group | | |
|---|---|---|---|---|---|---|
| | Coefficient | SE | P-value | Mean | SD | N |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ***Outcomes in the index*** | | | | | | |
| *In the last 4 wks, how often did you: (0=never to 5=daily)* | | | | | | |
| Talk to child while doing an activity w/ child around | 0.165 | 0.103 | 0.109 | 2.05 | 1.88 | 1,256 |
| Describe things to child when walking | 0.152 | 0.096 | 0.112 | 2.11 | 1.76 | 1,256 |
| Pointed, named object and asked child to repeat | 0.141 | 0.096 | 0.142 | 1.65 | 1.85 | 1,256 |
| *In the last 4 weeks, did any adult:* | | | | | | |
| Sang to child | -0.012 | 0.019 | 0.521 | 0.88 | 0.32 | 1,254 |
| Read to/looked at book with child | 0.065 | 0.028 | 0.019 | 0.44 | 0.50 | 1,256 |
| Told story to child | 0.045 | 0.026 | 0.089 | 0.31 | 0.46 | 1,251 |
| Played with child | 0.005 | 0.007 | 0.500 | 0.98 | 0.14 | 1,256 |
| Decribed things to child | 0.017 | 0.025 | 0.489 | 0.69 | 0.46 | 1,257 |
| *As percent of total play time:* | | | | | | |
| % of time playing w/ adult | 0.921 | 0.903 | 0.308 | 31.37 | 16.29 | 1,258 |

Note: Each line reports the result of a different regression for which the outcome (variable indicated on the left) is regressed on a dummy equal to 1 if the household received the main intervention, at baseline (treatment group). Column 1 reports the coefficient on treatment, column 2 the standard error, and column 3 the p-value. Columns 4 and 5 report the control group mean and standard deviation. Column 6 reports the number of observations. For outcomes in the panel "In the last 4 weeks, how often did you...", respondents were asked to choose from a 6-point Likert scale (never (0), rarely, a few times, once a week, multiple times a week, daily (5)). As for the main text Table 4, all regressions include clinic fixed effects and controls for the child's age (in days), survey date, and surveyor gender.

Table A.6: Treatment Effects on Mother-reported Child Language Score Components

| | Treatment Effect | | | Control Group | | |
|---|---|---|---|---|---|---|
| | Coefficient | SE | P-value | Mean | SD | N |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| # of words in list child understands | 0.475 | 0.180 | 0.008 | 6.95 | 4.99 | 1,258 |
| # of words in list child says | 0.305 | 0.124 | 0.014 | 1.04 | 2.42 | 1,258 |
| # of phrases in list child understands | 0.096 | 0.055 | 0.078 | 1.40 | 1.44 | 1,258 |
| Attempt to say words (yes/no) | 0.044 | 0.021 | 0.038 | 0.42 | 0.49 | 1,257 |
| *How often does child: (1=not yet to 3=often)* | | | | | | |
| Repeat/imitate words | 0.046 | 0.032 | 0.154 | 1.49 | 0.76 | 1,258 |
| Name/label things | 0.016 | 0.025 | 0.535 | 1.18 | 0.50 | 1,256 |

Note: Each line reports the result of a different regression for which the outcome (variable indicated on the left) is regressed on a dummy equal to 1 if the household received the main intervention, at baseline (treatment group). Column 1 reports the coefficient on treatment, column 2 the standard error, and column 3 the p-value. Columns 4 and 5 report the control group mean and standard deviation. Column 6 reports the number of observations. The Child language score is computed using questions derived from a version of the MacArthur-Bates Communicative Development Inventories Words and Gestures (MB-CDI-WG) adapted to Ghana by the Harvard Laboratory for Developmental Studies (Duflo et al., 2024). The adapted inventory includes a list of words and sentences across various domains. Mothers were asked if the child understands and/or can say (either spontaneously or upon prompting) the items in this list: *words*: "Ouch (wa aish)", "Toy (biebi)", "Bread (paa nu)", "Toffee (toffe)", "Biscuit (biscuit)", "Shoe (namda)", "Chair (ku'g)", "Cup (copu)", "Feed (dima/dibu)", "Beautiful (vilem)", "Egg (galili)", "Baby (bilegu)", "Bath (kom subu)", "Don't (dining)", "Shh (shh)", "Thank you (npahiya)", *phrases*: "Give me a hug (awa wa tuu)", "Go get _ (chamtik-pahi)", "Good girl/boy (bi sung)", and "Spit it out (tuhi bahi)". Additionally, mothers were asked questions about whether the child started to talk. For the child language score, we use binary versions of the variables presented above (dummies for each word/phrase indicating whether the child understands/says it and for whether the child attempts to speak, to imitate words, and to name/label things). We compute the child language score using Item Response Theory which involves estimating a one-parameter logistic model on the mother's responses to the adapted MB-CDI-WG, where the model assigns a difficulty level to each question and, then, a latent trait to each individual based on their answers to the questions adjusting for the question's difficulty level. As for the main text Table 4, all regressions include clinic fixed effects and controls for the child's age (in days), survey date, and surveyor gender.

Table A.7: Treatment Effects on Mother-reported Child Gestural Communication

| | Treatment Effect | | | Control Group | | |
|---|---|---|---|---|---|---|
| | Coefficient | SE | P-value | Mean | SD | N |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ***Outcomes in the index*** | | | | | | |
| *How often does child: (1=not yet to 3=often)* | | | | | | |
| Give toy when holding it | 0.057 | 0.038 | 0.133 | 1.82 | 0.79 | 1,256 |
| Point at interesting things | 0.022 | 0.040 | 0.585 | 1.61 | 0.84 | 1,258 |
| Wave when someone leaves | 0.026 | 0.035 | 0.446 | 1.60 | 0.82 | 1,257 |
| Shake head for no | 0.081 | 0.038 | 0.035 | 1.72 | 0.84 | 1,258 |
| Gesture shh | 0.049 | 0.029 | 0.096 | 1.22 | 0.55 | 1,257 |
| Blows kisses | 0.044 | 0.034 | 0.195 | 1.40 | 0.65 | 1,257 |

Note: Each line reports the result of a different regression for which the outcome (variable indicated on the left) is regressed on a dummy equal to 1 if the household received the main intervention, at baseline (treatment group). Column 1 reports the coefficient on treatment, column 2 the standard error, and column 3 the p-value. Columns 4 and 5 report the control group mean and standard deviation. Column 6 reports the number of observations. Components of the Mother-reported child gestural communication index are based on a selection of items from the "First Communicative Gestures" from the MB-CDI-WG. Respondents were asked to choose from a 3-point Likert scale (not yet (1), sometimes, and often (3)). As for the main text Table 4, all regressions include clinic fixed effects and controls for the child's age (in days), survey date, and surveyor gender.

Table A.8: Treatment Effect on Observed Child Development

| | Treatment Effect | | | Control Group | | |
|---|---|---|---|---|---|---|
| | Coefficient | SE | P-value | Mean | SD | N |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ***Outcomes in the index*** | | | | | | |
| *Child assessment: (1=worst to 3=best)* | | | | | | |
| Watches mother move | 0.031 | 0.032 | 0.326 | 2.78 | 0.59 | 1,172 |
| Watches toy placed in front | 0.009 | 0.023 | 0.706 | 2.90 | 0.42 | 1,179 |
| *Child assessment: (1=worst to 4=best)* | | | | | | |
| Identifies spoon correctly when asked | -0.034 | 0.059 | 0.562 | 2.25 | 1.04 | 1,141 |
| Imitates or tries to imitate bi-syllabic words | -0.010 | 0.035 | 0.763 | 1.28 | 0.62 | 1,132 |
| Reacts to name when playing | 0.091 | 0.071 | 0.202 | 2.74 | 1.27 | 1,174 |
| Stops reaching for toy when told no | -0.014 | 0.061 | 0.823 | 1.76 | 1.09 | 1,116 |
| Uses or mimics words in play context | 0.007 | 0.029 | 0.814 | 1.19 | 0.49 | 1,139 |
| Babbles or attempts to when prompted | 0.058 | 0.046 | 0.207 | 1.40 | 0.81 | 1,142 |
| Combines word and gesture (correctly or not) | -0.043 | 0.046 | 0.351 | 1.45 | 0.86 | 1,139 |

Note: Each line reports the result of a different regression for which the outcome (variable indicated on the left) is regressed on a dummy equal to 1 if the household received the main intervention, at baseline (treatment group). Column 1 reports the coefficient on treatment, column 2 the standard error, and column 3 the p-value. Columns 4 and 5 report the control group mean and standard deviation. Column 6 reports the number of observations. The assessment was administered to the child by the surveyor during the survey. The first two tasks were adapted from the problem-solving ASQ-3 (2 months) module and the others were adapted from the language and cognitive Ox-NDA modules. For the ASQ-like items, each task was evaluated from 1 (lowest score) to 3 (highest score). For the Ox-NDA-like items, each task was evaluated using a scale from 1 (lowest score) to 4 (highest score). Instead of giving a score, surveyor could indicate they were unable to assess the child (because the infant was out of sight, sleeping, crying, became too agitated, etc.) and those are recoded as missing. As for the main text Table 4, all regressions include clinic fixed effects and controls for the child's age (in days), survey date, and surveyor gender.

Table A.9: Robustness of Treatment Effects, 6 to 8 Months After Intervention

| | Mother's Interview | | | | Observed | | LENA | |
|---|---|---|---|---|---|---|---|---|
| | Mother's belief index | Mother's behavior index | Child language score | Child gestural communication index | Child development index | Child babbles (= 1) | Female adult words per minute | Child vocalizations per minute |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Panel A: Without Controls | | | | | | | | |
| Treatment | 0.125 | 0.126 | 0.115 | 0.108 | 0.051 | 0.054 | -0.008 | 0.047 |
| | (0.060) | (0.056) | (0.054) | (0.055) | (0.056) | (0.033) | (0.720) | (0.060) |
| | {0.036} | {0.026} | {0.034} | {0.051} | {0.368} | {0.098} | {0.991} | {0.428} |
| Control mean | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.51 | 16.61 | 1.53 |
| Controls | No | No | No | No | No | No | No | No |
| Clinic FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,258 | 1,258 | 1,258 | 1,258 | 1,184 | 888 | 774 | 775 |
| Panel B: With Clinic-Day Fixed Effects | | | | | | | | |
| Treatment | 0.137 | 0.133 | 0.108 | 0.103 | 0.034 | 0.061 | -0.037 | 0.035 |
| | (0.058) | (0.055) | (0.035) | (0.041) | (0.049) | (0.029) | (0.733) | (0.061) |
| | {0.018} | {0.016} | {0.002} | {0.012} | {0.491} | {0.035} | {0.959} | {0.570} |
| Control mean | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.51 | 16.61 | 1.53 |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clinic FE | No | No | No | No | No | No | No | No |
| Clinic-Day FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,258 | 1,258 | 1,258 | 1,258 | 1,184 | 888 | 774 | 775 |
| Panel C: With Surveyor Fixed Effects | | | | | | | | |
| Treatment | 0.180 | 0.119 | 0.085 | 0.073 | 0.013 | 0.032 | -0.044 | 0.040 |
| | (0.054) | (0.051) | (0.031) | (0.039) | (0.047) | (0.027) | (0.710) | (0.060) |
| | {0.001} | {0.021} | {0.006} | {0.059} | {0.783} | {0.232} | {0.951} | {0.508} |
| Control mean | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.51 | 16.61 | 1.53 |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clinic FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Surveyor FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,258 | 1,258 | 1,258 | 1,258 | 1,184 | 888 | 774 | 775 |

Note: Endline and LENA day 1 recording data. This table presents the results from running the main specification (presented in Table 4) without any control (Panel A); replacing clinic FE by clinic-day fixed effects (Panel B); and adding surveyor fixed effects (Panel C). Since 3 surveyors completed 20 or fewer endline surveys each, we grouped them as one surveyor when including Surveyor Fixed Effects (in Panel C). Robust standard errors in parenthesis, p-values in curly brackets.

Table A.10: Treatment Effects on Observed Child and Parental Behavior (LENA Outcomes)

| | Treatment Effect | | | Control Group | | |
|---|---|---|---|---|---|---|
| | Coefficient | SE | P-value | Mean | SD | N |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ***Count per minute*** | | | | | | |
| Adult words per minute | 0.070 | 0.786 | 0.929 | 20.37 | 11.75 | 775 |
| Female adult words per minute | 0.131 | 0.697 | 0.851 | 16.61 | 10.39 | 774 |
| Male adult words per minute | -0.058 | 0.215 | 0.788 | 3.76 | 3.00 | 774 |
| Focal child vocalizations per minute | 0.052 | 0.058 | 0.377 | 1.53 | 0.82 | 775 |
| Conversational turns per minute | 0.004 | 0.017 | 0.815 | 0.41 | 0.25 | 775 |
| ***Length as % of audio*** | | | | | | |
| Female adult speech | 0.094 | 0.302 | 0.754 | 7.32 | 4.49 | 774 |
| Male adult speech | 0.021 | 0.102 | 0.835 | 1.88 | 1.42 | 774 |
| Other children speech | 0.133 | 0.153 | 0.383 | 3.46 | 2.13 | 774 |
| Focal child sounds | 0.163 | 0.148 | 0.272 | 4.14 | 2.03 | 774 |
| Focal child vocalizations | 0.090 | 0.082 | 0.274 | 2.00 | 1.14 | 774 |
| Focal child non-vocalizations | 0.061 | 0.074 | 0.412 | 1.79 | 0.98 | 774 |
| Faint or overlapping speech | 0.750 | 1.108 | 0.499 | 39.07 | 15.67 | 775 |
| Silence | -0.336 | 1.120 | 0.764 | 29.79 | 15.31 | 775 |
| Background noise | -1.245 | 0.651 | 0.056 | 19.32 | 9.63 | 775 |
| Electronic media (TV, radio, etc.) | 0.157 | 0.663 | 0.813 | 9.37 | 10.39 | 775 |

Note: LENA day 1 recording data. Each line reports the result of a different regression for which the outcome (variable indicated on the left) is regressed on a dummy equal to 1 if the household received the main intervention, at baseline (treatment group). Column 1 reports the coefficient on treatment, column 2 the standard error, and column 3 the p-value. Columns 4 and 5 report the control group mean and standard deviation. Column 6 reports the number of observations. Given financial constraints, only a random subset of households could be included in the LENA measurement (N=900 households sampled). For households sampled to keep the LENA device for two days, only the first day recording is kept in the analysis. The analysis is further restricted to recording times between 10 a.m. to 7 p.m. (this excludes 10/785 LENA day 1 recordings which have less than 9 hours (rounded up) of recording). Some variables are missing for one recording. The focal child sounds as well as adults' and other children's speech only include sounds from those sources categorized as near by the LENA algorithm. Faint or overlapping speech includes faint female and male adults' sounds, other children's vocalizations, and overlapping speech (both near and faint). Background noise and electronic media each include both near and far sounds from those sources. As in columns 7 and 8 of main text Table 4, regressions include clinic fixed effects and controls for the child's age in days, the day of the week the audio was recorded (dummies), the total time (min) the shirt/LENA device was removed from the child, the total time (min) the child was held on someone's back while wearing the device, and the household size. For further details on the LENA outcomes, please refer to Section A.2.

Table A.11: Treatment Effect Split by Susceptibility to Experimenter Demand Effects

| | Mother's Interview | | | | LENA |
|---|---|---|---|---|---|
| | Mother's belief index (1) | Mother's behavior index (2) | Child language score (3) | Child gestural communication index (4) | Female adult words per minute (5) |
| Treatment | 0.181 | 0.143 | 0.112 | 0.104 | -0.344 |
| | (0.060) | (0.059) | (0.037) | (0.044) | (0.745) |
| | {0.003} | {0.015} | {0.003} | {0.019} | {0.645} |
| Treatment x did not associate intervention w/ surveyor | -0.260 | -0.033 | -0.030 | -0.013 | 2.498 |
| | (0.121) | (0.103) | (0.052) | (0.069) | (1.124) |
| | {0.032} | {0.749} | {0.566} | {0.850} | {0.027} |
| P-val total effect for did not associate | 0.509 | 0.278 | 0.101 | 0.172 | 0.051 |
| Control mean | -0.00 | -0.00 | -0.00 | 0.00 | 16.61 |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Clinic FE | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,248 | 1,248 | 1,248 | 1,248 | 773 |
| Observations: did not associate | 131 | 131 | 131 | 131 | 72 |

Note: Endline and LENA day 1 recording data. See main text Table 4 for details on specifications and outcomes. "Did not associate the intervention with survey" is a dummy equal to 1 if the respondent mentioned neither the video nor the calendar when asked about the baseline survey (without prompting). The dummy is always equal to 0 for control respondents as they did not receive the main intervention and were not asked those questions. N=1,248 instead of1,258 because 10/625 treatment respondents who consented to the endline survey did not reach the intervention recall module at the end of the endline activities, hence are dropped from the sample in this table. See main text Table 3 for further details on recall questions and sample. Robust standard errors in parentheses, p-values in curly brackets.

Table A.12: Robustness of Treatment Effects to Double Lasso Approach

| | Mother's Interview | | | | Observed | | LENA | |
|---|---|---|---|---|---|---|---|---|
| | Mother's belief index | Mother's behavior index | Child language score | Child gestural communication index | Child development index | Child babbles (= 1) | Female adult words per minute | Child vocalizations per minute |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Treatment | 0.187 | 0.106 | 0.082 | 0.066 | 0.026 | 0.040 | -0.085 | 0.066 |
| | (0.051) | (0.051) | (0.030) | (0.038) | (0.048) | (0.027) | (0.703) | (0.060) |
| | {0.000} | {0.040} | {0.007} | {0.083} | {0.593} | {0.135} | {0.904} | {0.273} |
| Control mean | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.51 | 16.61 | 1.53 |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clinic FE | No | No | No | No | No | No | No | No |
| Observations | 1,258 | 1,258 | 1,258 | 1,258 | 1,184 | 888 | 774 | 775 |

Note: Endline and LENA day 1 recording data. We use the double Lasso approach of Belloni, Chernozhukov and Hansen (2013) as implemented by Ahrens, Hansen and Schaffer (2019) to flexibly choose control variables for each regression. See main text Table 4 for further details on outcomes. Robust standard errors in parenthesis, p-values in curly brackets.

## Table A.13: Surveyors: Balance Check and Influence on Outcomes

| | Mother's Interview | | | | Observed | | LENA | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mother's belief index | Mother's behavior index | Child language score | Child gestural communication index | Child development index | Child babbles (= 1) | Female adult words per min | Child vocalizations per minute | Treatment |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Surveyor 2 | 0.650 | 0.519 | 0.128 | 0.200 | 0.025 | 0.002 | -0.771 | -0.287 | -0.028 |
| | (0.134) | (0.165) | (0.086) | (0.096) | (0.170) | (0.083) | (1.922) | (0.162) | (0.083) |
| | {0.000} | {0.002} | {0.139} | {0.038} | {0.883} | {0.984} | {0.689} | {0.076} | {0.735} |
| Surveyor 3 | 0.781 | -1.128 | -0.442 | -0.406 | -0.061 | -0.014 | 0.999 | -0.140 | -0.031 |
| | (0.147) | (0.103) | (0.072) | (0.091) | (0.114) | (0.064) | (1.760) | (0.145) | (0.067) |
| | {0.000} | {0.000} | {0.000} | {0.000} | {0.593} | {0.830} | {0.570} | {0.333} | {0.642} |
| Surveyor 4 | 0.081 | -0.223 | -0.228 | -0.162 | 0.059 | 0.068 | 1.672 | -0.028 | 0.032 |
| | (0.151) | (0.108) | (0.075) | (0.095) | (0.116) | (0.065) | (1.712) | (0.147) | (0.064) |
| | {0.594} | {0.038} | {0.003} | {0.090} | {0.611} | {0.297} | {0.329} | {0.849} | {0.623} |
| Surveyor 5 | -0.845 | 0.575 | 1.137 | 0.925 | 0.103 | -0.042 | 1.545 | -0.272 | -0.005 |
| | (0.143) | (0.167) | (0.079) | (0.108) | (0.148) | (0.080) | (2.025) | (0.144) | (0.078) |
| | {0.000} | {0.001} | {0.000} | {0.000} | {0.488} | {0.601} | {0.446} | {0.059} | {0.945} |
| Surveyor 6 | -0.566 | -0.215 | 0.665 | 0.381 | 0.243 | 0.405 | 1.376 | -0.079 | 0.022 |
| | (0.147) | (0.186) | (0.083) | (0.096) | (0.127) | (0.071) | (1.814) | (0.144) | (0.076) |
| | {0.000} | {0.249} | {0.000} | {0.000} | {0.056} | {0.000} | {0.449} | {0.586} | {0.772} |
| Surveyor 7 | -1.001 | 0.293 | 0.592 | 0.781 | 0.291 | 0.452 | 1.782 | -0.115 | 0.135 |
| | (0.125) | (0.159) | (0.081) | (0.104) | (0.113) | (0.065) | (1.696) | (0.155) | (0.070) |
| | {0.000} | {0.065} | {0.000} | {0.000} | {0.010} | {0.000} | {0.294} | {0.457} | {0.055} |
| Surveyor 8 | -0.271 | 1.078 | 0.947 | 0.510 | 0.701 | 0.024 | -0.449 | 0.128 | 0.047 |
| | (0.133) | (0.172) | (0.087) | (0.099) | (0.133) | (0.070) | (1.691) | (0.174) | (0.075) |
| | {0.041} | {0.000} | {0.000} | {0.000} | {0.000} | {0.734} | {0.791} | {0.461} | {0.532} |
| Surveyor 9 | -0.589 | 0.421 | 0.475 | 0.185 | 0.189 | 0.234 | 0.669 | -0.159 | 0.013 |
| | (0.139) | (0.171) | (0.084) | (0.096) | (0.133) | (0.068) | (1.655) | (0.158) | (0.072) |
| | {0.000} | {0.014} | {0.000} | {0.054} | {0.157} | {0.001} | {0.686} | {0.313} | {0.852} |
| Surveyor 10 | 0.790 | -0.819 | -0.108 | -0.096 | -0.287 | -0.151 | 0.510 | -0.037 | -0.018 |
| | (0.153) | (0.107) | (0.071) | (0.096) | (0.122) | (0.066) | (1.676) | (0.149) | (0.064) |
| | {0.000} | {0.000} | {0.127} | {0.317} | {0.019} | {0.023} | {0.761} | {0.803} | {0.777} |
| Surveyors 11-13 | -0.090 | 0.994 | 0.027 | 0.246 | -0.518 | -0.027 | -5.469 | -0.368 | -0.214 |
| | (0.199) | (0.176) | (0.139) | (0.114) | (0.249) | (0.084) | (2.470) | (0.224) | (0.112) |
| | {0.652} | {0.000} | {0.846} | {0.032} | {0.038} | {0.743} | {0.027} | {0.101} | {0.055} |
| Surveyor 14 | | | | | | | 2.310 | -0.065 | |
| | | | | | | | (1.891) | (0.160) | |
| | | | | | | | {0.222} | {0.687} | |
| F-stat surveyor FE | 35.65 | 29.73 | 39.31 | 14.90 | 6.49 | 14.02 | 1.40 | 1.35 | 1.48 |
| P-val F-stat surveyor FE | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.167 | 0.192 | 0.143 |
| Control mean | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.51 | 16.61 | 1.53 | 0.00 |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clinic FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,258 | 1,258 | 1,258 | 1,258 | 1,184 | 888 | 774 | 775 | 1,258 |

Note: Endline and LENA day 1 recording data. For columns 1-6 and 9, regressions include controls for the child's age in days, day of the survey, and surveyor gender (as in Table 4). In columns 7 and 8, regressions include controls for the child's age in days, the day of the week the audio was recorded (dummies), the total time (min) the shirt/LENA device was removed from the child, the total time (min) the child was held on someone's back while wearing the device, and the household size (as in Table 4). All regressions include baseline clinic fixed effects (not shown) and surveyor fixed effects (shown). The dummy for one surveyor ("Surveyor 1") is omitted. Since 3 surveyors completed 20 or fewer endline surveys each, we grouped them ("Surveyors 11-13"). Note that we still use endline surveyors for the LENA outcomes (as opposed to LENA surveyors) since, at the time the LENA measures were collected, the LENA surveyors had minimal contact with respondents (only dropping off and setting up the LENA devices). See main text Table 4 for further details on outcomes. Robust standard errors in parenthesis, p-values in curly brackets.

Table A.14: Treatment Effect on Perceived Barriers to Parent-Infant Conversations

| | Agreed with main barrier | | | Does not |
|---|---|---|---|---|
| | Beliefs | Hard to form habit | Risk of social scorn | agree there is any barrier |
| | (1) | (2) | (3) | (4) |
| Main intervention | -0.019 | 0.043 | 0.086 | -0.104 |
| | (0.028) | (0.045) | (0.041) | (0.058) |
| | {0.497} | {0.341} | {0.037} | {0.074} |
| Endline intervention | -0.050 | 0.041 | 0.197 | -0.191 |
| | (0.021) | (0.038) | (0.039) | (0.049) |
| | {0.018} | {0.275} | {<0.001} | {<0.001} |
| Pure control mean | 0.07 | 0.15 | 0.10 | 0.66 |
| P-value Endline=Main | 0.150 | 0.967 | 0.020 | 0.142 |
| Controls | Yes | Yes | Yes | Yes |
| Clinic FE | Yes | Yes | Yes | Yes |
| Observations | 780 | 780 | 780 | 780 |

Note: Endline data. Sample restricted to respondents who received a LENA device hence answered the IDS barrier module after 1 or 2 days of recording (see Figure A.3 for further details on the experimental design and timing of the "Module: IDS Barriers"). Respondents were asked about barriers that may prevent families from talking to their babies. Respondents were specifically asked about three barriers: "it's hard to remember to do it, it takes effort to make it a habit" (habit), "it's not clear that it makes any difference for the child" (belief), and "it's frowned upon /mocked in the community" (social sanctions/scorn). Respondents could also suggest other barriers. Regressions include baseline clinic fixed effects and controls for the child's age (in days), endline survey date, days elapsed between the endline survey and the LENA activity, and LENA randomization strata. Robust standard errors in parenthesis, p-values in curly brackets.

## Table A.15: Treatment Effect Split by Child Birth Order

| | Mother's Interview | | | | Observed | | LENA | |
|---|---|---|---|---|---|---|---|---|
| | Mother's belief index | Mother's behavior index | Child language score | Child gestural communication index | Child development index | Child babbles (= 1) | Female adult words per minute | Child vocalizations per minute |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Treatment | 0.143 | 0.090 | 0.066 | 0.047 | 0.040 | 0.030 | 0.272 | 0.050 |
| | (0.070) | (0.064) | (0.040) | (0.047) | (0.057) | (0.033) | (0.821) | (0.068) |
| | {0.041} | {0.156} | {0.099} | {0.315} | {0.482} | {0.362} | {0.741} | {0.459} |
| Treatment x 1st-time mother | -0.051 | 0.132 | 0.129 | 0.186 | -0.012 | 0.092 | -0.672 | -0.001 |
| | (0.128) | (0.127) | (0.077) | (0.094) | (0.110) | (0.063) | (1.521) | (0.131) |
| | {0.691} | {0.300} | {0.092} | {0.048} | {0.914} | {0.143} | {0.659} | {0.997} |
| 1st-time mother | 0.116 | -0.000 | -0.079 | -0.037 | -0.043 | -0.023 | -0.482 | -0.027 |
| | (0.084) | (0.090) | (0.054) | (0.064) | (0.075) | (0.044) | (1.052) | (0.096) |
| | {0.166} | {0.997} | {0.140} | {0.560} | {0.565} | {0.609} | {0.647} | {0.778} |
| P-val total effect for 1st-time mother | 0.385 | 0.045 | 0.003 | 0.004 | 0.765 | 0.021 | 0.757 | 0.656 |
| Control mean not 1st-time mother | -0.02 | -0.01 | -0.06 | -0.06 | -0.04 | 0.49 | 17.02 | 1.55 |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clinic FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,258 | 1,258 | 1,258 | 1,258 | 1,184 | 888 | 774 | 775 |
| Observations: 1st-time mother | 347 | 347 | 347 | 347 | 335 | 240 | 213 | 214 |

Note: Endline and LENA day 1 recording data. See main text Table 4 for details on specifications and outcomes. "1st-time mother" is a dummy equal to 1 if the focal child is the first born of the respondent. Robust standard errors in parenthesis, p-values in curly brackets.

Table A.16: Baseline Characteristics and Balance for Sample with First-Time Mothers

| | **Full Sample** | | | Treatment | | Control | | Treatment = Control |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | N | Mean | SD | Mean | SD | P-value |
| Age (years) | 23.25 | 3.37 | 392 | 23.25 | 3.34 | 23.25 | 3.39 | 1.000 |
| Dagomba ethnie | 0.81 | 0.40 | 393 | 0.81 | 0.39 | 0.80 | 0.40 | 0.882 |
| Main langage spoken: Dagbani | 0.86 | 0.35 | 393 | 0.88 | 0.33 | 0.85 | 0.36 | 0.422 |
| Highest level of education: | | | | | | | | |
| None | 0.21 | 0.41 | 393 | 0.21 | 0.41 | 0.22 | 0.41 | 0.832 |
| Primary school | 0.32 | 0.47 | 393 | 0.33 | 0.47 | 0.31 | 0.46 | 0.665 |
| Secondary school | 0.33 | 0.47 | 393 | 0.33 | 0.47 | 0.33 | 0.47 | 0.988 |
| Can read (English/Dagbani) | 0.75 | 0.43 | 393 | 0.76 | 0.43 | 0.74 | 0.44 | 0.660 |
| Housewive/no occupation | 0.35 | 0.48 | 393 | 0.30 | 0.46 | 0.39 | 0.49 | 0.067 |
| Married | 0.98 | 0.12 | 393 | 1.00 | 0.00 | 0.97 | 0.17 | 0.014 |
| Polygamous | 0.20 | 0.40 | 361 | 0.23 | 0.42 | 0.18 | 0.38 | 0.229 |
| Partner is home whole month | 0.76 | 0.43 | 388 | 0.75 | 0.43 | 0.77 | 0.42 | 0.573 |
| Partner passed primary school | 0.86 | 0.35 | 388 | 0.86 | 0.35 | 0.86 | 0.34 | 0.909 |
| Household size | 7.56 | 5.43 | 390 | 7.82 | 5.74 | 7.33 | 5.14 | 0.375 |
| # of household members: under-5 | 1.44 | 1.46 | 393 | 1.51 | 1.49 | 1.38 | 1.44 | 0.387 |
| # of household members: 5-15 y/o | 1.29 | 1.81 | 391 | 1.43 | 2.01 | 1.16 | 1.60 | 0.149 |
| # of household members: above-16 | 4.84 | 3.14 | 390 | 4.90 | 3.27 | 4.79 | 3.03 | 0.739 |
| Has children | 0.66 | 0.47 | 393 | 0.69 | 0.46 | 0.63 | 0.48 | 0.221 |
| Has child 6 years or younger | 0.51 | 0.50 | 393 | 0.53 | 0.50 | 0.49 | 0.50 | 0.432 |
| Has child older than 1 month | 0.50 | 0.50 | 393 | 0.51 | 0.50 | 0.48 | 0.50 | 0.586 |
| Has child older than 3 months | 0.40 | 0.49 | 393 | 0.40 | 0.49 | 0.41 | 0.49 | 0.841 |
| Age at first child (years) | 23.03 | 3.39 | 259 | 22.80 | 3.43 | 23.27 | 3.35 | 0.266 |
| # of children | 0.74 | 0.71 | 393 | 0.81 | 0.82 | 0.67 | 0.60 | 0.056 |
| Age youngest child (months) | 5.47 | 3.43 | 259 | 5.25 | 3.34 | 5.69 | 3.51 | 0.305 |
| Youngest child eligible | 0.66 | 0.47 | 393 | 0.69 | 0.46 | 0.63 | 0.48 | 0.221 |
| Pregnant with an eligible child | 0.34 | 0.47 | 393 | 0.31 | 0.46 | 0.37 | 0.48 | 0.221 |
| Focal child is first born | 1.00 | 0.00 | 393 | 1.00 | 0.00 | 1.00 | 0.00 | . |
| F-test p-value | | | | | | | | 0.431 |
| Observations | 393 | | | 184 | | 209 | | |

Note: Baseline data. The sample is restricted to mothers whose child enrolled in the study was their first (alive) child. Treatment is a dummy equal to 1 if the respondent received the intervention at baseline. The question on polygamy was added after the start of the data collection, hence is missing for some observations. Note that some mothers had twins or adopted children from their relatives, hence # of children is on average slightly higher than the % of women who have children. The F-test p-value reported at the bottom of the table is for the joint significance of the differences between the treatment and control groups for all of the variables reported in the table. For the F-test, missing values (due to refusal/don't know or a logic skip (e.g., age of youngest child when no children)) are replaced by the variable average value and flagged by a dummy.

Table A.17: Baseline IDS beliefs and Behavior for First-Time Mothers Sample

| | **Full Sample** | | | Treatment | | Control | | Treatment = Control |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | N | Mean | SD | Mean | SD | P-value |
| **Beliefs on IDS and Child Development:** | | | | | | | | |
| Time/attention is more important than money to a child's success | 0.35 | 0.48 | 393 | 0.36 | 0.48 | 0.34 | 0.48 | 0.686 |
| *Child's age (in mo) when:* | | | | | | | | |
| a child starts responding with noise/babbles | 7.83 | 10.40 | 366 | 8.01 | 9.67 | 7.68 | 11.02 | 0.762 |
| a child starts saying meaningful words | 20.97 | 14.54 | 351 | 21.12 | 14.40 | 20.84 | 14.71 | 0.858 |
| it becomes clear a child is smart | 32.67 | 21.34 | 369 | 32.42 | 22.54 | 32.90 | 20.24 | 0.830 |
| *Child's age (in mo) when parents should start:* | | | | | | | | |
| talking to their child | 10.83 | 11.37 | 375 | 12.08 | 12.22 | 9.74 | 10.48 | 0.049 |
| talking in full sentences to their child | 25.66 | 19.11 | 345 | 27.01 | 20.86 | 24.53 | 17.50 | 0.237 |
| telling stories to their child | 23.32 | 18.63 | 356 | 23.74 | 17.72 | 22.94 | 19.44 | 0.684 |
| **Self-Reported IDS Behavior:** | | | | | | | | |
| Tells stories to youngest child | 0.43 | 0.50 | 201 | 0.39 | 0.49 | 0.47 | 0.50 | 0.264 |
| Ask youngest child to repeat words | 0.45 | 0.50 | 201 | 0.45 | 0.50 | 0.46 | 0.50 | 0.917 |
| When child was 1m/o: Described objects when cleaning/organizing | 0.36 | 0.48 | 195 | 0.34 | 0.48 | 0.39 | 0.49 | 0.509 |
| When child was 3m/o: Described things to child when walking | 0.58 | 0.49 | 158 | 0.64 | 0.48 | 0.53 | 0.50 | 0.147 |
| **Inequality Aversion:** | | | | | | | | |
| It is best to treat/invest in children equally | 0.46 | 0.50 | 393 | 0.45 | 0.50 | 0.48 | 0.50 | 0.516 |
| A mother should feel bad for 1st child if she provides better care to 2nd child | 0.72 | 0.45 | 393 | 0.77 | 0.42 | 0.68 | 0.47 | 0.054 |
| F-test p-value | | | | | | | | 0.361 |
| Observations | 393 | | | 184 | | 209 | | |

Note: Baseline data. Sample restricted to mothers whose child enrolled in the study was a first (alive) child. In the panel "Beliefs on IDS and Child Development", questions "child's age (in months) when parents should start..." were only asked to respondents who reported that the respective activities were important to children's brain development. Child's age outcomes are in months. In the panel "Self-Reported IDS Behavior", questions were only asked to a subset of respondents based on their youngest child's age. "Tell stories to youngest child" and "Asks youngest child to repeat words" were only asked to respondents with a child aged 6 years or less, and the two subsequent questions to parents with a child aged between 1 month and 6 years, and between 3 months and 6 years. The F-test p-value reported at the bottom of the table is for the joint significance of the differences between the treatment and control groups for all of the variables reported in the table. For the F-test, missing values (due to refusal/don't know or a skip pattern such as age of youngest child when there are no children) are replaced by the variable average value and a missing flag is included.

## Table A.18: Treatment Effect Split by Use of Calendar to Track Habit

| | Mother's Interview | | | | Observed | | LENA | |
|---|---|---|---|---|---|---|---|---|
| | Mother's belief index | Mother's behavior index | Child language score | Child gestural communication index | Child development index | Child babbles (= 1) | Female adult words per minute | Child vocalizations per minute |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Treatment | 0.102 | 0.075 | 0.088 | 0.053 | 0.061 | 0.033 | -0.055 | 0.094 |
| | (0.066) | (0.068) | (0.041) | (0.048) | (0.056) | (0.032) | (0.835) | (0.070) |
| | {0.123} | {0.267} | {0.032} | {0.270} | {0.275} | {0.314} | {0.947} | {0.180} |
| Treatment x Colored stars | 0.135 | 0.170 | 0.044 | 0.127 | 0.011 | 0.056 | 0.851 | -0.097 |
| | (0.100) | (0.086) | (0.057) | (0.072) | (0.075) | (0.042) | (1.066) | (0.095) |
| | {0.178} | {0.049} | {0.443} | {0.079} | {0.878} | {0.176} | {0.425} | {0.309} |
| P-val total effect for colored stars | 0.012 | 0.002 | 0.012 | 0.007 | 0.290 | 0.020 | 0.418 | 0.970 |
| Control mean | -0.01 | -0.00 | 0.02 | 0.02 | 0.02 | 0.51 | 16.62 | 1.53 |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clinic FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,136 | 1,136 | 1,136 | 1,136 | 1,076 | 881 | 723 | 724 |
| Observations: colored stars | 188 | 188 | 188 | 188 | 174 | 172 | 117 | 117 |

Note: Endline and LENA day 1 recording data. See main text Table 4 for details on specifications and outcomes. "Colored stars" is a dummy equal to 1 if the respondent reported keeping track of her IDS-practice by coloring stars on the calendar given at baseline to treated respondents. The dummy is always equal to 0 for control respondents as they did not receive the main intervention and were not given the IDS-themed calendar nor asked those questions. The sample size is smaller than in other tables as we added question on calendar use mid-survey so do not have this data for 90 treatment respondents (see Table 3). See Figure A.1 for further details on the calendar. Robust standard errors in parenthesis, p-values in curly brackets.

## References

**Ahrens, Achim, Christian B. Hansen, and Mark E Schaffer**, "PDSLASSO: Stata module for post-selection and post- regularization OLS or IV estimation and inference," *Statistical Software Components*, 1 2019.

**Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen**, "Inference on treatment effects after selection among high-dimensional controls," *Review of Economic Studies*, 2013, 81 (2), 608–650.

# Appendix B: LENA Technology Description

## What is a LENA device?

A LENA device is a small recorder children wear for a day in the front pocket of a "LENA shirt" (see Figure B.1). It functions as a sort of "talk pedometer". The audio is processed by a cloud-based LENA software which provides detailed information on the child's audio environment. Available outcomes include lengths of the audio capturing sounds coming from 8 different sources: focal child, other children, male and female adults, overlapping speech, electronic media,non-speech noise, and silence. The LENA software further categorizes those segments as near or far/faint (those with a lower probability of being attributed to the right source of origin) and extracts additional outcomes such as lengths of meaningful speech (including near sounds emitted by the focal child, adults, or other children), faint or overlapping speech (including overlapping speech or far sounds emitted by the focal child, adults, or other children), lengths of focal child vocalizations and non-vocalizations segments, number of focal child vocalizations, adult words, and conversational turn counts. Those are further described in the next section.

Figure B.1: LENA device and shirt

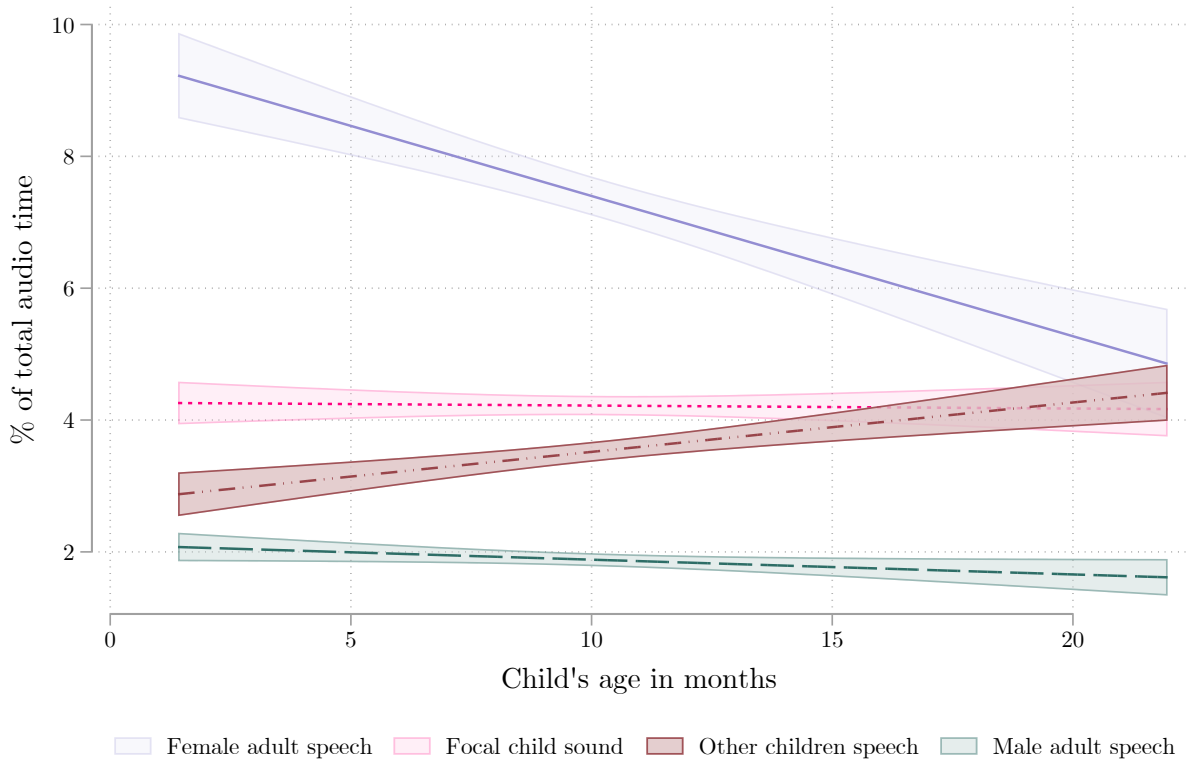
Note: A child participating in the study and wearing the LENA shirt with the device inserted in the front pocket. The LENA device was purchased from the LENA Foundation and the shirt was designed by the research team.

## Description of LENA outcomes

In the paper, we focus on the two LENA outcomes summarized below. For further details and more complete descriptions of all LENA outcomes, please refer to the LENA technical reports LTR 12 (available at https://www.lena.org/wp-content/uploads/2020/07/LTR-12_How_LENA_Works.pdf) and LTR-05-2 (available at https://www.lena.org/wp-content/uploads/2016/07/LTR-05-2_Reliability.pdf) (Gilkerson and Richards, 2020; Xu, Yapanel and Gray, 2009)

- **Adult words count:** estimated number of words spoken by post-pubescent males and females in the child's vicinity.
- **Focal child sounds:** any sound from the child wearing the device. The LENA software further categorizes the focal child sound segments into 1) **vocalizations** (including words, babbles, and pre-speech communicative sounds or "protophones" such as squeals, growls, or raspberries) and 2) **non-speech** sounds like fixed signals and vegetative sounds (such as breathing, burping, crying, etc.).
- **Vocalization count:** estimated number of speech-like utterances the focal child wearing the device emits (including words, babbles, and pre-speech communicative sounds or "protophones" such as squeals, growls, or raspberries). It excludes non-speech sounds like fixed signals or vegetative sounds (such as breathing, burping, crying, etc.).
- **Conversational turn count:** estimated number of back-and-forth alternations between the focal child wearing the device (any alternation including a vocalization) and an adult.
- **% audio of meaningful speech:** includes all segments of the audio labeled as near sounds by adults or focal children as well as near vocalizations from other children.

Figure B.2: Proportion of Speech by Speaker Over Time




Note: LENA data. N=775 recordings. Given financial constraints, only a random subset of households could be included in the LENA measurement (N=900 households sampled). For households sampled to keep the LENA device for two days, only the first day recording is kept in the analysis. The analysis is further restricted to recording times between 10 a.m. to 7 p.m. (this excludes 10/785 LENA day 1 recordings which have less than 9 hours (rounded up) of recording). Outcomes are % of total audio time. Focal child sound (as % of total audio) is the share of the recording tagged by the LENA software as emitted by the child, including both vocalizations and non-vocalizations (cry, fixed signals, vegetative sounds). Lines indicate linear best fit, and shaded areas indicate 95% confidence intervals.

Table B.1: LENA Debrief Survey by Treatment Status (Only 1st Day of LENA Recording)

| | Treatment | | | Control | | | Treatment = Control |
|---|---|---|---|---|---|---|---|
| | Mean | SD | N | Mean | SD | N | P-value |
| Shirt/device removed during recording | 0.93 | 0.25 | 385 | 0.94 | 0.24 | 390 | 0.625 |
| # times shirt/device removed | 1.42 | 0.76 | 385 | 1.44 | 0.74 | 390 | 0.707 |
| Device removed [10h,18h] | 0.52 | 0.50 | 385 | 0.50 | 0.50 | 390 | 0.540 |
| # times deviced removed [10h,18h] | 0.66 | 0.72 | 385 | 0.64 | 0.74 | 390 | 0.759 |
| Total min device removed [10h,18h] | 42.79 | 78.89 | 346 | 39.85 | 70.85 | 347 | 0.606 |
| Device removed during [10h,18h] but invalid duration | 0.10 | 0.30 | 385 | 0.11 | 0.31 | 390 | 0.686 |
| Child carried on someone's back with device | 0.52 | 0.50 | 385 | 0.51 | 0.50 | 390 | 0.853 |
| # times child carried on back with device | 0.80 | 0.93 | 385 | 0.82 | 0.98 | 390 | 0.765 |
| Held on back [10h,18h] | 0.47 | 0.50 | 385 | 0.47 | 0.50 | 390 | 0.922 |
| # times held on back [10h,18h] | 0.66 | 0.83 | 385 | 0.67 | 0.83 | 390 | 0.875 |
| Total min held on back [10h,18h] | 29.18 | 53.23 | 361 | 35.10 | 69.68 | 351 | 0.204 |
| Child held on back during [10h,18h] but invalid duration | 0.06 | 0.24 | 385 | 0.10 | 0.30 | 390 | 0.055 |
| Day was unusual for child | 0.09 | 0.29 | 385 | 0.09 | 0.28 | 390 | 0.856 |
| *Reason why day was unusual for child:* | | | | | | | |
| Child was sick | 0.26 | 0.44 | 35 | 0.29 | 0.46 | 34 | 0.736 |
| Child cried throughout day for no reason | 0.29 | 0.46 | 35 | 0.24 | 0.43 | 34 | 0.639 |
| Child was uncomfortable with LENA | 0.37 | 0.49 | 35 | 0.41 | 0.50 | 34 | 0.736 |
| Child took immunizations | 0.06 | 0.24 | 35 | 0.06 | 0.24 | 34 | 0.977 |
| Number of LENA 1 recordings | 385 | | | 390 | | | 775 |

Note: LENA debrief survey (collected the day after the LENA recording). The sample is restricted to recordings with data from 10 a.m. to 7 p.m. (this excludes 12/785 LENA day 1 audio which have less than 9 hours (rounded up) of recording). Variables in the "LENA Debrief Survey" come from questions asked to primary caregiver the day following the recording.

## References

**Gilkerson, Jill and Jeffrey A Richards**, "A Guide to Understanding the Design and Purpose of the LENA System," *LENA Foundation Technical Report*, 2020

**Xu, Dongxin, Umit Yapanel, and Sharmi Gray**, "Reliability of the LENA Language Environment Analysis System in Young Children's Natural Home Environment" *LENA Foundation Technical Report*, 2009

# Appendix C: Descriptive analyses of outcome variables

This appendix highlights cross-sectional patterns in our data to offer insight into our outcome measures. We focus on data from the control group, i.e., a group not informed about the benefits of IDS. In Table C.1, we estimate multivariate regressions to understand whether socioeconomic status predicts our outcome variables. In Figures C.1 to C.8, we present scatter plots and bivariate correlations between the LENA-recorded measures, the mother-reported measures, the surveyor-observed measures, and child age. We discuss cross-sectional patterns in the data revealed by these analyses.

## C.1 Associations with socioeconomic status

Table C.1 shows that maternal education and indicators of wealth (having a home with concrete walls rather than mud walls) are *negatively* associated with both female adult words per minute and child vocalizations per minute as recorded by the LENA. This is the exact opposite of what has been observed in higher-income countries. A key question is whether this is indicative that the LENA provides poor measures of the true outcomes in our context, or whether the gradient is truly different across higher-income and lower-income contexts.

The lack of a positive SES gradient appears common to all measures of child development (whether LENA, reported by the mother or observed by the surveyor), which strongly suggests that the underlying relationship is truly different in lower-income contexts. (It is only for mother-reported measures of parental behavior that we do observe a positive SES gradient, but this may be due to social desirability bias.) Overall, this evidence suggests that even the more educated and/or wealthier parents may under-invest in the early cognitive development of their children.

Table C.1: Correlation Between Outcome Variables and Socioeconomic Status

| | LENA | | Observed | | Mother's Interview | |
|---|---|---|---|---|---|---|
| | Female adult words per min | Child vocalizations per minute | Child development index | Child babbles (= 1) | Child language score | Mother's behavior index |
| Panel A: Maternal education | | | | | | |
| Mother's years of education | -0.220 | -0.015 | 0.007 | 0.007 | 0.000 | 0.022 |
| | (0.089) | (0.007) | (0.006) | (0.003) | (0.004) | (0.007) |
| | {0.014} | {0.038} | {0.248} | {0.042} | {0.909} | {0.002} |
| Control mean | 16.61 | 1.53 | 0.00 | 0.51 | -0.06 | 0.00 |
| Observations | 388 | 389 | 598 | 435 | 628 | 628 |
| Panel B: Indicator of wealth | | | | | | |
| Concrete walls | -1.319 | -0.095 | 0.057 | 0.057 | -0.014 | 0.068 |
| | (1.030) | (0.087) | (0.069) | (0.041) | (0.048) | (0.078) |
| | {0.201} | {0.277} | {0.403} | {0.166} | {0.766} | {0.384} |
| Control mean | 16.61 | 1.53 | 0.00 | 0.51 | -0.06 | 0.00 |
| Observations | 389 | 390 | 603 | 439 | 633 | 633 |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Clinic FE | No | No | No | No | No | No |

Note: Endline and LENA day 1 recording data. Sample is restricted to control group respondents. Panel A and Panel B report results from separate regressions. Regressions include controls for the child age group (under 5 months old, 6-9 months old, 10-14 months old, over 15 months old dummies). See Table 4 notes for further details about the sample and Online Appendix A.2 for details on the LENA device and outcomes. "Mother's years of education" is the number of years of education completed by the child's mother. "Concrete walls" is equal to 1 if the child's home has concrete walls and 0 otherwise. Robust standard errors in parenthesis, p-values in curly brackets.

## C.2 Reliability of measures

Figure C.1 shows a weak relationship between child verbal output and child age. This also contrasts with findings from higher-income countries (Gilkerson and Richards, 2008). We also observe falling mother verbal input with child age (Figure C.1). This latter result may explain the former: lower mother verbal input means the child has less reason to vocalize even if their language skills have improved. As such, this would indicate that the LENA measure of child vocalizations is not a good proxy for child language development. Consistent with this, we find that child vocalizations per minute has weak correlations with the surveyor-observed child development index or child babbling (Figure C.3) as well as the mother-reported child language score or child gestural communication index (Figure C.4). Figure C.2 also shows a

weak correlation between female adult word count and mother-reported behavior.

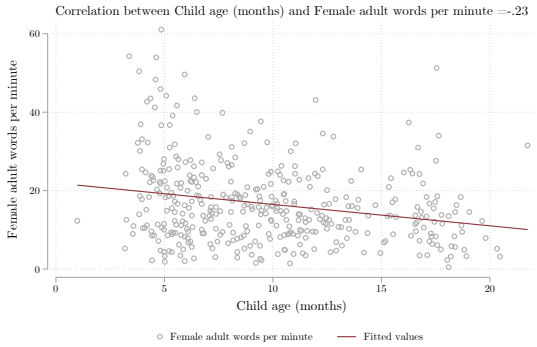
The correlation between LENA-recorded child verbal outputs and mother-reported/surveyor-observed measures is stronger for children 15 months or older. In Figure C.5, we estimate bivariate correlations between our main outcomes for four age groups: 5 months or younger, 6-9 months old, 10-14 months old, and 15 months or older. Children 15 months or older exhibit the strongest relationships between child vocalizations per minute and mother-reported measures (e.g., the child language score) as well as surveyor-observed measures (e.g., the child development index).

As expected, there are strong relationships between surveyor-observed/mother-reported measures of child development and child age. The bivariate correlations for child age and the child development index, child babbling, the child language score, or the child gestural communication index range from 0.54 to 0.77 (Figure C.6 and Figure C.7).

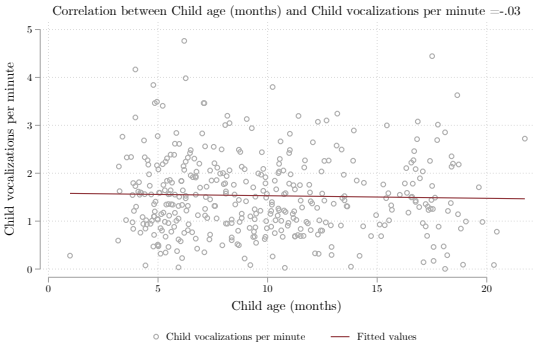
The surveyor-observed and mother-reported child development measures are also predictive of each other. Figure C.8 estimates correlations of 0.44 to 0.52 between the surveyor-observed measures (child development index/child babbling) and the mother-reported measures (child language score/the child gestural communication index). These correlations are present across age groups, though they are weaker for children under 5 months old.

Figure C.1: Correlation Between LENA-recorded Outcomes and Child Age

(a) LENA-recorded Parental Verbal Inputs



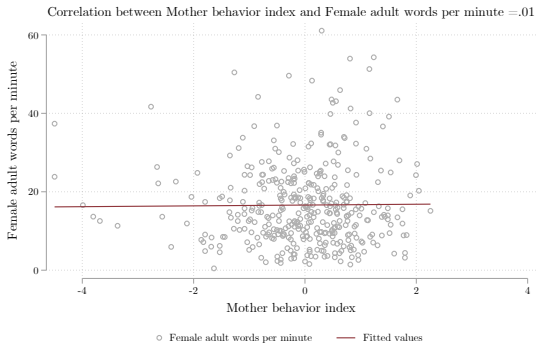

(b) LENA-recorded Child Verbal Outputs



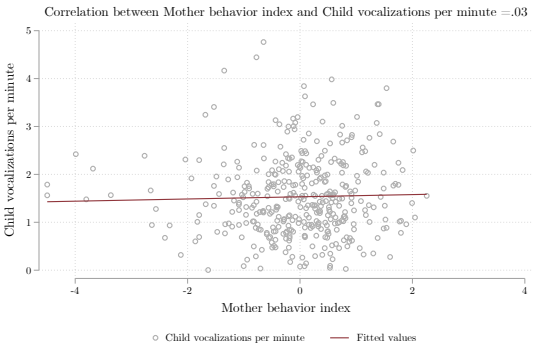

Notes: Endline and LENA data. Using data from the control group only. For details on outcomes and samples, please refer to Table 4. All correlations presented are bivariate correlations.

Figure C.2: Correlation Between LENA-recorded Outcomes and Mother-reported Behavior Index

(a) LENA-recorded Parental Verbal Inputs



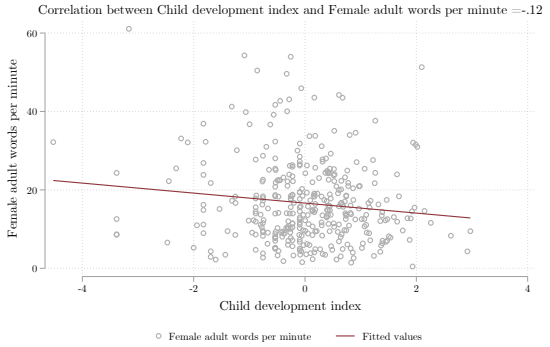

(b) LENA-recorded Child Verbal Outputs



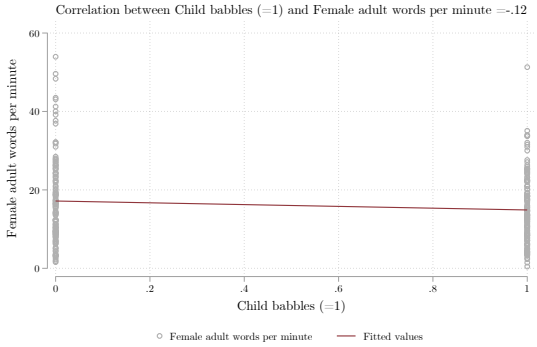

Notes: Endline and LENA data. Using data from the control group only. For details on outcomes and samples, please refer to Table 4. All correlations presented are bivariate correlations.

Figure C.3: Correlation between LENA-recorded Outcomes and Surveyor-observed Child Development

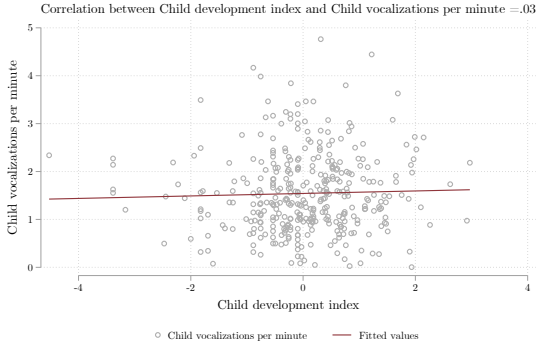
(a) LENA-recorded Parental Verbal Inputs and Surveyor Observed Child Development



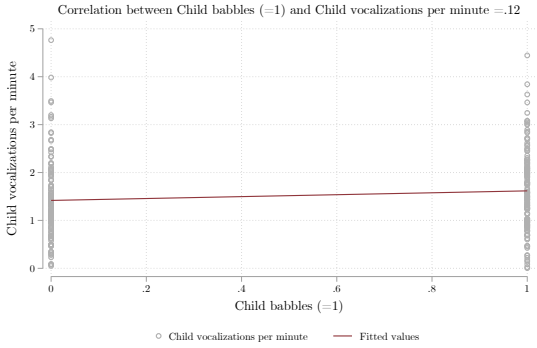

(b) LENA-recorded Parental Verbal Inputs and Surveyor-observed Child Babbling




(c) LENA-recorded Child Verbal Outputs and Surveyor-observed Child Development



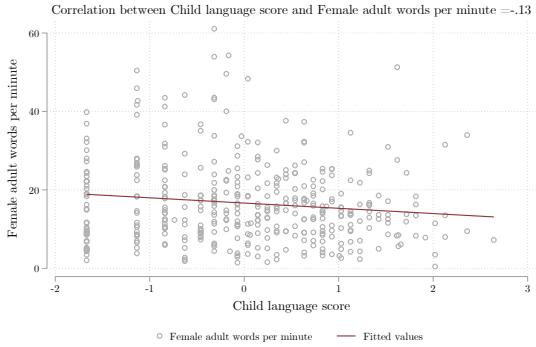

(d) LENA-recorded Child Verbal Outputs and Surveyor-observed Child Babbling



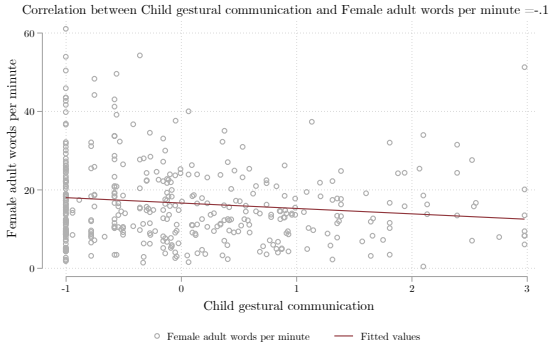

Notes: Endline and LENA data. Using data from the control group only. For details on outcomes and samples, please refer to Table 4. All correlations presented are bivariate correlations.

Figure C.4: Correlation Between LENA-recorded Outcomes and Mother-reported Child Communication

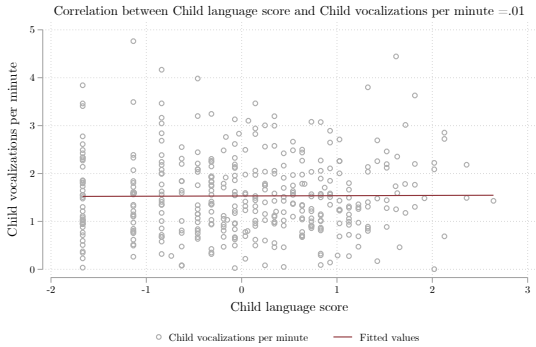
(a) LENA-recorded Parental Verbal Inputs and Mother-reported Child Language



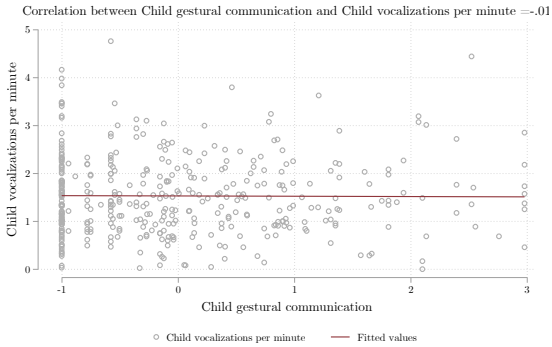

(b) LENA-recorded Parental Verbal Inputs and Mother-reported Child Gestural Communication




(c) LENA-recorded Child Verbal Outputs and Mother-reported Child Language



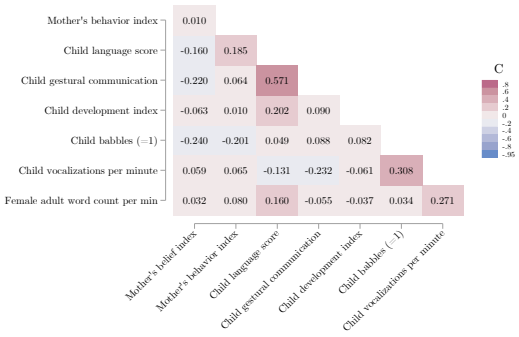

(d) LENA-recorded Child Verbal Outputs and Mother-reported Child Gestural Communication



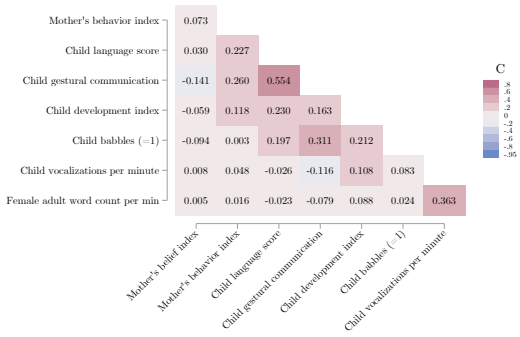

Notes: Endline and LENA data. Using data from the control group only. For details on outcomes and samples, please refer to Table 4. All correlations presented are bivariate correlations.

Figure C.5: Correlations of Outcome Variables by Age Group

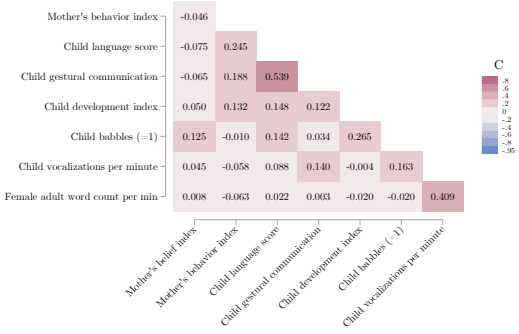
(a) 5 months or younger



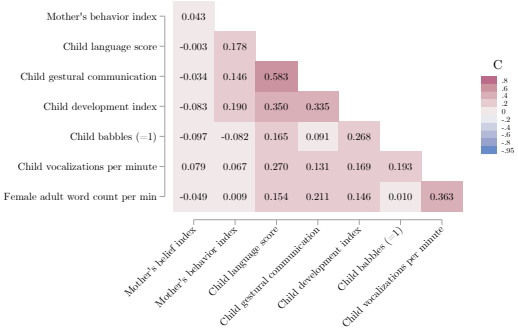

(b) 6-9 months old




(c) 10-14 months old



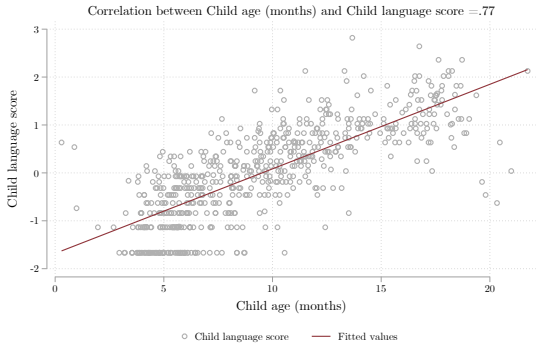

(d) 15 months or older



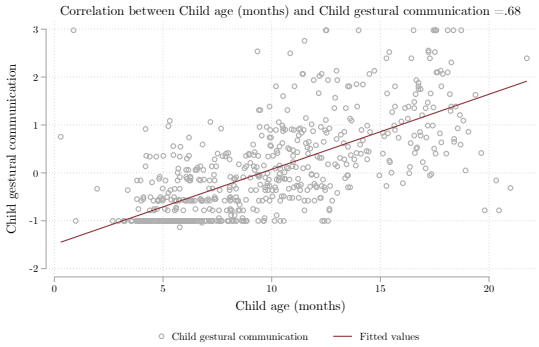

Notes: Endline and LENA data. Using data from the control group only. For details on outcomes and samples, please refer to Table 4. All correlations presented are bivariate correlations.

Figure C.6: Correlation Between Mother-reported Child Communication and Child Age

(a) Mother-reported Child Language



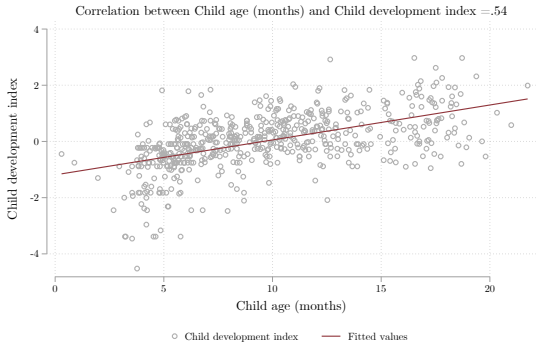

(b) Mother-reported Gestural Communication



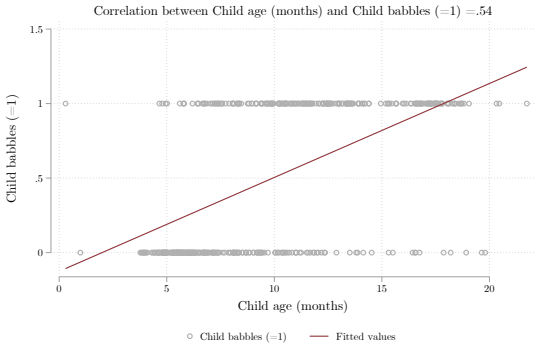

Notes: Endline. Using data from the control group only. For details on outcomes and samples, please refer to Table 4. All correlations presented are bivariate correlations.

Figure C.7: Correlation Between Surveyor-Observed Child Development and Child Age

(a) Surveyor-observed Child Development Index



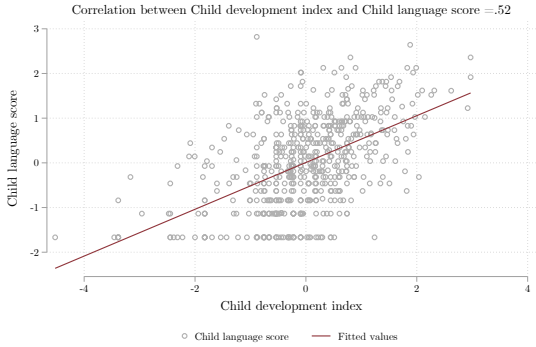

(b) Surveyor-observed Child Babbling



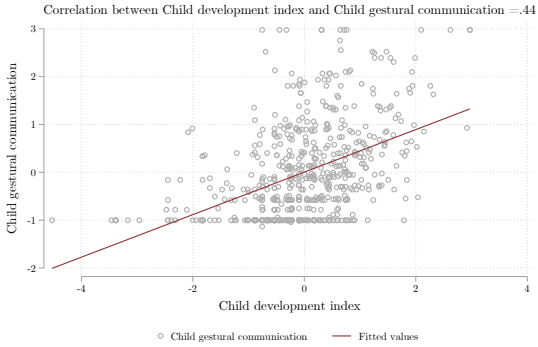

Notes: Endline. Using data from the control group only. For details on outcomes and samples, please refer to Table 4. All correlations presented are bivariate correlations.

Figure C.8: Correlation Between Surveyor-observed and Mother-reported Child Development Outcomes

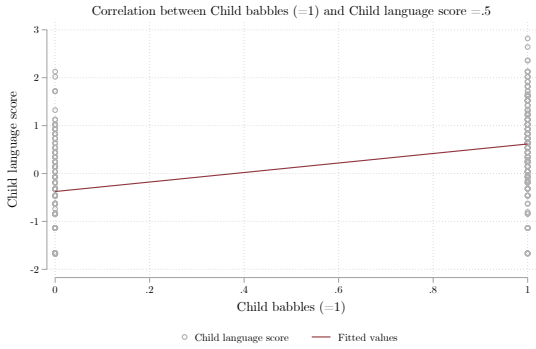
(a) Surveyor-observed Child Development Index and Mother-reported Child Language



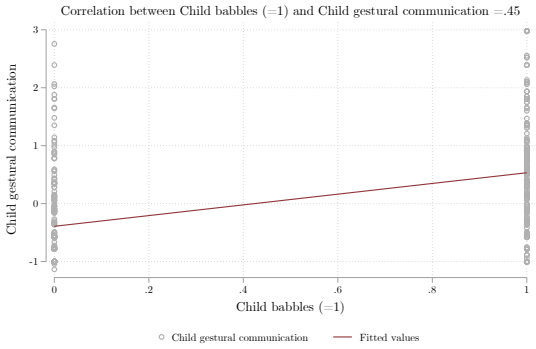

(b) Surveyor-observed Child Development and Mother-reported Child Gestural Communication




(c) Surveyor-observed Child Babbling and Mother-reported Child Language




(d) Surveyor-observed Child Babbling and Mother-reported Child Gestural Communication




Notes: Endline and LENA data. Using data from the control group only. For details on outcomes and samples, please refer to Table 4. All correlations presented are bivariate correlations.

## References

**Gilkerson, Jill and Jeffrey A Richards**, "The LENA natural language study" *LENA Foundation Technical Report*, 2008