# Alternative Dataset Supplemental Material

The typical development and autism recording data from the main text and above were the same as in Oller et al. (2010). A somewhat different but overlapping dataset was used by Warren et al. (2010). Both datasets were obtained by the LENA Foundation, and some of the recordings are the same in the two sets. Our choice to focus on the Oller et al. data was motivated primarily by the fact that all the recordings in that set were made using the same version of the LENA DLP recorder. Different versions of the recorder had somewhat different features, for example different microphones, which could in turn have systematic effects on segmenting and labeling of the recordings. The Oller et al. TD and ASD dataset also had the benefit of consisting of more recordings (1153 compared to 438). The Warren et al. data was not matched for recorder type but had the advantage that TD participants and participants with ASD were matched with respect to age, gender, and maternal education. Below we report the results obtained when the Warren et al. dataset was used instead of Oller et al. data. The following results show that although there are some differences in the findings obtained with the two datasets, the primary findings are the same.

## Demographics and sound types in the alternative dataset

A mixed effects regression was run for each sound type, with participant ID as a random effect and ASD, age, maternal education, gender, recorder version, interaction between ASD and age, and interaction between maternal education and age as fixed effects. The results of this analysis are given in the table below.

*Effects of ASD, age, maternal education, gender (male), and recorder version on prevalence of each sound type, measured in terms of total time per 12 hour recording, for the multiple recorder, age- and education-matched dataset used in Warren et al. (2010).*

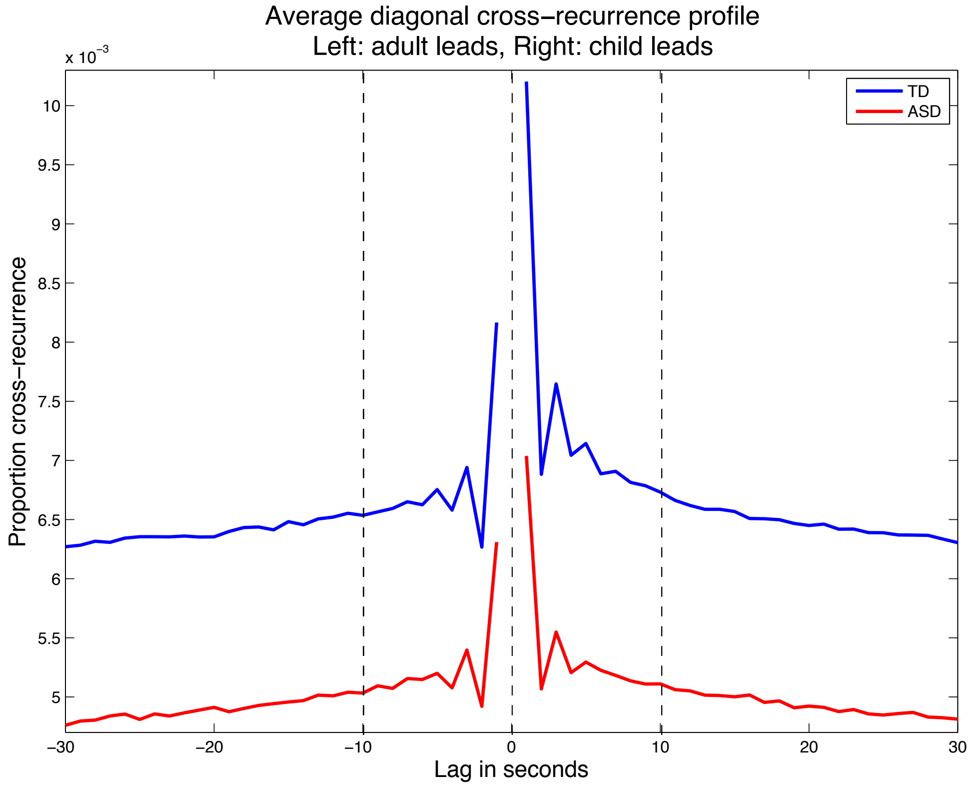
| Sound type | ASD | Age | MomEd | Gender | Recorder | Age*ASD | Age*MomEd |
|---|---|---|---|---|---|---|---|
| Child total | $\beta = -0.381$<br>$p < .001$ | $\beta = 0.237$<br>$p < .001$ | $\beta = 0.159$<br>$p = .005$ | $\beta = -0.124$<br>$p = .030$ | $\beta = 0.082$<br>$p = .164$ | $\beta = -0.093$<br>$p = .109$ | $\beta = 0.023$<br>$p = .683$ |
| Child speech-related | $\beta = -0.455$<br>$p < .001$ | $\beta = 0.301$<br>$p < .001$ | $\beta = 0.174$<br>$p = .001$ | $\beta = -0.133$<br>$p = .014$ | $\beta = 0.188$<br>$p = .001$ | $\beta = -0.124$<br>$p = .026$ | $\beta = 0.030$<br>$p = .559$ |
| Child cry and vegetative | $\beta = 0.070$<br>$p = .357$ | $\beta = -0.025$<br>$p = .687$ | $\beta = 0.010$<br>$p = .879$ | $\beta = -0.037$<br>$p = .549$ | $\beta = -0.090$<br>$p = .171$ | $\beta = 0.021$<br>$p = .736$ | $\beta = -0.012$<br>$p = .850$ |
| Adult | $\beta = -0.111$<br>$p = .130$ | $\beta = 0.090$<br>$p = .139$ | $\beta = 0.282$<br>$p < .001$ | $\beta = 0.047$<br>$p = .440$ | $\beta = 0.038$<br>$p = .545$ | $\beta = 0.053$<br>$p = .391$ | $\beta = -0.081$<br>$p = .165$ |
| Other child | $\beta = -0.294$<br>$p < .001$ | $\beta = 0.312$<br>$p < .001$ | $\beta = 0.191$<br>$p = .002$ | $\beta = 0.083$<br>$p = .108$ | $\beta = -0.271$<br>$p < .001$ | $\beta = -0.025$<br>$p = .633$ | $\beta = 0.079$<br>$p = .115$ |
| Overlap | $\beta = 0.146$<br>$p = .021$ | $\beta = 0.140$<br>$p = .001$ | $\beta = 0.083$<br>$p = .110$ | $\beta = 0.067$<br>$p = .205$ | $\beta = 0.350$<br>$p < .001$ | $\beta = 0.038$<br>$p = .463$ | $\beta = -0.019$<br>$p = .700$ |
| Electronic | $\beta = -0.134$<br>$p = .063$ | $\beta = 0.054$<br>$p = .358$ | $\beta = -0.131$<br>$p = .027$ | $\beta = 0.061$<br>$p = .310$ | $\beta = 0.241$<br>$p < .001$ | $\beta = -0.040$<br>$p = .502$ | $\beta = 0.055$<br>$p = .340$ |
| Noise | $\beta = 0.099$<br>$p = .170$ | $\beta = -0.176$<br>$p = .004$ | $\beta = -0.103$<br>$p = .084$ | $\beta = -0.186$<br>$p = .002$ | $\beta = -0.049$<br>$p = .427$ | $\beta = -0.046$<br>$p = .454$ | $\beta = 0.034$<br>$p = .553$ |
| Silence | $\beta = -0.040$<br>$p = .525$ | $\beta = -0.242$<br>$p < .001$ | $\beta = 0.018$<br>$p = .724$ | $\beta = 0.061$<br>$p = .244$ | $\beta = -0.435$<br>$p < .001$ | $\beta = 0.013$<br>$p = .794$ | $\beta = -0.024$<br>$p = .639$ |

Although Other Child vocalization segments were less frequent in the ASD group's recordings, the number of siblings living in the home was not very different across groups, with the mean number being .95 ($SD$ = 0.86) for the TD group's recordings and .93 ($SD$ = 1.04) for the ASD group's recordings.

Comparing the table to the results reported in the main text, we see that in both cases, quantity of speech-related vocalization was lower for children with ASD compared to TD and increased with age and maternal education. In the main dataset, we found an interaction between diagnosis and age, such that growth in speech-related vocalization was slower for children with ASD, but we only found this effect for proportion of vocalizations that were speech-related. The test when run on raw amount of speech-related vocalization was not statistically significant. However, in the alternate dataset, the interaction was significant for raw values but not for proportions. In both cases, the non-significant interactions were in the expected direction.

**Cross recurrence in the alternative dataset**

Recall that the overall height of the diagonal cross recurrence profile (DCRP) gives a measure of the amount of child-adult interactivity. The ratio of the height of the right side of the DCRP to the height of the left side gives a measure of the extent to which the child tended to initiate as opposed to follow. The figure below shows the average DCRP for the alternative dataset TD recordings and for the alternative dataset ASD recordings. The overall height of the DCRP was lower for the ASD group than the TD group, $\beta = -0.255$, $p < .001$, indicating reduced levels of child-caregiver interaction in ASD. The ratio of the child leading side of the DCRP to the caregiver leading side was lower for the ASD group than the TD group, $\beta = -0.383$, $p < .001$. The latter difference was also found for the primary dataset.




*Diagonal cross recurrence profiles, averaged across TD recordings (blue) and ASD recordings (red). Overall height of the plot indicates the overall level of child-adult interactivity. Displacement along the diagonal refers to the diagonal of the recurrence plots, not shown here, on which this profile is based (see Method Supplementary Material for further details). Displacement along the diagonal measures the difference between pairings of child and adult vocalizations in seconds.*

## Contingency of adult responses on content of child vocalizations in the alternative dataset

As was the case with the main dataset recordings, in the alternative dataset the presence of speech-related vocalization within a child segment did indeed correspond to increased likelihood of receiving an immediate adult response. For the typically developing children, the difference between the proportion of speech-related vocalizations receiving an adult response and the proportion of non-speech-related child vocalizations receiving an adult response was 0.054, $p < .001$. The same pattern held for the recordings of children with autism, with the difference between the two proportions being 0.040, $p < .001$. The contingency of adult response on child vocalization speech-likeness was higher for the TD group than the ASD group, $\beta = -0.151$, $p = .014$. It also increased as maternal education level increased, $\beta = 0.260$, $p < .001$ and was higher for males than females, $\beta = 0.148$, $p = .013$.

**The effect of contingent adult responses on subsequent child vocalizations in the alternative dataset**

As was the case again with the main dataset recordings, in the alternative dataset a child vocalization was more likely to be speech-related if the child's previous speech-related vocalization had received an immediate adult response. For TD children, the average difference between the proportion of child vocalizations that was speech-like when the preceding speech-related vocalization was responded to and the proportion when the preceding speech-related vocalization was not responded to was 0.031, $p < .001$. For the ASD children, it was 0.041, $p < .001$. This contingency of child vocalization type on previous adult response was significantly greater for male children, $\beta = 0.153$, $p = .005$, but did not differ significantly across diagnostic groups, age, or maternal education.

**References**

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

Warren, S., Gilkerson, J., Richards, J., Oller, D. K., Xu, D., Yapanel, U., & Gray, S. (2010). What automated vocal analysis reveals about the language learning environment of young children with autism. *Journal of Autism and Developmental Disorders*, *40*(5), 555–569. doi:10.1007/s10803-009-0902-5