Main research questions
1. Does expressive language development differ in children with and without siblings?
2. Is there an expressive language difference between girls and boys?
3. Is the expressive language development of children with siblings correlated to the expressive language ability of their siblings?

Background
What is expressive language? Expressive language is a spontaneous production of language. Mean Length of Utterance (MLU). Mean length of utterance as measured in words (MLU-w) is the average number of words produced in an utterance. An utterance is one turn-taking episode. MLU can also be measured in morphemes (MLU-m) (Ezeizabarrena & Fernandez 2018). MLU-w and MLU-m are effective as measures of language development in young children (Kaderavek 2015). MLU-w is easier and less complex to calculate than MLU-m (Parker & Bronson 2005). Here we use MLU-w as the dependent variable.

A new look at MLU: massive, natural recordings. Analyzing MLU from daylong audio recordings allows assessment of a child’s speech under natural conditions using a larger sample than is possible with other methods.

Method
Participants
16 traditional, two-parent families with typically developing children (7 boys, 9 girls) with an average age of 29.8 months. The socio-economic status (SES) averaged 7.81 (on a 15-point scale).

Materials
1. All audio, transcriptions, and metadata were collected from the HomeBank archive (VanDam et al 2016) using the VanDam Public Corpus (VanDam 2018).
2. Number of utterances and MLU-w were calculated using the CLAN software (Computerized Language Analysis) and the KidEVAL subroutine using the default settings.

Procedure
Raw, daylong audio was collected using the LENA system (Language ENVironment Analysis; LENA Foundation, Boulder, CO). The VanDam Public Corpus consists of 159 files from 53 families. The database includes files from families with children with a hearing loss not considered here. All recordings in the database are segments excised from daylong recordings that were processed with automatic speech recognition (ASR) software, then post-processed for linguistic and communicative content, and archived in five-minute segments that are fully transcribed by trained judges in CHAT format.

For this work, a total of 48 five-minute files were collected, all representing families of children who are typically-developing. All 48 segments were input into CLAN and analyzed for number of child utterances and both child and sibling mean length of utterance (MLU-w).

Data analysis, statistics
Mann-Whitney U-Tests were computed to test for differences between groups. A 95% confidence interval about the mean was computed and plotted in Fig. 1. Spearman’s ρ using exact permutation test is shown in Fig. 2. spearman’s ρ and least squares regression lines are shown.

Results
1. Only-children have a larger MLU than children with siblings, Fig. 1A.
2. Boys without siblings have a larger MLU than boys with siblings (Fig. 1D), but there is not a difference for girls (Fig. 1B).
3. MLU of children is not correlated to their siblings’ MLU, Fig. 2.

Discussion
Only children seem to have larger MLU, but the effect may be driven by boys, not girls. The main effect favoring the only-child families may be explained by greater opportunity for language use and practice. The interaction favoring boys who are only-children is unexplained. Volubility seems to be a personal trait, not a household or family trait for children. A larger sample will improve the ability to generalize findings.

Future directions

Acknowledgements
For raw data and to download the acoustic MAY be, get it from http://homebank.lilithlab.org/. 

References
VanDam, M. (2018). VanDam Public Corpus, Available at https://van dam.group.nest.berkeley.edu/

Acknowledgements
For raw data and to download the acoustic MAY be, get it from http://homebank.lilithlab.org/. 

Work supported by Tea Washington Research Foundation, SBE RIDIR, and NIH/NIDCD R01DC009569, 376.

For more information, see https://talkbank.homebank.org/.

For an online repository of daylong child centered audio recordings. See HomeBank (VanDam 2018).

For more information, see https://talkbank.homebank.org/.