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
Complexity measures. Language input quality has been shown to improve language outcomes (Rowe 2012; Weizman & Snow 2001; Huttonlocher et al 2010; Hurtado et al 2008).

Mean Length of Utterance. 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 means of measuring language development in young children (Kaderavek 2015), although its diagnostic clinical value has been challenged (Eisenberg, Fersko, & Lundgren (2018)).

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 LENA automatic speech processing (ASCP) software, 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 transcriptions were processed with CLAN for 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 distributions are shown with OLS regression lines in Fig. 2.

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 (240 minutes), 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 were calculated using CLAN software (MacWhinney 2000) and the KidEVAL subroutine using the default settings.

Results
1. Only-children have a larger MLU than children with siblings, Fig. 1A.
2. Only-child-boys 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).

Future directions
Examine differences in expressive language development in a hearing impaired (or disordered) population with and without siblings. 
Explore possible differences in expressive language development in children with younger versus older siblings, and two children of different age and sex distributions.

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. Vulnerability seems to be a personal trait, not a household or family trait for children. A larger sample will improve the ability to generalize findings.

This work is also a proof-of-concept to show how massive-database, ecologically valid recordings can bear on (old) questions in new ways. This work is extensible to a wide variety of similarly structured questions.