

RCC Computational Support for the Language Development Project

Richard Williams IV¹ and Susan Goldin-Meadow²

¹Research Computing Center, ²Department of Psychology, University of Chicago, Chicago, IL, USA



INTRODUCTION

Professor Susan Goldin-Meadow (principal investigator) and her research collaborators in the Language Development Project (LDP) investigate the development of human language, focusing on linguistic and gestural inputs and their consequences for child language and gesture in typically developing and brain injured children. This longitudinal study investigates caregiver and child language, the evolution of each over time, and the influence of interactions between caregivers and children on development. RCC provides data management, analytics, and application development support for the LDP.

Data Management and Analytics

Data management:

RCC manages the central SQLite database for the LDP containing annotated interactions between caregivers and children collected over the course of several years (over 1 million coded utterances). RCC also develops and implements data management systems for the backup and long-term archival of LDP primary video source data, derived data (transcripts and databases), and project documentation.

Analytics:

RCC computational scientists also provide analyses and reports from the LDP dataset and have processed over 108 requests for analyses on behalf of 27 researchers affiliated with the project.

Application Development

RCC develops applications and workflows to facilitate and validation annotations of spoken utterances and gestures. In order to better examine aspects of language and gesture usage in spontaneous interactions between caregivers and children a custom Python library was developed to provide advanced querying and reporting tools for the LDP SQLite database. RCC has developed tools in Ruby for validating coding annotations of transcribed caregiver/child interactions. A version of this coding validator has been developed for use within Datavyu, a video coding and data visualization software.

Summary and Acknowledgements

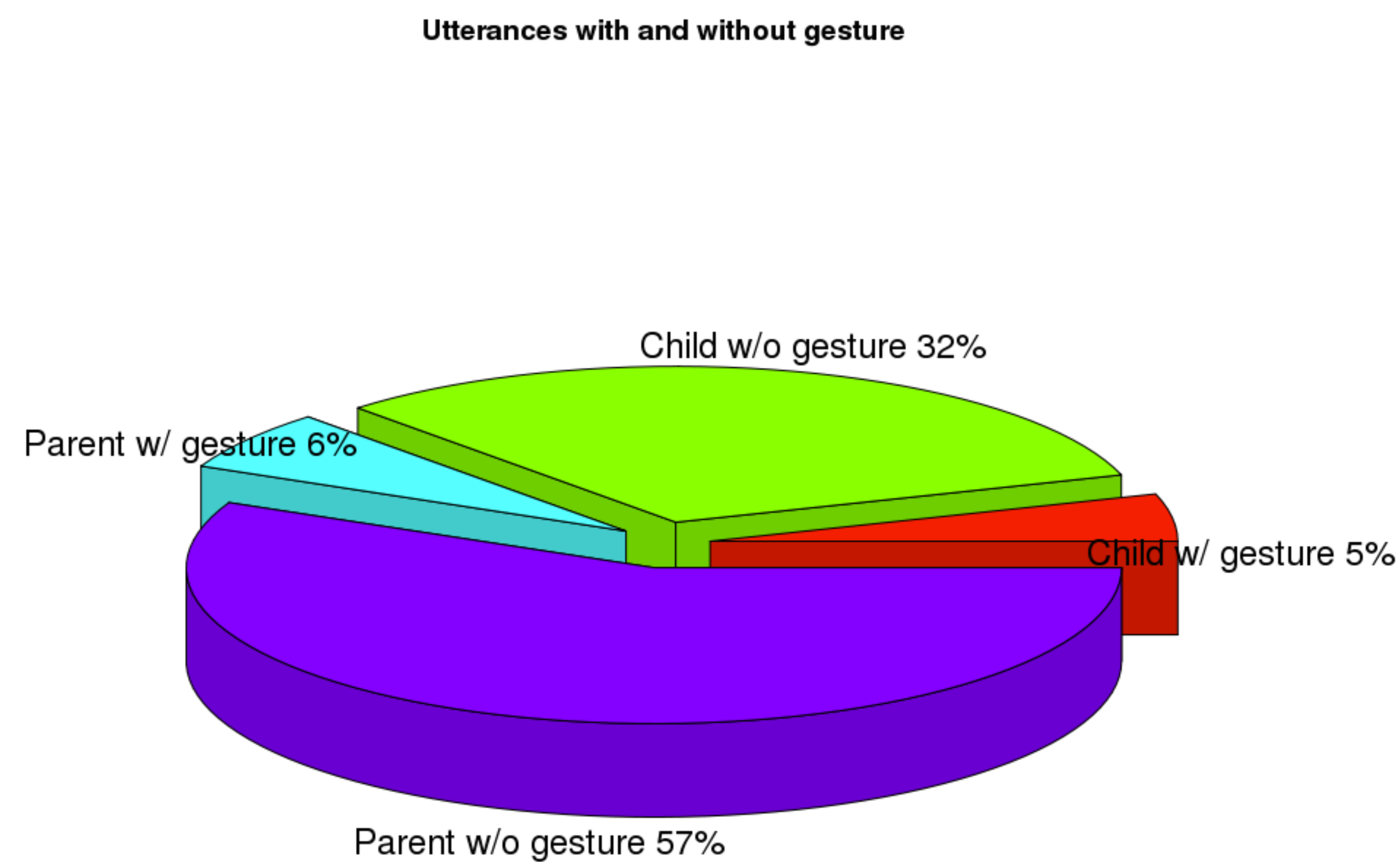
Summary:

In support of the Language Development Project, RCC provides technical and analytical expertise to enable robust analyses and the continued expansion of the multifaceted LDP longitudinal dataset. The tools and data management systems developed by RCC allow for increased accessibility of the LDP data by researchers in development, gesture, and linguistics.

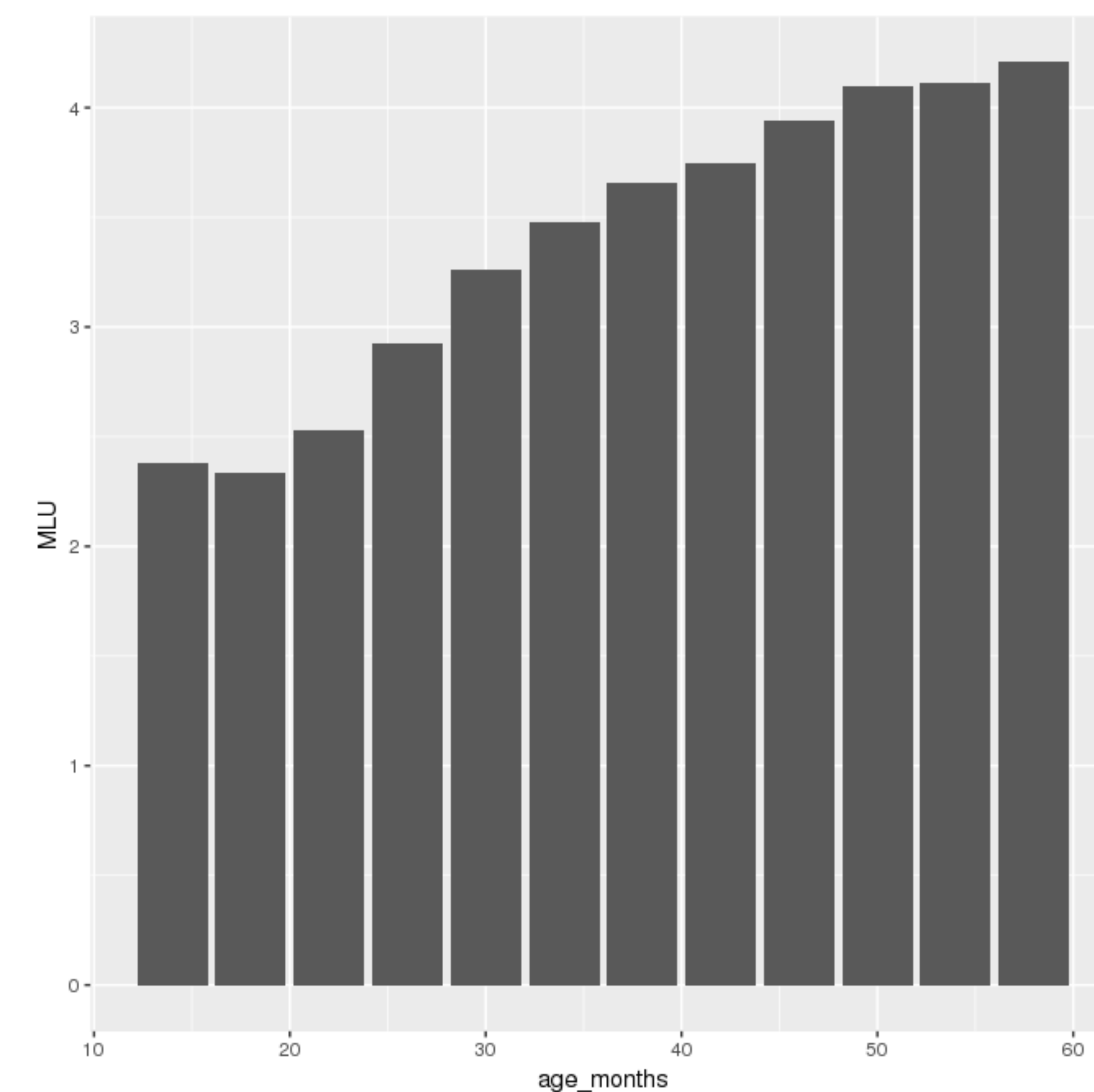
Acknowledgements:

Funding for this computational support is provided by the Goldin-Meadow Laboratory. We thank Kristi Schonwald for her technical assistance and Jason Voigt for developing the LDP Python library and initial LDP SQLite database.

1. Percentages of all utterances spoken by child or parent, with and without gesture.



3. Mean length of child utterances as a function of child age (in months)



2. Wordcloud visualization example Udisplaying most frequently spoken 100 words by one child (age 58 months) during a recorded session.



4. Mean number of gestures per session as a function of child age (in months) for children (A) and parents (B).

