**Demographic Biases in Naturalistic Language Recordings in the CHILDES Database**

Camila Scaff1,2, Georgia Loukatou3,, Alejandrina Cristia2, and Naomi Havron4

1 Institute of Evolutionary Medicine, University of Zurich

2 Laboratoire de Sciences Cognitives et de Psycholinguistique,

Département d’études cognitives, ENS, EHESS, CNRS, PSL University

3 Stanford University

4 Haifa University

**Email**: camillescaff@gmail.com

**Acknowledgements**

We are grateful to Lara Oliel and Clara Daoud for help entering corpus-level data; Sara Pisani for organizing communication with corpus curators; Brian MacWhinney for support and encouragement when retrieving information from corpus curators. We particularly want to thank the literally hundreds of researchers who contributed data to CHILDES, without which this project would be unthinkable; and the 103 curators who replied to our request.

AC acknowledges financial and institutional support from Agence Nationale de la Recherche (ANR-16-DATA-0004 ACLEW, ANR-17-EURE-0017, ANR-14-CE30-0003 MechELex); the J. S. McDonnell Foundation Understanding Human Cognition Scholar Award; and the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (ExELang, Grant agreement No. 101001095). GL acknowledges financial support from Fondation Fyssen.

**Conflict of interest statement**

None of the authors have a conflict of interest to disclose.