Stanford University

MICHAEL FRANK
Associate Professor
Department of Psychology

July 2nd, 2014

Dear Colleague,

My lab is beginning a new data-aggregation and data-sharing project that we call Wordbank. Wordbank (wordbank.stanford.edu) is a free and open database containing the full, item-level responses for MacArthur-Bates Communicative Development Inventory instruments (CDI forms). We hope that by pooling detailed word-learning data across labs, we can create a database of unprecedented size that will lead to new insights about the shape of vocabulary development in early childhood.

I am writing to ask whether you would consider contributing data to Wordbank. I hope that this contribution will be valuable both for you and for the research community more broadly. By contributing, you gain a permanent digital copy of your CDI data, available both in tabular form and in detailed reports for reanalysis and comparison with other data. In addition, by helping to increase the size of the overall Wordbank dataset, you help further research on the large-scale dynamics of vocabulary learning.

There are two routes to contribute. If you already have a digital version of your dataset available at the item level—meaning that you have already entered individual children's responses to particular words—you can share your digital data with us directly and we will upload and convert them. If you have paper surveys, we will be glad to digitize them as soon as we obtain sufficient funding. We will send you mailing materials, and will digitize, re-key, and return your data at no charge. The minimum demographic information we require to enter data into the database is the age and gender of each participant, the approximate date on which the survey was given, and the location.

Privacy is always a concern when creating open repositories for experimental data. Wordbank will contain only anonymized data. No identifiable information will ever be entered into the database. For example, we will store children's exact ages but neither their exact dates of birth nor the exact date on which they were tested. Wordbank has been approved by the Stanford University Institutional Review Board (IRB), Protocol #30911. We believe that sharing of anonymized data should be compatible with other IRB standards; if you are concerned, please check with your local IRB prior to contributing. More details are available at http://wordbank.stanford.edu/contribute.

I hope that you choose to contribute! Please do not hesitate to contact me if you have any further questions or concerns.

Sincerely,

Michael C. Frank
Associate Professor of Psychology