

Title: CHILDES and PhonBank resources for studying child language development

Begun in 1984, the CHILDES database includes language sample data from thousands of children speaking 34 languages. These corpora have been contributed by hundreds of child language researchers who have committed themselves to the scientific value of data-sharing. Most of the recently contributed corpora are linked to audio and video data for direct playback. There are now over 7500 published articles that have made use of CHILDES data. The corpora include dense longitudinal case studies, cross-sectional studies comparing children with language disorders with normally developing controls, narrative samples, studies of bilingual development, daylong recordings in the home, studies of the details of phonological development, and studies of dialect and social class variation. Because all of this data is transcribed using the CHAT format, the entire database can be analyzed quickly through the new TalkBankDB corpus search and analysis facility.

This workshop will cover these topics:

1. An introduction to the use of web pages to browse, download and analyze all of the freely available data in CHILDES, PhonBank, and FluencyBank.
2. An introduction to the use of the CLAN programs for data analysis.
3. How to use CLAN for language sample analysis (LSA) in the clinic.
4. How to use KIDEVAL to compare a given transcript with the larger database.
5. How to use TalkBankDB to download relevant datasets for further statistical analysis by R/SPSS/Python.

Webpages: <https://talkbank.org>, <https://childes.talkbank.org>, <https://phonbank.talkbank.org>, <https://fluency.talkbank.org>