Thursday April 8th, 2pm EDT
Presentation in Zoom, accessible via the C-STAR website:
http://cstar.sc.edu/lecture-series/

Data-Sharing to Understand Aphasia
Davida Fromm, PhD & Brian MacWhinney, PhD
Carnegie Mellon University

AphasiaBank is an open-access resource for the study of functional spoken language communication in people with aphasia (PWA). Supported since 2007 by a grant from NIDCD, the database and programs have been used in over 250 published papers. The largest segment of the database includes hour-long videotaped interviews with 397 English-speaking PWAs and 277 control participants. The interview follows a consistent protocol including the stroke story, picture descriptions, a retelling of the Cinderella story, and a procedural discourse task. A set of formal and informal tests are administered as well. Localized versions of the discourse protocol have been used with Croatian, French, Italian, Mandarin, Romanian, and Spanish. Additional data include the Western Aphasia Battery, and the verb naming test.

AphasiaBank is one of 14 databases in the larger TalkBank system, which also includes data from language in RHD, TBI, ASD, stuttering, dementia, SLI, and phonological disorders. Each of these banks uses data in the same consistent transcription format and the bulk of the data is linked to either audio or video media for playback both over the web and on the desktop. We have also constructed a set of analysis and profiling programs that rely on this consistent transcription format. Using methods from computational linguistics, the transcripts can be automatically analyzed for part-of-speech, morphological composition, and grammatical dependency structure. This information then further supports automatic methods for profiling through computational implementations of language assessment measures such as TTR, vocD, MATTR, MLU, QPA, NNLA, IPSyn, DSS, SUGAR, and others. There are several tools that support discourse analysis. The entire corpus can be searched online using the TalkBankDB database search engine that can track grammatical structures and n-grams and can output data in spreadsheet format to R and Python.

For educational purposes, segments of AphasiaBank and the other banks for language disorders include "Grand Rounds" instructional pages that exemplify clinical types, error types, and disfluency patterns. Transcripts with linked media can be played back directly online using the TalkBank Browser.

Planned extensions of AphasiaBank focus on the inclusion of other languages and speaker groups, work with dementia and apraxia, the analysis of conversational and gestural patterns, and greater reliance on automatic speech recognition.

_____

The online lecture can be followed online from your computer, tablet or smartphone, in **Zoom**. The zoom link is accessible via the C-STAR website: http://cstar.sc.edu/lecture-series/

For more information, or to be added to the C-STAR mailing list, contact Dirk den Ouden:
denouden@sc.edu