**History of Preparing the NSCM Data for Analysis**

1. Anjali created Master Data Files from the Log files in Matlab format. From these Master Files, she then created Excel spreadsheets of each trial per subject, with the stimulus, Examiner’s response(s), and Transcriber’s response (s) (which had been converted from hand-transcribed IPA to Darpabet, by Cynthia’s lab team in the spring and summer of 2010 (check these dates).
2. From the Master Data Files, Anjali generated confusion matrices (CM) in Excel, that Cynthia could read. There were four CMs per subject (initial consonants, final consonants, initial vowels, and final vowels). Tony was missing files from Oct. 12 and Nov. 2, even though there was hand transcription for these dates.
   1. Anjali noted “Exceptions” in the Excel CM’s for each subject, indicating the line on the Excel spreadsheet where something other than a CV or VC occurred.
3. Mary and Cynthia worked together and resolved every exception for every child. Cynthia sent these to Anjali.
4. On May 30, 2011, Anjali sent Cynthia an e-mail outlining the new work she had done. This included merging the two corrected Excel files from Cynthia and Mary (“corrected” means that all the Exceptions were resolved into the closest CV or VC approximation). Also, Anjali added in Tony’s missing sessions from Oct. 12 and Nov. 2. Finally, she said she generated:
   1. New CMs.
   2. New Master Data Files
   3. New scatter plots and bar plots.

These were all attached to her e-mail as “Reading Group 5\_30\_11.zip”

1. On June 6, 2011, Cynthia sent Anjali an e-mail listing 7 Exceptions that she and Mary found where the “Exceptions Resolved” were still not appearing in the Excel Spreadsheets.
2. On July 20, 2011, Anjali updated her NSCM analysis files and sent Cynthia an e-mail stating what she had done. This included:
   1. Updating the data file NSCM\_data
   2. Converting the Master files into an Excel Spreadsheet, with the total number of trials per subject.
   3. Updated the NSCM file.
   4. Updated the CMs in the folder, CM.
3. Around 1-30-12 and 2-3-12, Christoph plotted new CMs (four per subject) and percent correct tables (four per subject). He thought the numbers were off in the CMs, because he did not realize that we were plotting ALL responses from both the Examiner and Transcriber.
4. On 3-19-12 and 4-3-12, Christoph revised the percent correct tables for each subject, to include Cynthia and Mary’s hand-written corrections (from Excel files that Cynthia had sent to Anjali, to resolved the Exceptions). Evidently, Anjali had included these “Exceptions Resolved” (i.e., “corrections”) only in her Master Files (not in the Log Files in Matlab format).
5. We still need to revise the CMs to include the “Exceptions Resolved”. This includes adding the few hundred “Exceptions Resolved” into the Log Files in Matlab Format.

Log Files

↓

“Log Files in Matlab Format”

↙ ↘

Master Data Files (summary files) 1. CMs,

2. Trial by Trial Excel Spreadsheets (per subject):

Stimulus, Examiner and Transcriber’s Responses

↓

Mary and Cynthia’s “Exceptions Resolved”

(plus Toni’s Nov. 2 and Oct. 12 sessions,

which were previously omitted,

and Mary/Cynthia’s 7 remaining unresolved

exception listed in Cynthia’s 6-6-11 e-mail to

Anjali)

↓

CMs Corrected (to include “Exceptions Resolved”

and other corrections listed just above)