**Overview**Task: For each data set (28 total), segment the original file into individual files that each contain a single elicitation repetition. Then trim the silence at the end of every 3rd repetition. Then rename the segmented files according to this format:

List-##\_###A\_######.wav

* Replace ## with the list number (such as 06).
* Replace ###A with the elicitation number and repetition letter (as in 337b).
* Replace ###### with the speaker ID (such as CTB501).

Data should be organized into 4 lists (5-8), each with it’s own dedicated folder. Nested within each list folder should be 7 more folders, each representing 1 of the 7 speakers. These folders should be named according to the speaker’s 6-character ID (i.e. MFG512). Segmented files should be placed within the corresponding speaker folder.

Programs used: Audacity, Pratt, Bulk Rename Utility, SSH Secure Shell

A variety of procedures were used to segment and rename the lists:

**Attempted Automation**  
The original procedure involved using Audacity’s ‘find silence’ and ‘export multiple’ features to automatically segment the files. This procedure worked perfectly for about 5-6 data sets (of of 28 data sets). The first problem was that the silence-finder couldn’t trim the long silence following the 3rd repetition of each elicititation- after the segmenting task I had to manually trim every 3rd file (initially using Pratt, later adapting a more rapid solution with Audacity, explained below in greater detail). Problems were also encountered when speakers fluctuated in speech tempo: the silence finder struggled to separate the repetitions correctly. These errors in the original recording were transferred to the segmented files. Consequently, I had to comb through the segmented files manually, often deleting or splitting files as necessary. As a result, the bulk rename feature became useless since the segmented files were no longer in numerical order (due to deletions and splittings). The combination of these errors resulted in a lot of wasted time. Eventually, this procedure was abandoned and a new procedure was adopted.

**Crawling Over Soundwaves**This procedure involved scanning each original file, moving from left-to-right across the soundwave, inserting labels (using Ctrl+B) between each repetition. To avoid having to trim the silence after every 3rd repetition, an extra label was inserted. The downside is that, after exporting the segmented files, I had to manually delete every 4th file (which contained mostly silence and the prompt for the next elicitation). This meant that, once again, the bulk-rename-utility became useless. I just renamed each file manually using a combination of hotkeys:

Highlight first file to rename >> F2 to rename >> Ctrl+V to paste the name template (i.e. List-07\_000a\_MMT517) >> arrow keys to move cursor to number/letter to edit (124a, for example) >>> enter new number/letter >> tab to move to next file.

While doing this manual renaming procedure, I stopped every-so-often to check the the file name against the audio itself, comparing it against the list document (the table of elicitations). Any errors that occurred were fixed on the spot.

**Trimming Files with Audacity**This method is faster than trimming files with Praat, mostly because of the available keyboard shortcuts.

Run Audacity >> Ctrl+O for open dialog box >> click desired file to trim >> press enter >> click-drag to highlight silence at the end of the recording >> Ctrl+X or Delete to trim the silence >> Ctrl+Shift+E for export dialog box >> tap Enter twice >> Ctrl+W for close dialog box >> right arrow then Enter to close Audacity (if you don’t do this the windows will quickly pile up and strain the computer).

**Results**There were a few imperfections in the resulting files:

* List-6-EGS505: only 2 repetitions for #092
* List-6-FNL520: only 2 repetitions for #101, extra repetition for #094
* List-6-MMT517: only 2 repetitions for #129
* List-7-MFG512: missing elicitation #271
* List-8-CTB501: only 2 repetitions for #339
* List-8-FNL520: extra repetition for #314 and #315 (1 each)
* List-8-ECG503: only 2 repetitions for #206, missing elicitations #253, #254, and #270
* List-8-MFG512: only 2 repetitions for #328
* List-8-MMT517: only 2 repetitions for #327
* List-8-RCG500: only 2 repetitions for #295