

**Cleanup:** Batchalign output will need to be checked and cleaned up in various ways.

The outputs from **morphotag** and **align** should be checked using the CHECK and Chatter programs as described in the CLAN manual.

The output from **transcribe** will pass CHECK and Chatter, but it will need to be corrected in several ways.

You should use the ROLES command in CLAN to convert speaker codes with names such as PAR0 and PAR1 to things like INV and PAR, based on checking the first lines of the file to see how to best assign these. Please consult the CHAT manual section 7.2 for the list of possible roles. First make a backup version of your files. Then, the command is: `roles +c *.cha +1` which uses a `roles.cut` file like this:

PAR0	INV	Investigator
PAR1	PAR	Participant

The next steps are done by opening each file in the CLAN Editor. With the media file in the same folder as the transcript file, use CLAN's continuous playback command (Esc-8) to play back and listen through the whole transcript. This will take some time. As you go, you will want to fix these things.

#### 1. Utterance termination decisions

- split utterances where needed
- join utterances where needed

#### 2. Utterance termination markers to replace periods: ? ! +...

#### 3. Speaker ID -- using Tiers menu

#### 4. Accuracy of words transcribed

#### 5. Revision and repetition coding

#### 6. Phonological fragments and fillers

#### 7. Overlapping communicators

#### 8. Add Gem markers as needed

#### 9. Add [+ exc] post-code for non-task utterances (only for picture description, Cinderella, PBJ; not for free speech tasks)

For a list of CHAT coding symbols, see the cheatsheet at

<https://aphasia.talkbank.org/cheatsheet.doc>

Useful editor commands:

F4 – plays current bulleted line, also command-click

F5 – redo bulleting

Esc-8 – continuous play

Esc-A – open/hide bullets