

Batchalign can be downloaded from <https://github.com/talkbank>. It is a Python script, created by Houjun Liu, that uses either Whisper or Rev-AI automatic speech recognition (ASR), the Montreal Forced Aligner (MFA) and the Unix version of CLAN to go from raw audio to a well-formatted CHAT transcription for analysis by CLAN and inclusion in TalkBank. Batchalign also has the capacity to use [Universal Dependency](#) taggers within the [Stanza](#) framework to add %mor and %gra morphological tiers to a CHAT transcript in dozens of languages.

Installation - Docker

Batchalign is available either through Anaconda (macOS or Linux) or through a Docker container for all platforms. Because of difficulties installing Anaconda under some versions of Windows 11, use of the Docker version could be a good choice in some cases. To install the Docker version, go to <https://github.com/talkbank>, select Batchalign Docker and follow the instructions.

Installation - Anaconda for macOS

Batchalign is also available through the TalkBank channel in Conda. To get started, begin by installing Anaconda for macOS from [this site](#). You can find detailed Anaconda installation instructions [here](#). It is best if you are logged in as the admin user when installing Anaconda and for the rest of this procedure. The *pip* program should be installed as a part of the Anaconda installation.

To install and run Batchalign, you need to open a Terminal window. You can find the macOS Terminal application in /Applications/Utilities. This terminal window is where you will enter Batchalign commands. When you first open this window, it will display the word "base" at the beginning of the command line prompt. This reflects the fact that Anaconda takes over the shell, which is usually zsh.

Conda will assume that your default working directory is /Users/yourname. Terminal commands refer to this location as ~ (tilde). So, a location like ~/mfa_data is equivalent to /Users/yourname/mfa_data To make sure that you have zsh initialized, just run "conda init" after completing installation.

Within the Terminal window, please enter the following commands in order, pressing the "enter" key after each one.

```
conda install mamba -c conda-forge
conda create -n batchalign
conda activate batchalign
mamba install batchalign -c talkbank -c conda-forge
pip install rev_ai
```

At this point, Batchalign should be installed.

Usage

Once installed, the usage of Batchalign follows these steps. These instructions supercede the earlier descriptions in our 2023 published article in JSLHR.

1. At the level of ~/ you should create a folder using: `mkdir mfa_data`. Then use `cd mfa_data` to go inside that folder and create subfolders using: `mkdir input` and `mkdir output`.
2. Next, you will need to prepare your audio or video file and put it inside the ~/mfa_data/input folder. Batchalign uses ffmpeg to convert video to .wav. If you record with iPhone, the format is m4a. Since batchalign only accepts mp3, mp4, and wav, we need to convert .m4a to .wav. You can do this using an online converter site such as this one: <https://cloudconvert.com/m4a-to-wav>
3. Batchalign supports different processes with different verbs. The three most used are marked with an asterisk:
 - ***align** produces utterance- and word-level alignment of a bulleted text when you place both the media and transcript files into /input.
 - ***morphotag** uses Stanza/UD to add %mor and %gra lines to a transcript. It does not require a media file.
 - ***transcribe** provides transcription directly from audio or video. This only requires raw media files (audio or video) in /input.
 - **bulletize** creates time bullets for files that were not originally linked to media.
 - **clean** empties the input and output folders.
 - **recursive** allows you to process a full hierarchy of folders in the /input folder
 - **version** lists the version of batchalign.
 - **sph** allows for .sph and .stm file format in /input.
 - **benchmark** compares ASR output with human transcription in /input
4. You can use either Whisper or Rev-AI for transcription. The default mode uses Rev-AI. For this, you will need to open a rev.ai account. Rev-AI provides you with 6 free hours for your new account. Charges are \$.02/minute of audio for this service. Go to rev.ai, sign up, and on the left side of your dashboard, you will find a tab called Access Token. Click generate to generate a new token, copy and paste the key to somewhere you can find later.
5. We are only one step away! Now make sure you are in the Batchalign environment by typing "conda activate batchalign" and then run the basic Batchalign command:
 - `batchalign transcribe ~/mfa_data/input ~/mfa_data/output`
6. The first time you run, the program will take about 5 minutes to download the material that will go into ~/mfa_data/model. After that, the program will ask for your rev.ai key that you obtained in step 5. You can cut and paste that from the place where you saved it earlier.
7. Next you wait as MFA works to align the Rev-AI transcript with the audio. When complete, you get back to the shell "batchalign" prompt and you should see a transcribed CHAT (*.cha) file in the ~/mfa_data/output folder!
8. We have adjusted Whisper to perform nearly as well as Rev-AI. Also, Whisper does better than Rev-AI for languages other than English, such as Spanish. Although Whisper runs much more slowly than Rev-AI, some projects may prefer Whisper's local mode of operation. To

use Whisper instead of Rev_AI, you must use this command to install several additional utilities:

- `pip install plotly imblearn pydub eyed3 hmmlearn pyAudioAnalysis`
9. Then you need to add the `--whisper` switch (two dashes) to your transcribe command, along with the `-n` (one dash) command and a numerical argument to specify the number of speakers, as in this command:
- `Batchalign transcribe --whisper -n 2 ~/mfa_data/input ~/mfa_data/output`
 - Note: your first download of Whisper can take up to an hour.

Cleanup

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

1. The outputs from **morphotag** and **align** should be checked using the CHECK and Chatter programs as described in the CLAN manual.
2. The output from **transcribe** will pass CHECK and Chatter, but it will need to be corrected in several ways.
3. To facilitate this process, it is best to first remove the %wor line using this CLAN command: `trim -t%wor *.cha +1` You can add it back later in a second pass of **align**.
4. The next steps are done by opening each file in the CLAN Editor.
5. First, you will need to correct the 4-letter Speaker IDs created by Rev-AI from the form PAR0 or PAR1 to standard CHAT roles such as CHI for Target_Child, MOT for Mother, or INV for Investigator. Please consult the CHAT manual section 7.2 for the list of possible roles. You need to correct this both in the @Participants line and the @ID fields and later on throughout the transcript. You can get new correct @ID fields by deleting the old ones and then running escape-L to check the file which then automatically enters new ones based on the @Participants line.
6. Once you have the correct set of Participant IDs, you can choose “update” in the Tiers menu to get shortcuts for insertion of each ID as needed. Also, you can use CLAN’s query-replace function (command-R) to change forms like *PAR0: to *CHI: throughout the file as needed.
7. 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 is the major job you need to do and it will take some time. As you go, you will want to fix three things.
 - a. You can fix incorrect IDs using the shortcuts such as command-1 in the Tiers menu.
 - b. You can correct incorrect words by retyping the correct forms.
 - c. You can correct incorrect utterance segmentation by joining two utterances into one or breaking up an utterance into two pieces.
8. Once you are done with these fixes, you can run the file through the **align** version of Batchalign to restore the %wor lines.

Further Explanations:

1. Examples:

```
batchalign align ~/mfa_data/input ~/mfa_data/output
batchalign transcribe ~/mfa_data/input ~/mfa_data/output
batchalign transcribe ~/mfa_data/input ~/mfa_data/output --lang=es
batchalign morphotag --lang=nl ~/mfa_data/input ~/mfa_data/output
```

2. Updating:

```
(base) conda activate batchalign
(batchalign) mamba update batchalign -c talkbank -c conda-forge
```

3. **Languages:** To run the three basic commands (align, transcribe, morphotag) for languages other than English, you need to add the `--lang` switch, as in `--lang=de` for working with German (Deutsch). The two-letter ISO-639 abbreviations can be found in the CHAT manual.

4. **Morphotag:** To tag transcripts for %mor and %gra using Universal Dependency models, you use commands in this form:

```
(batchalign) batchalign morphotag --lang=es ~/mfa_data/input ~/mfa_data/output
```

5. **Recursive:** This verb lets you place a collection of folders into your /input folder to be run recursively. You then enter the actual command in a further quoted string, as in:

```
(batchalign) batchalign recursive ~/mfa_data/input ~/mfa_data/output "morphotag lang=de"
```

6. Fixing:

```
(batchalign) conda deactivate
(base) conda env remove -n batchalign
(base) conda create -n batchalign
(base) conda activate batchalign
(batchalign) mamba install batchalign -c talkbank -c conda-forge
(batchalign) pip install rev_ai
```

Installation - Anaconda for Windows

The Anaconda versions of Batchalign works for some but not all Windows systems. It works for Windows 10, but not all versions of Windows 11. Once installed, you should have a program exactly named "Anaconda Prompt". Use that instead of "Anaconda PowerShell" or "Anaconda Shell". In the "Anaconda Prompt" environment, run these commands:

- `conda install mamba -c conda-forge`
- `conda create -n batchalign`
- `conda activate batchalign`
- `mamba install batchalign -c talkbank -c conda-forge`

- `pip install rev_ai`

Next, you should create two folders `C:\Users\yourname\mfa_data\input` and `C:\Users\yourname\mfa_data\output`. Place your input file in `C:\Users\yourname\mfa_data\input`, then, run:

- `conda activate batchalign`
- `batchalign transcribe C:\Users\yourname\mfa_data\input C:\Users\yourname\mfa_data\output`

Possible glitches that can occur during installation include failure to install punkt. To do this, run these instructions:

- `conda activate batchalign`
- `python`
- `import nltk`
- `nltk.download("punkt")`

Also, if you wish to use Whisper, you need to run this command:

```
pip install plotly imblearn pydub eyed3 hmmlearn pyAudioAnalysis
```