Workshop 1: Transcription with ELAN

Zac Boyd

zacboyd@cc.au.dk

AU Phonetics & Phonology Research Group

01 Dec. 2022

Funded by
Carlsberg Foundation

Plan

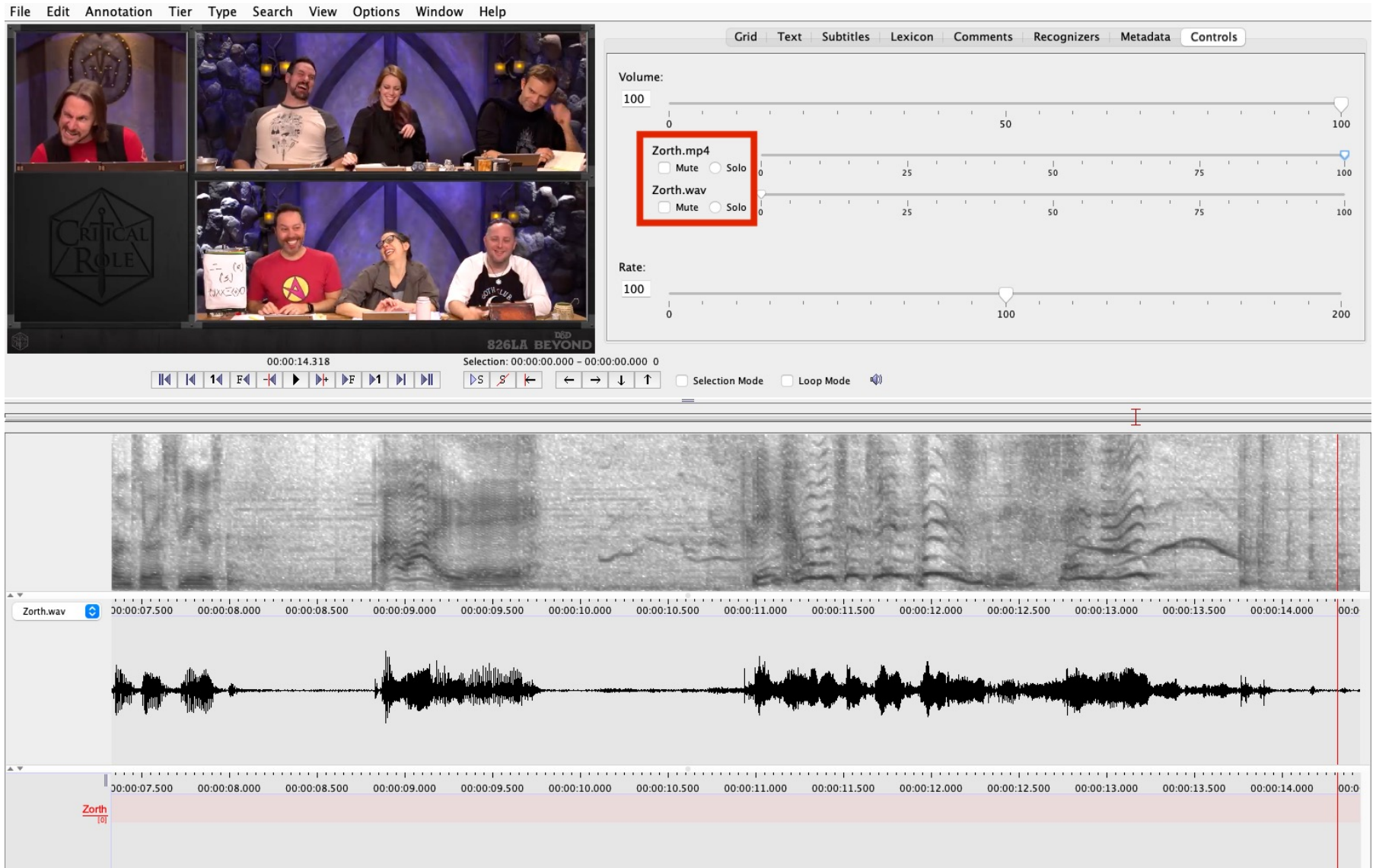
- What is ELAN & Why do I suggest it for Linguistic Analysis?
- Transcribing with Forced Alignment in mind
- Basics of ELAN
- ELAN as a tool for analysis
  - Setting up Tiers
  - Coding for Dependent and Independent Factors
- Time to play with ELAN!

Why ELAN

ELAN is an annotation software which can time-align any number of annotations on audio/visual recording data.

- Transcribing hours and hours of audio is soul suckingly boring and it sucks super hardcore – ELAN at least makes it suck ‘slightly’ less
  - the average time to transcribe one hour of audio is approximately four hours or more
  - this can greatly increase with multiple (overlapping) speakers and any number of other factors
- Ease of use
  - Esp. when you have the key shortcuts memorized
- Exports time aligned transcriptions directly to .TextGrid files (among others if you need that for some reason)
  - This is what allows the aligner to work

ELAN – Audio & Video



Note: You will need to import both the Audio and the Video to get both the video and waveform/spectrogram to appear

This can be done when starting a new ELAN project via File -> New -> 'Add media files…'
– just add both at the time of creation

(Audacity is capable of extracting audio from .mp4 and other video files - just save as .wav)

Transcription Guidelines

Best practices always follow transcription guidelines

- Consistency (Within the dataset & across various annotators working with the same dataset)
- Ensures that Forced Alignment will work consistently

  (or will, at the least, fail consistently so you can manually fix it where you know it needs fixing)
- (Generally – and for Montreal Forced Alignment (MFA)) The exact set of guidelines you follow don't matter as much as the fact that you follow some

Transcription Guidelines

- I use a modified version of FAVE Guidelines
  - Logical & Easy to remember
  - Works with the Montreal Forced Aligner after little modification

https://www.ling.upenn.edu/~wlabov/L560/Transcription_guidelines_FAAV.pdf

This should be obvious, but…

With all transcription, remember to write down exactly what was said, not what would be "correct" or what you think should have been said.

This includes mispronunciations, false starts, etc.

Transcribing for Forced Alignment

- If the end goal is Forced Alignment:
  - Strictly follow the same conventions – Essential for Consistency in Alignment
  - Know what guidelines your aligner requires (if any)

    (For example, FAVE-Align has stricter guidelines than MFA – MFA does not have 'official' guidelines – that I know of - which lets us adapt)
  - (In my experience) I've found that aiming for 5-10 seconds of audio per phrase has the best results
    - You can split a phrase into smaller sections if need be – just find a phoneme & word boundary with no co-articulation (Voiceless stops work great for this)
    - Very long phrases and very short phrases tend introduce a greater chance that the alignment will be off (esp. in conversational & faster speech)
    - This tends to be less of an issue for short word list items (the careful speech style helps here)

Transcribing for Forced Alignment

- The aligner doesn't care about grammar or punctuation
- It *does* care about disfluent speech
  - If there is sound in the file that overlaps with a time-aligned phrase, the aligner is going to try to force align it (hence the need for consistency in your guidelines regarding disfluent speech and noise)
- Contractions
  - The only time punctuation might matter
    - *E.g. - it's* vs. *its* would be aligned the same
    - For something like *wouldn't* the apostrophe is necessary (because *wouldnt* without the punctuation is likely not a word in the pronunciation dictionary)

Transcribing for Forced Alignment

- Conventions I use that work with MFA:

Transcription	Description of the Noise	
{BR}	breath	(Speaker takes an audible breath)
{CG}	cough	(The speaker coughs, or clears their throat)
{LS}	lip smack	(The speaker smacks their lips)
{LG}	laughter	(The speaker laughs)
{NS}	noise	(Audible background noise that overlaps speech)

- MFA will interpret these as an unknown word and will bypass that section of the speech when aligning until it comes to the next known word

Transcribing for Forced Alignment

Other best practices:

- Numbers – Written out (generally good practice but essential for aligners)
  - *i.e.* forty-two
- Acronyms and Spoken Letters
  - Use all caps separated by spaces: "All this transcription makes we want to drink an I P A"
- *Pronounced* Acronyms
  - Spell out in all capital letters with no space – *e.g.* NASA
- Partial words
  - the conventions will vary based on aligner… for MFA just put in the partial word (it may or may not align that word)
  - "Well I suppo- I think so yes"
- (For MFA) Unclear speech can be left as an empty phrase, noted with {NS}, or I occasionally write what I think it is within double parentheticals (( ))
  - MFA doesn't recognize the double parenthetical as potentially wrong speech and will attempt to align what you wrote (again, it doesn't care about punctuation – (( )) is to make it easier to find and check after alignment).

Transcribing for Forced Alignment

- In general, just think: how would someone dead inside listen to this and transcribe it for them
  - i.e., make your transcription explicitly clear and you should be fine
  - Any issues where an aligner doesn't like your transcription can generally be fixed (More on that next week)

Tiers – aka "How to code variables with ELAN"

- Annotations can be created on multiple layers, or "Tiers"
  - Tiers can be hierarchically created
- Annotations are time-aligned to the media – but can also refer back to other existing annotations

Why would you use this?

- Can independently transcribe overlapping speakers
- Can code for speech type/task in longer recording session
  - i.e., you can have a separate tier which indicates Semantic Differentials, Minimal Pairs, Reading Passage, etc. for easy reference.
  - Coding for speech action in Conversation Analysis

Getting Started with ELAN

- Files are organized into “projects”
- ELAN projects consists of: one or more media file(s) (sound or video), and one .eaf annotation file
- ELAN links annotation files with their corresponding sound files.
  - Every time you open an ELAN file, it will automatically search for the sound file belonging to your transcription file.
  - If ELAN cannot find the sound file at the specified, it will ask you for the location of the sound file
  - Basically, just always keep the .wav and .eaf files in the same folder with the same name

ELAN Basics - Lets get some transcription

- The rest of today is just basics of transcription in ELAN
- ELAN has a huge amount of functionality – I exclusively use it with Forced Alignment in mind
  - As that's next week's workshop, that's what I'm focusing on

Getting Started with ELAN

- To begin annotation go to **File -> New…**
- Click on **Add Media File…** and locate the audio you wish to transcribe. Lets start with Audio1_Marion.
- You should now see the ELAN window. First thing, Save the .eaf in the same folder as the audio with the same name as the .wav
- At the top of the window you'll see "Grid", "Text", "Subtitles", "Lexicon", "Comments", "Recognizers", "Metadata", and "Controls".
  - The lexicon, recognizer, and metadata don't really serve any purpose for us

Getting Started with ELAN

Controls window:

- Basic self-explanatory functions
  - Adjust the volume
    - If you have multiple media files you can adjust them individually.
  - Playback speed (“Rate”)

Grid, Text, and Subtitles:

- offer different ways to view your transcriptions.

Lexicon:

- I’ve never once touched this tab and have no idea what it’s for.

ELAN Basics

- To create a new tier, go to **Tier** -> **Add new tier…**
- Tier name: This could be the speaker (or whatever pseudonym or code for the speaker you have). You may also have a “Noise” Tier, or anything else you’re coding.
- Participant: Can be used to identify the participant if not included within the Tier Name
- Transcriber: Good for projects with multiple annotators.

- This is also where you would add sub-tiers if using those.

Tier Name is the only place you need to fill out. The rest can be empty.

CARLSBERG FOUNDATION

AARHUS UNIVERSITY

ELAN Basics

ELAN Basics

The waveform is a bit small. It's easier to work with once you can see it better. Right click (or control click) on the waveform and go to **Vertical Zoom** and change that to 300% (Sometimes you won't need this, but it's good to know!)

Now let's get to work on some transcription. First we need to segment it so we know where to place the annotations.

In **Options** you'll see various *Modes*.

- Annotation, Transcription, and Segmentation are what we'll look at today.

ELAN Basics

Go to **Options** -> **Segmentation Mode**

At the top you'll see various options. The only two that matter for us are **Two-keystrokes per annotation** & **One-keystroke per annotation**

This comes down to personal preference. I prefer One-keystroke because I'm lazy.

Select one and then click on the waveform below.

You can now begin to add boundaries. Do so when there is no speaking sound- you don't want to interrupt word or phoneme boundaries here. They can be deleted with Right-click or moved with the mouse by clicking and dragging.

ELAN Basics

- Once the file is segmented we can begin transcribing
- Go to **Options** -> **Transcription Mode**
- You'll then select the type of tiers you wish to transcribe (you may only have one option, or multiple depending on the audio) – Usually this is just default.
- Click **Apply**
- Click the first empty box, and begin transcribing
  - The audio will not move on from the current selection until you hit **enter**

ELAN Basics

- Congrats, you've transcribed something. Well done.

- Lets export it.
- **File -> Export As -> Praat TextGrid…**
  - Here you can select which speakers you want (if you have multiple), among other things.
  - Export it into the same folder as everything else, and you now have a time-aligned .TextGrid file

Shortcuts to know and remapping

For best results with ELAN it's good to learn some shortcuts.

Specifically:

Play media -> **Shift + Space**

Play Selection -> **Control** (Mac) **Command** (PC)

+/- 1 second -> **Shift +** ←/→

You can view and rebind the shortcuts in:

Edit -> Preferences -> Edit Shortcuts…

(be warned…. There's a LOT of them)

More Info

ELAN's Full 425 Page Manual

- https://www.mpi.nl/tools/elan/docs/ELAN_manual.pdf