

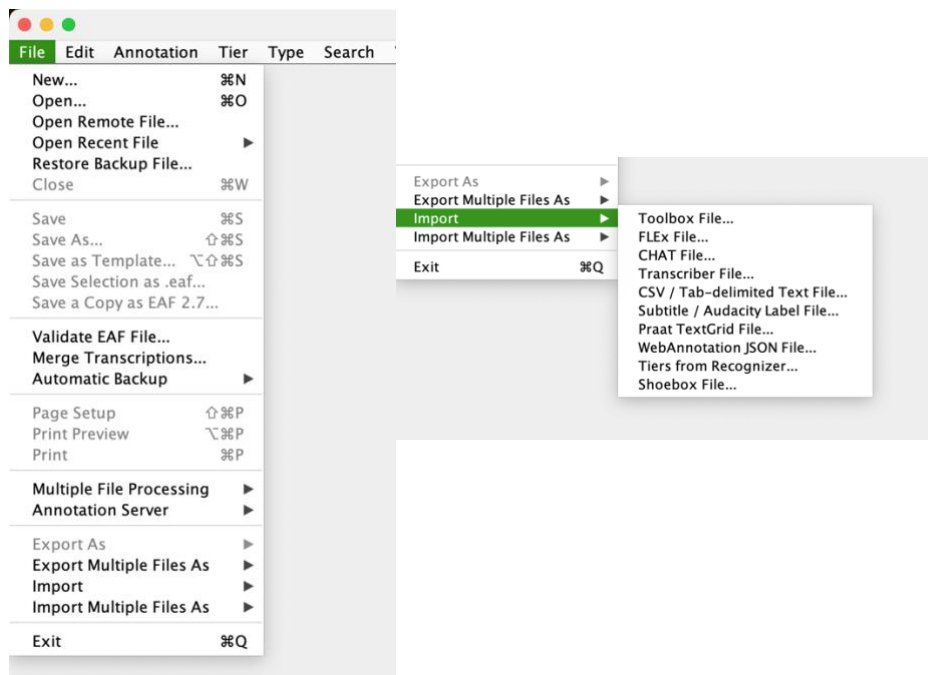
How to Import .TextGrids into ELAN

Usually, the .eaf base file is anonymized before an external program for the force-alignment of individual phones, for example WebMaus, uses it to produce a corresponding .TextGrid in PRAAT. Ideally this allows for a 1:1 match between the tiers of the .eaf and the .TextGrid allowing for easier segment analysis.

Keep in mind:

- Make sure to use the latest version of ELAN to minimize glitches and crashes during the importing process (it is not recommended to keep multiple versions of ELAN on the same computer)
- Make sure you are importing ANONYMIZED .TextGrids into ANONYMIZED .eaf's
- Make sure to import FORCED ALIGNED .TextGrids (e.g. WebMaus), not manually aligned nor VOT .TextGrids
- The alignment between the .eaf tiers and the .TextGrid tiers may not be exactly one-to-one (this may be due to further tier manipulation during the anonymization process)

1. Locate the ANONYMIZED .eaf file
2. Open the .eaf file in the latest version of ELAN
3. Open File > Import > Praat TextGrid File



4. Browse your files for the corresponding ANONYMIZED .TextGrid file (or with the Webmaus indicator)
 - a. It's handy to store the .eaf and the .TextGrid into the same working directory for a structured workflow.

5. Check the box next to “Include Praat PointTiers” as this forces ELAN to check individual phones.
6. Keep the default PointTier annotation duration to the pre-set 40ms.
7. (Optional) You may check the box next to “Skip empty interval / annotations” (I haven’t seen a significant difference if you check that box or not).

Select a Praat TextGrid file containing Interval tiers

Include Praat PointTiers

Default PointTier annotation duration (ms)

Skip empty intervals / annotations

Attempt to auto detect wav file

8. Click “Next” > Select Encoding
 - a. It seems that the Encoding you select depends on the language you’re looking at; each HLVC team used a different encoding
 - i. Italian uses “default” encoding
 - ii. Cantonese, Russian use “UTF-8” encoding
 - iii. Tagalog uses “UTF-16” encoding
 - b. Click “Okay”

Include Praat PointTiers

Default PointTier annotation duration (ms)

Skip empty intervals / annotations

Attempt to auto detect wav file

9. Elan will now create tiers and annotations from the Praat source file, the speed at which this process loads depends on the size of the .TextGrid.
 - a. Files over 100 Mb tend to take 10-20 minutes on the Lab Macs or freeze on older computers
 - b. Files under 100 Mb tend to resolve rather fast
10. Check to see if the tiers roughly match up to the .eaf tiers (there may be some delay or mismatch); but the content must be the same.
11. Make sure to save the .eaf file as SPEAKERCODE_anon_aligned.eaf.
12. Transfer these files to the correct language, under the “ANON_ALIGNED .eaf’s” folder.
13. Make sure to update the HL catalog to show that file has been force-aligned.