

Exercise 8

Segmenting & Exporting Files Quickly

In this exercise we will be using (semi-)automated tools in ELAN & Audacity to segment a wordlist and export them as separate files. We will learn how to work with ‘Audio Recognizer’ in ELAN and how to export ELAN files into Audacity’s ‘labels’.

Today’s files are found in the folder called ‘SNA-Velarization-Wordlist’.

Step one: Create a new ELAN file

As you have done many times now, add/change tiers & types. You should have two types & two tiers:

Types: *text & translation*

Tiers: *Shona Word & English Gloss*

Step two: Set up the ‘Audio Recognizer’

In the top pane, there are a number of tabs: ‘Grid’, ‘Text’, ‘Subtitles’, ‘Audio Recognizer’, ‘Meta Data’, ‘Controls’.

1. Select ‘Audio Recognizer’.

You will see a number of separate drop-down menus & boxes.

Recognizer: ‘Silence recognizer MPI-PL’ (We will use this recognizer).

File(s): ‘SNA-Velarization-Wordlist.wav’ should already be selected.

Selections: Find a flat (silent) section of the waveform and select a small portion—it need not be long. Click ‘Add’. This will be used as the baseline for which silence is detected by the audio recognizer.

Parameters: ‘Minimal Silence Duration’ & ‘Minimal Non Silence Duration’ are set to 200 Milliseconds. You can leave this for now.

Progress: You can now click ‘Start’.

Once you hit start, blue lines will appear on the waveform some with an ‘x’ (non silence) and others with an ‘s’ (silence).

Try adjusting the ‘Minimal Silence Duration’ to 300 or 400 Milliseconds. What happens?

Step three: Create tiers from ‘Audio Recognizer’

Now you can create tiers from the silences and/or non-silences.

1. Click on ‘Create Tier(s)...’ in the bottom left corner of the ‘Audio Recognizer’ tab in the top pane.

This opens up the ‘Create tiers from segments’ dialog box. You’ll see several things in this box:

- **Select a segmentation**, which allows you to choose the a separate channel if you have a stereo recording. Because we have a mono recording, we only have one option here.
- **Select and configure segments**, which allows you to select whether you want to create tiers from noise or silence or both. We want to make annotations from noise, so uncheck ‘s’ under ‘Include in tier’.
- Leave everything else as is & click **Create**.

Step four: Delete, Adjust & Input annotations

You will see that ELAN automatically created annotations for you. However, these will probably not be 100% correct, so you may need to do some cleanup. Use the file: `SNA-Velarization-Wordlist.txt` and input your words.

Step five: Export to Audacity

Once you have all of your annotations, you will now export your wordlist to Audacity.

1. Go to **File > Export As > Tab-delimited text...**
2. On the top of the dialogue box, check the box with the tier ‘Shona Word’ and make sure all other tiers are unchecked.
3. Make sure that ‘Exclude tier names from output’, ‘Begin time’, ‘End time’, & ‘ss.msec’ are all checked. Nothing else should be checked.
4. Name the file something like ‘SNA-Velarization-Wordlist-Export.txt’

Step six: Import into Audacity & Export to individual .wav files

1. Open Audacity.
2. Go to **File > Open...** and select the audio file ‘SNA-Velarization-Wordlist.wav’
3. Next go to **File > Import > Labels...**

4. You will now see the waveform with all of your labels (annotations). Next you want to export your files. Go to **File > Export Multiple...**
5. In the **Export Multiple** dialogue box, keep the '**Export Multiple**' as '*Other uncompressed files*'. In the box entitled '**Split files based on**' choose '*Labels*' & uncheck all other boxes. In the box entitled '**Name files:**' choose '*Use Label/Track Name*'. Finally, uncheck '**Overwrite existing files**'.
6. Click OK. (Audacity may ask you to do this a number of times.)

Now, you should have a folder full of individual .wav files!