

Instructions for Reviewers

The aim is to see how well human reviewers perform speaker diarization based on hearing only, not on visual elements (i.e. no inspecting the signal graph or the spectrogram).

The time for the review is rigorously limited to 45 minutes for a speech signal of 5 minutes to provide a realistic view of human performance in realtime (more than 5 minutes needed for the practicalities of setting the labels and rewinding the audio at times).

Absolutely no collaboration or discussion between reviewers.

Pauses of the same speaker during and between sentences should not be considered breaks unless the pauses are significant - 0.3 s and longer is a significant pause, though there will not be time in the review to accurately measure the pauses.

What the speakers are saying is not important, though complete sentences can be used to help determine whether it is the same person speaking.

Non-lexical speaker sounds (e.g. laughter, coughs, sneezes, breathing, lipsmacks) and non-speech sounds (e.g. door slams) should be ignored.

Vocal sounds that are not proper words do count though (e.g. “hmm”, “um”).

Reviewers will be asked what they think the environment is.

Legal Points

Results will be anonymised on GitHub page.

Audacity Hints and Shortcuts

Simon as administrator will open the file, set the 30 to 330 s range and demonstrate a simple example of how to annotate a speech file in Audacity.

Two windows: (a) the usual signal window; and (b) the labels window.


In the signal window, click outside the 1 to -1 range to remove visible signal.

Use spacebar to start/stop and the mouse cursor to move the time bar.

Use pause or its shortcut “p” (when cursor is in signal window) to stop time bar but keep it visible.

Ctrl-b will put the label start where the signal window cursor is when stationary.

Ctrl-m will put the label start where you are currently listening, and can be moved back a bit if you are slow to respond.

Drag the right side of  to move the label end. Do not drag the middle part as that moves the whole thing.

To delete a label, right click in the box and select “Delete Label” from the dropdown menu. This will not affect any other labels.

Number the speaker in the label depending on whether you have heard them before. i.e. start with “1” for the first speaker you hear, “2” for the second etc., reusing numbers if you hear a speaker again.

Expert Audacity users could use scrubbing toolbar to place labels more accurately, but in general this is an unnecessary complication.