

Identifying Unaligned Utterances

Under some circumstances, [forced alignment](#) can fail to produce alignments for some utterances; i.e. the utterance has no phone annotations created, the words are not aligned, and no htk annotation is created. This can happen because of the following factors:

- Not enough data (if you're using the 'train-and-align' approach)
- Poor quality recording, background noises, etc.
- Simultaneous speech (ignored by default)
- Inaccurate transcripts
- Inaccurate utterance alignment
- Lack of pause marking in the transcripts
- Mismatched phonology between dictionary and speech
e.g. using a rhotic dictionary to align non-rhotic speech

You can identify the utterances for which alignment has failed using LaBB-CAT's search and export functionality:

1. Click *search* and select the speaker(s) you aligned.
2. The search should be "the first word of each utterance that doesn't have an htk annotation" - i.e.:
 - *orthography* layer: matches .+
 - *utterance* layer: tick the left-hand checkbox that anchors the word to the beginning of the utterance

Search Matrix:

The screenshot shows the 'Search Matrix' interface. It consists of three main sections, each with a title and a search configuration area. The first section is titled 'utterance' and has a purple checkbox with a white checkmark on the left. The second section is titled 'htk' and has a white radio button on the left. The third section is titled 'orthography' and has a dropdown menu set to 'match' and a text input field containing '.*'. Each section has a small square button on the right side.

- *htk*¹ layer: doesn't match .+.

3. When the results are listed, click *CSV Export*

¹ *htk* or whatever the phrase tag layer is in the forced alignment configuration

The resulting file has the start and end time of each utterance in the *Line* and *LineEnd* columns. If you want to know the total duration of the unaligned utterances, use Excel or R to calculate the difference between *LineEnd* and *Line* to get the line duration, and then sum these durations to get the total, which is in seconds.