

There are three main steps: 1) calculating the PSDs, which takes a long time. Main function is `lv_preparepsds`. There are others to curate the 3000 syllables to build the model, but you can use your own batch files and segmentation here. Best to have no segmentation problems or noise files included in the set of 3000 syllables. 2) build the model: `lv_clusterpsds`. I don't just run this but go through cell by cell and check if the clustering looks good. 3) apply the model to more data: `lv_labelpsds` (need to have psds prepared for that as well)

1. Prepare PSDs

```
lt_make_batch(1)
```

```
lv_randsample_batch('batch.keep',30,9) %random sample a number of songs
```

Possibly: `lv_plotallsongs` to handselect songs without noise files

```
lv_resegment('batch.keep.lvrnd.keep','evtaf')
```

Check segmentation `evsonganaly('batch.keep.lvrnd.keep')`

```
lv_countsegments('batch.keep.lvrnd.keep')
```

Make sure you have ≥ 3000 segments, otherwise go back and `randsample` more

```
lv_preparepsds('batch.keep.lvrnd.keep','evtaf',5) - this will take a while
```

2. Build model from 3000 syllables

Select index of 20 basis syllables by hand, make sure you have at least 1 of each type, put it in `lv_clusterpsds`;

```
lv_clusterpsds
```

Go through cell by cell, check parameters

When you get to the step where all the clusters are shown for your evaluation, for each figure type one letter (the label for that cluster) or 0 (if that cluster should be left out from the model). The top left panel should turn green after you gave a label (make sure zoom/pan tools etc are turned off). Then move the figure aside without closing. OK to use same letter for multiple clusters, these will be merged.

Look at labels with `evsonganaly`, possibly adjust number of clusters etc

3. Apply model to new data

```
lv_resegment('batch.keep.lvrnd.ignore','evtaf')
```

```
lv_preparepsds('batch.keep.lvrnd.ignore','evtaf',5)
```

```
lv_labelpsds('~/.data/birds/wh09pk88/baseline/170217/smallmdl4','batch.keep.lvrnd.ignore')
```

Some parameters to play around with (have not systematically tested anything):

Number of basis syllables, number of training syllables

Number of PSDs per syllable (more is better but takes longer), probably can be reduced

Resolution of PSDs in frequency dimension can probably be reduced a lot (need to adapt tolerance window for cross-correlation accordingly)

Most obvious improvement missing: segmentation, local threshold seems necessary for some bids

Did not look into code that goes through and calculates the PSDs (got from Dave/Hamish), can maybe be sped up?