Procedure for extracting frames from .dv

Jeff Berry

March 1, 2011

1 Extracting the audio track

The audio track can be extracted from a .dv file using the following command:

```
ffmpeg -i filename.dv -vn -acodec copy filename.wav
```

This command will create a .wav file extracted from the .dv file.

2 Using Praat to create a label file

Once the sound file is extracted, Praat can be used to label the interesting parts of the file. To do this, open the .wav file in Praat as a Sound file. On the right, there will be an option called "Annotate -". Select "To TextGrid (silences)" as in figure 1.

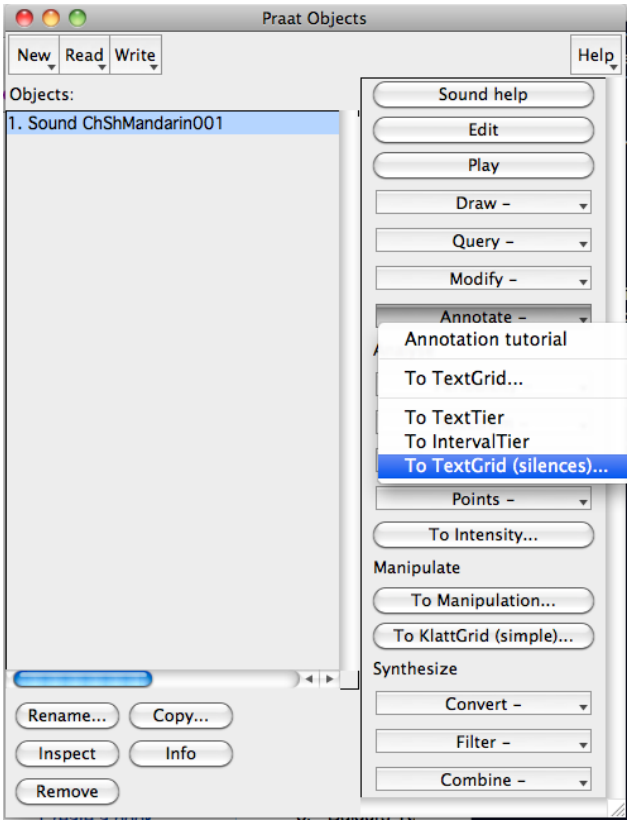

Figure 1: Select "To TextGrid (silences)"

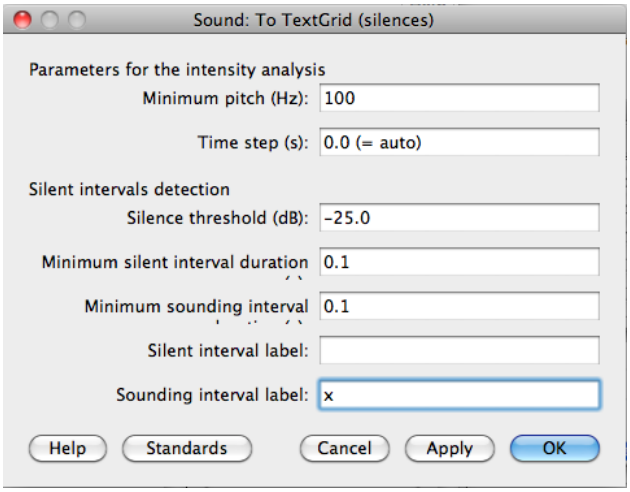

Figure 2: Change the values in the last 2 fields as shown

In the next window, change the field "Silent interval label" to be empty, and the field "Sounding interval label" to 'x' as in figure 2.

In the main Praat window, select both the Sound and the TextGrid, and press "Edit", which will open a new window where you will see the waveforms and the labels.

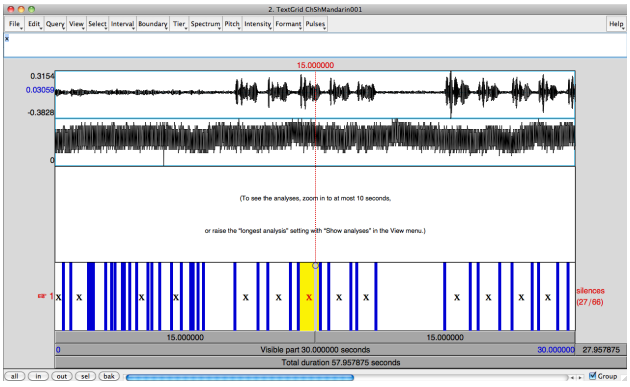

Figure 3: The edit window where labels will be fixed

Use the buttons in the bottom left corner of the edit window to zoom in and out. Check that all the label boundaries match actual stimuli. If labels need to be deleted, delete the text inside the interval, and delete the interval boundaries by selecting the boundary and clicking "Boundary" and then "Remove". You can also use the keyboard shortcut option+del. If a new interval needs to be created, select the portion of the waveform that should have an interval around it, and click "Interval" and then "Add interval on tier 1" from the drop-down menu, or use the keyboard shortcut command+1. This will create new boundaries. click inside the interval and add an 'x'.

Once all the intervals are correctly labeled, save the label file by clicking "File" and then "Write TextGrid to text file", or use the keyboard shortcut command+s. Select a filename and save. You can now close Praat.

3 Using ExtractFrames.py to segment the .dv file

`ExtractFrames.py` provides a simple interface for using the `.TextGrid` created with Praat to extract image files and `.wav` files from the `.dv`. When you run `python ExtractFrames.py`, a window will appear as in figure 4. Use the Open buttons to choose the `.dv` file and `.TextGrid` file. The Item List is an optional text

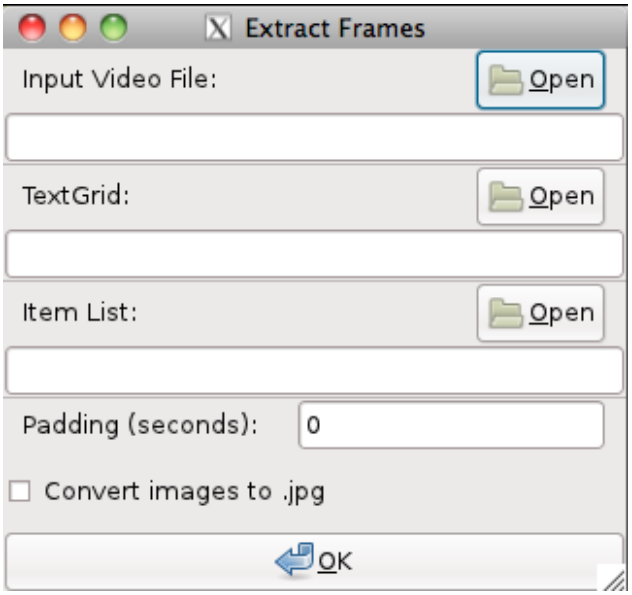

Figure 4: the ExtractFrames window.

file that contains the names of the outputfiles. Each line of the item list should contain a name corresponding to the interval in the `.TextGrid`. For example, if the `.TextGrid` file has 20 labeled intervals, there should be 20 lines in the item list, each specifying a unique name. A sample item list is shown here:

```
miaojian1
miaojian2
miaojian3
miaojian4
miaojian5
zhouchen1
zhouchen2
zhouchen3
zhouchen4
zhouchen5
loujiang1
loujiang2
loujiang3
loujiang4
loujiang5
liuchi1
liuchi2
liuchi3
liuchi4
liuchi5
```

The Padding text entry box allows the user to specify how much time to add to the edges of the labeled intervals in the `.TextGrid` file. Remember, 0.1 seconds = 100 ms. Using the method in section 2 above often places interval boundaries very close to vowels, which can cut off the beginnings of gestures in the onset. Rather than correcting the intervals by hand in Praat, the Padding box allows the user to add extra time to account for this.

Finally, the option to convert to `.jpg` allows the user to specify whether or not to create `.jpg` files. By default, `ExtractFrames.py` will extract `.png` images. If the convert to `.jpg` option is checked, both `.png` and `.jpg` files will be created.

Once all options and files are specified, and OK is clicked, the extraction process will begin, which can be monitored in the terminal window. A separate `.avi` file is created for each labeled interval, along with a corresponding `.wav` file located in a new WAV/ directory. A PNG/ directory is created as well, with `.png` images extracted from the `.avi` clips. If the convert option is selected, a JPG/ directory with `.jpg` images will be created. After all extraction has been done, and notification window will appear. This completes frame extraction.