



The Speech Solution

GENERAL OVERVIEW OF THE BABEL DEMONSTRATOR SYSTEM

RESPITE PROJECT

www.babeltech.com

www.infovox.se



Some General features

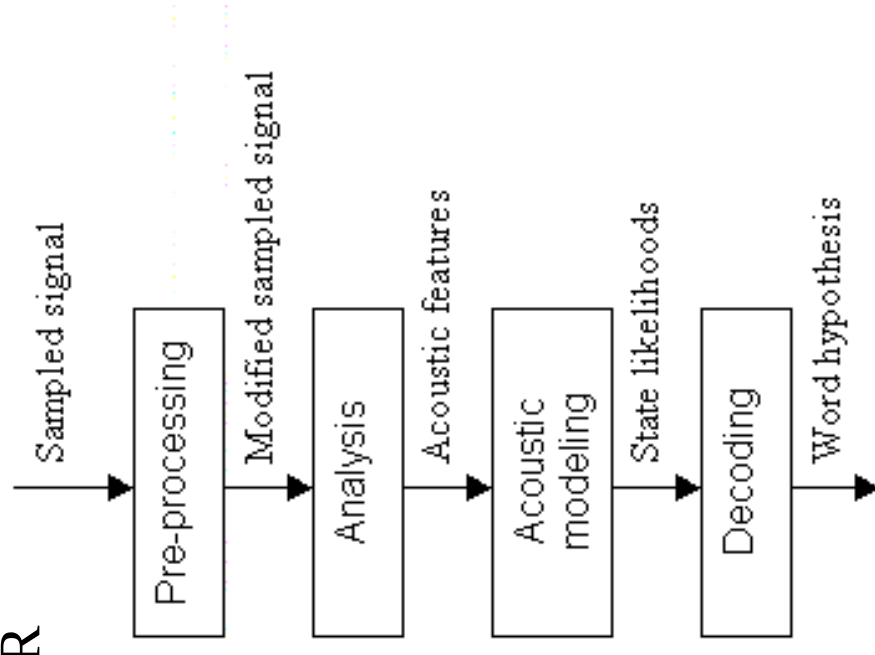
- Based on the wavesurfer program developped by KTH.
- Why ?
 - Platform independent – TCL–TK based programming
 - Plug–in based –> easy to extend
 - Free
- What ?
 - Plug–in for the integration of the work of all the partners in a single program



Babel demonstrator

The demo interface provides access to the ASR process at different anchor points such as :

- sampled speech signal,
- acoustic features,
- state likelihoods,
- recognized sentences.





Customisable interface

- Each block can be processed independently by calling a user-defined external program,
- The only constraint is the compatibility with the input/output data format.

Modifying the Script commands

Block	Name	Command line
Preprocessing	pre-processing	
Analysis	feature extraction	c:/ris/respite/respite2strut
Acoustic modeling	acoustic modeling	c:/ris/respite/respite2strut
Decoding	decoding	c:/ris/respite/respite2strut

OK Cancel



Data display

- Display is fully handled by the ASR interface by specific plug-ins for the 3 different data types:
 - samples, acoustic features, probabilities/likelihoods.*
 - Automatic time-alignment of the different data streams on the speech signal.
 - Allow display of block's internal data (multi-stream format).
 - Dynamic internal data management : Automatic update of data display when necessary
- The definition of the data format must include all the information required by these constraints.



Data format

- Frame based
- Each frame can contain several synchronized data streams of any of the pre-defined data-types
- Binary header:

Sample rate in Hz
Frame length in ms
Frame shift in ms
Number of streams
Name stream #1
Type stream #1
Frame size stream #1
Name stream #2
Type stream #2
Frame size stream #2

unsigned long
float
float
unsigned long
32 char string
unsigned long
unsigned long
32 char string
unsigned long
unsigned long

- The stream names are used for identifying displayed windowpane.



Data format (con't)

- The stream type is one of the following:

- 1 – samples (PCM16) signed short
- 2 – features float
- 3 – probabilities float

- The actual data are then formatted the following way, for each frame:

```
Time index in milliseconds for current frame – float
Data stream #1
Data stream #2
...
```

- The time index is used for time alignment of different data streams. A time index of –1 means the end of the current utterance.



Data format (con't)

- The word hypothesis are written in Timit format:

`start_time(samples) end_time(samples) word_hyp`

for instance:

```
0 8480 Sil
8480 13280 one
13280 20480 eight
20480 34880 six
```



Status point

The interface consists in 4 plug-ins:

- samples.plugin
- features.plugin
- recognize.plugin
- probabilities.plugin

Those plug-ins are compatible with WAVESURFER v1.2 and higher.

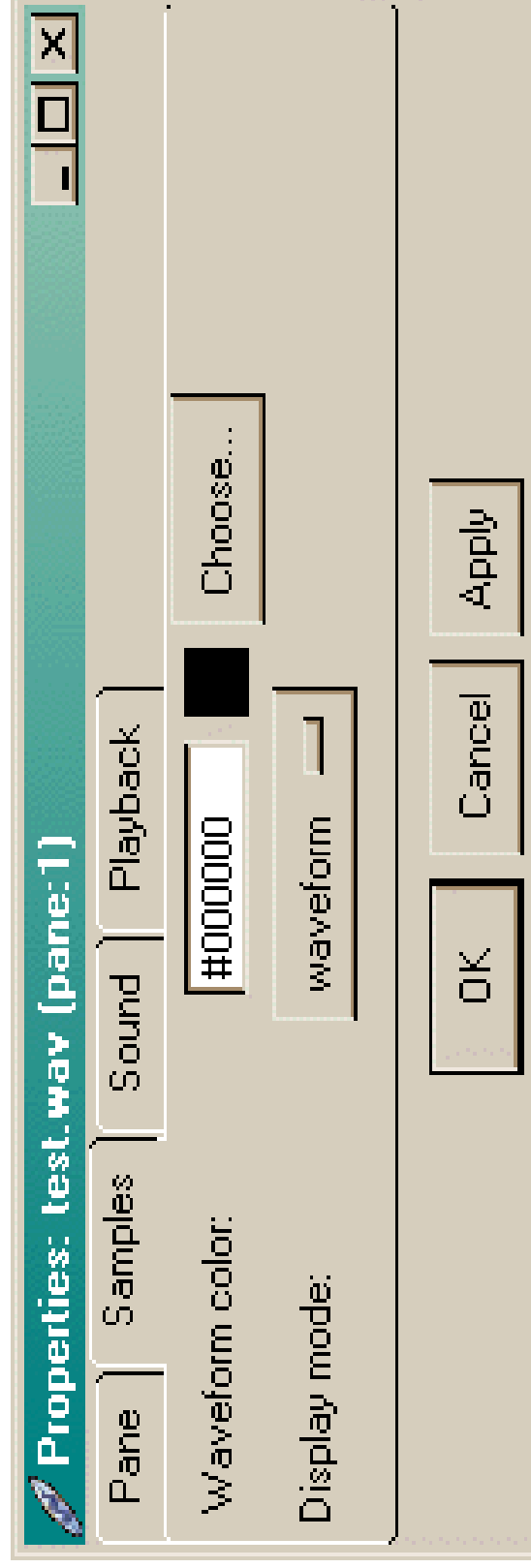
Under Windows, copy these files in
%HOME%\.wavesurfer/1.3/plugins

Under Linux, copy these files in
\$HOME/.wavesurfer/1.3/plugins



Display of samples

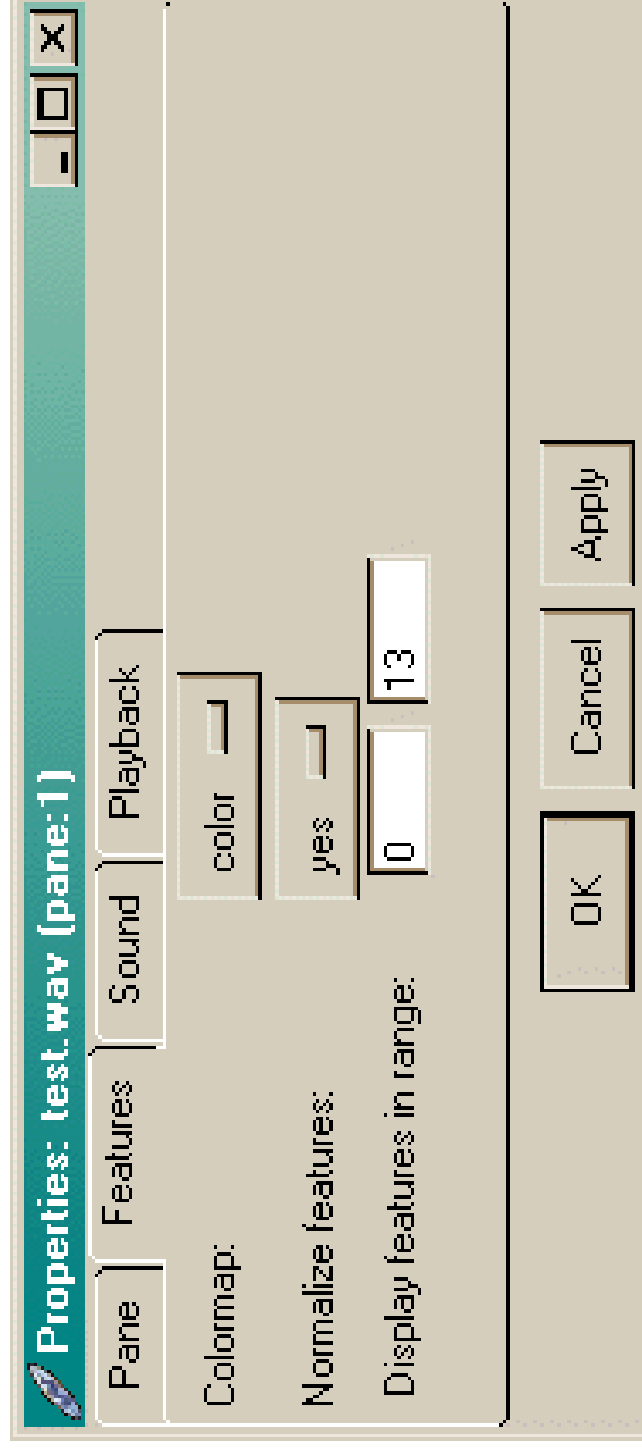
The display of the samples is very similar to the one of WAVESURFER. You can either plot the waveform or the spectrogram by selecting the menu <Properties> of the samples window pane.





Display of features

The features are displayed as pseudo-spectrograms. By default, the features values are normalized, i.e. each feature parameter is normalized over time. This can be modified by selecting the <Properties> menu of the features window pane. You can also select the range of feature parameters you want to display.





Display of probabilities

The probabilities/likelihoods can either be plotted or displayed as a pseudo-spectrogram. Again some options are available from the <Properties> menu. You can, for instance, specify the name of a file containing the symbols related to each probability.





Status point

- Integration of other partners work :
 - FPMS : integration of the multi-band approach.
- Display of the frequency bands features and probabilities.
- ICP : ?
- Sheffield University : Missing data ?



Demonstration ...

- KTH agreed to integrate the demonstration package into the distribution of the wavesurfer program.
 - <http://www.speech.kth.se/wavesurfer/>
- Link to the RESPITE web page ?
- Publically available for research purpose