WinSnoori version 1.34

A Speech research tool

1 Introduction

WinSnoori is a speech research software aimed at assisting researchers in the fields of speech recognition, phonetics, perception and signal processing. The tools provided by WinSnoori were designed with a principal concern of making their use on speech signals very easy. This explains and justifies the overall organization of the software and the set of default parameters retained, such as those needed for the calculation of spectrograms.

WinSnoori can run on a Windows XP multimedia machine.

2 What WinSnoori can do?

WinSnoori is able to edit speech signals, calculate and display spectrograms, to which must be added phonetic and orthographic labeling of speech signals, the possibility of dynamic visualization of the results of diverse spectral analyses, automatic exploration of speech corpora, formant monitoring, fundamental frequency extraction, speech rate modification and fundamental frequency contour modification. Most of the results of WinSnoori, are put in text form and therefore can be exploited with other software.

WinSnoori provides a complete graphical interface for generating stimuli for the Klatt synthesizer (edition of all the Klatt parameters, automatic formant tracking, copy synthesis).

2.1 The speech editor

The editor is designed to enable very fine modification of signals with the help of the spectrogram. The spectrogram is indeed able to highlight spectral discontinuities and so facilitates their correction and removal. To avoid the limitations relating to memory capacity and file size, WinSnoori works on 8 second window intervals. In addition to traditional editing and speech restitution, WinSnoori allows both signal filtering (FIR and OLA filtering wich offers the possibility of drawing filters directly onto the spectrogram) and signal attenuation. These functions are especially relevant to stimuli preparation for perception experiments. WinSnoori is capable of handling a very wide variety of speech and label files.

2.2 Spectrograms

WinSnoori calculates and displays wide and narrow band spectrograms, cepstrally smoothed spectrograms and spectrograms obtained through linear prediction. In addition, it is now possible to compute reassigned spectrograms that improve the distribution of energy. The signal window size, just as the minimum and the maximum spectrogram energy levels, can be adjusted interactively. Fine edition of speech signals is applied directly on spectrograms (see Fig. 1 where Snoori is written by filtering noise).

2.3 Spectral analyses

The user can open a window to display spectral slices such as Fourier transforms, simple or selective linear predictions, cepstrally smoothed slices and also a variety of critical bands analyzes (Mel or Bark filters followed by liftering of linear prediction to enable Mel cepstrally or PLP smoothed spectra to be calculated). It is possible to adjust all spectral slice related parameters. Slices are themselves calculated by following the movement of the mouse on the spectrogram. The user can also open a text file into which to recuperate numerical results. In addition, WinSnoori can display the peaks of cepstrally smoothed spectrums, or the roots of linear prediction polynomials on the spectrogram and thereby highlight formants.

2.4 Labeling

WinSnoori provides the user with the possibility of both phonetic and orthographic labeling. The user uses the mouse and a menu containing the list of the phonemes of the language under investigation to construct phonetic labels for the speech. These labels appear at bottom of the spectrogram. The list of phonemes can be modified at will, thus making WinSnoori multilingual with the possibility of application to any set of phonemes or to any type of marks (prosodic, acoustic,. . .), irrespective of the language considered. Furthermore, WinSnoori offers complementary editing and file management tools such as displacing, deleting, searching phonetic and orthographic labels, as well as reading and saving label files.

2.5 Prosody

Fundamental frequency extraction is based on an algorithm which uses a frequency comb and is completed by a pitch correction algorithm based on dynamic programming. In order to take into account the sex of the speaker in the variation of the fundamental, WinSnoori offers the possibility of specifying three types of speakers, namely, female, male and speakers with unknown sex.

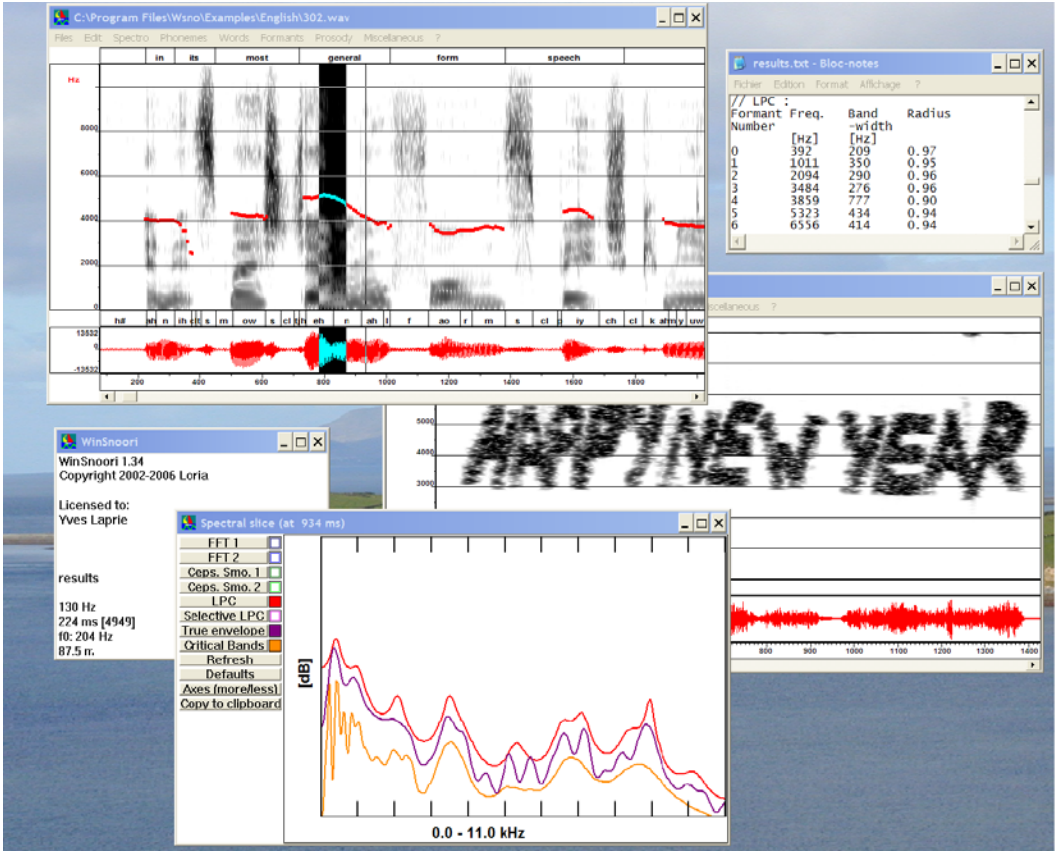

Figure 1: Main window of WinSnoori with speech signal. large band spectrogram, phonetic and orthographic labelings. ”Happy new year” is written by filtering white noise by means of an overlap and add method using filters drawn by hand.

WinSnoori can be used to study prosody from a descriptive point of view but the most interesting point is probably the possibility of changing the speech rate and the fundamental frequency contour. Changing speech rate especially slowing down) is an efficient way to improve oral comprehension of a foreign language and modify the relative importance of the speech rate compared with that of the fundamental frequency. Changing the fundamental frequency can be carried out either on the whole utterance by applying a multiplicative factor or on specific regions by editing the F0 contour and then resynthesize a new signal.

2.6 Graphical interface of the Klatt synthesizer

The Klatt synthesizer is an invaluable tool for studying acoustics and perception of speech. WinSnoori, therefore, provides users with a powerful graphical interface which includes a complete toolbox for creating and editing files of parameters as well as tools for copy synthesis for generating stimuli close to natural speech signals. All the parameter trajectories can be modified point by point, moved and smoothed. Formant trajectories can be registered automatically at the closest roots of the linear prediction polynomials or the peaks of cepstrally smoothed spectra. Synthetic and original spectrograms or spectral slices can be compared easily and formant trajectories can be displayed in the F1-F2 plane to facilitate the assessment of the amplitude of coarticulation.

WinSnoori now offers tools for copy synthesis, i.e. automatic formant tracking and automatic adjustment of formant amplitudes. This enables the creation of stimuli very close to natural speech.

The synthesizer is independent of WinSnoori and is the GPL version developed by Jon Iles and Nick Ing-Simmons (the Windows version is available at http://www.loria.fr/equipes/parole/Html/klatt.html).

2.7 Using speech corpora

WinSnoori recognizes the more commonly used standards in phonetic and orthographic labeling (NIST, SAM...). WinSnoori is thus able to read the majority of speech corpora in use. To work on a corpus, the user only needs to indicate the location of the speech file and the corresponding label file. This avoids the need for complete specification of paths at each file loading session. WinSnoori provides a mechanism for the automatic exploration of phonetically and orthographically labeled corpora to allow the construction of speech files and the corresponding label files associated with all occurrences of a given sequence of phonemes or text.

3 How to get WinSnoori?

WinSnoori can be freely downloaded at http://www.winsnoori.fr.