Phonexia Language Identification from spoken speech

User's guide for GUI application

This documentation applies to BS-API revision 331 and GUI application version 4.2.1

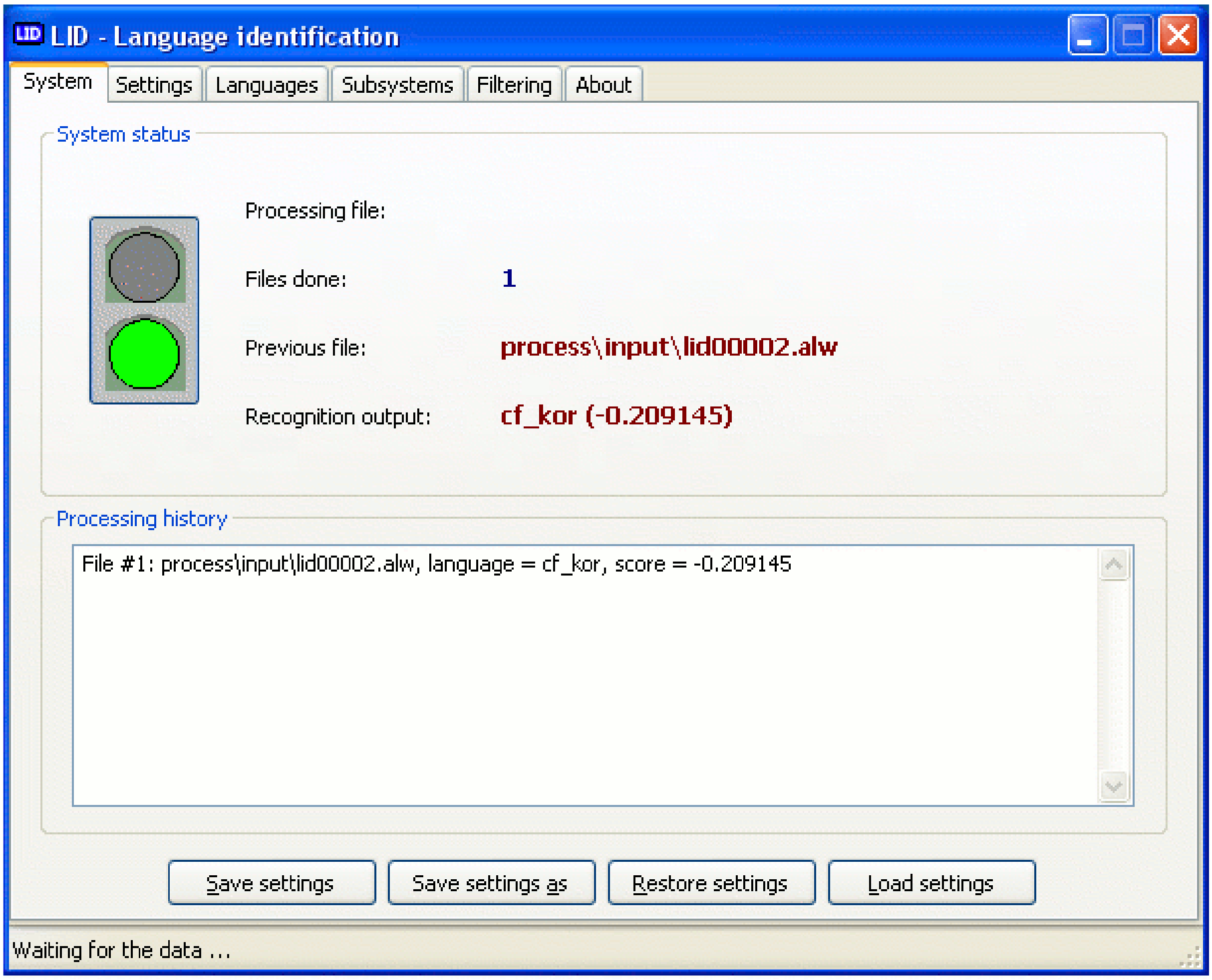
The graphical user interface for Phonexia Language Identification system allows to easily use Phonexia's technology with no programming. To use the application, follow these three easy steps:

1. Unpack the archive
2. Copy a license file (obtained during registration and download) to the program's root directory (next to lid.exe)
3. Run lid.exe

If everything went correctly – the license file was found, there is an internet connection and the license has not expired yet, you should get a multi-page application window. In opposite situation, a dialog window showing a contact to our technical support is displayed.

The initial page has name "System". The page shows lights. By clicking on the lights, the processing of files can be turned on or turned off. Red light means "the system is stopped". Green light means "the system is running ant it is waiting for incoming data or it is processing data." The page also contains information about number of processed files, about the last identified file, its language, and about currently processed file. Outputs for few last processed files are kept in the "Processing history" box.

The application has four buttons on the bottom for managing configurations. The buttons can be used anytime but a new configuration applies after restart of the processing (by lights). "Save settings" saves the configuration after a field is changed by user. "Save settings as" saves the settings to a new configuration file. "Restore settings" restores all edit boxes to the state before the user start editing them. And "Load settings" loads the configuration from a chosen configuration file.

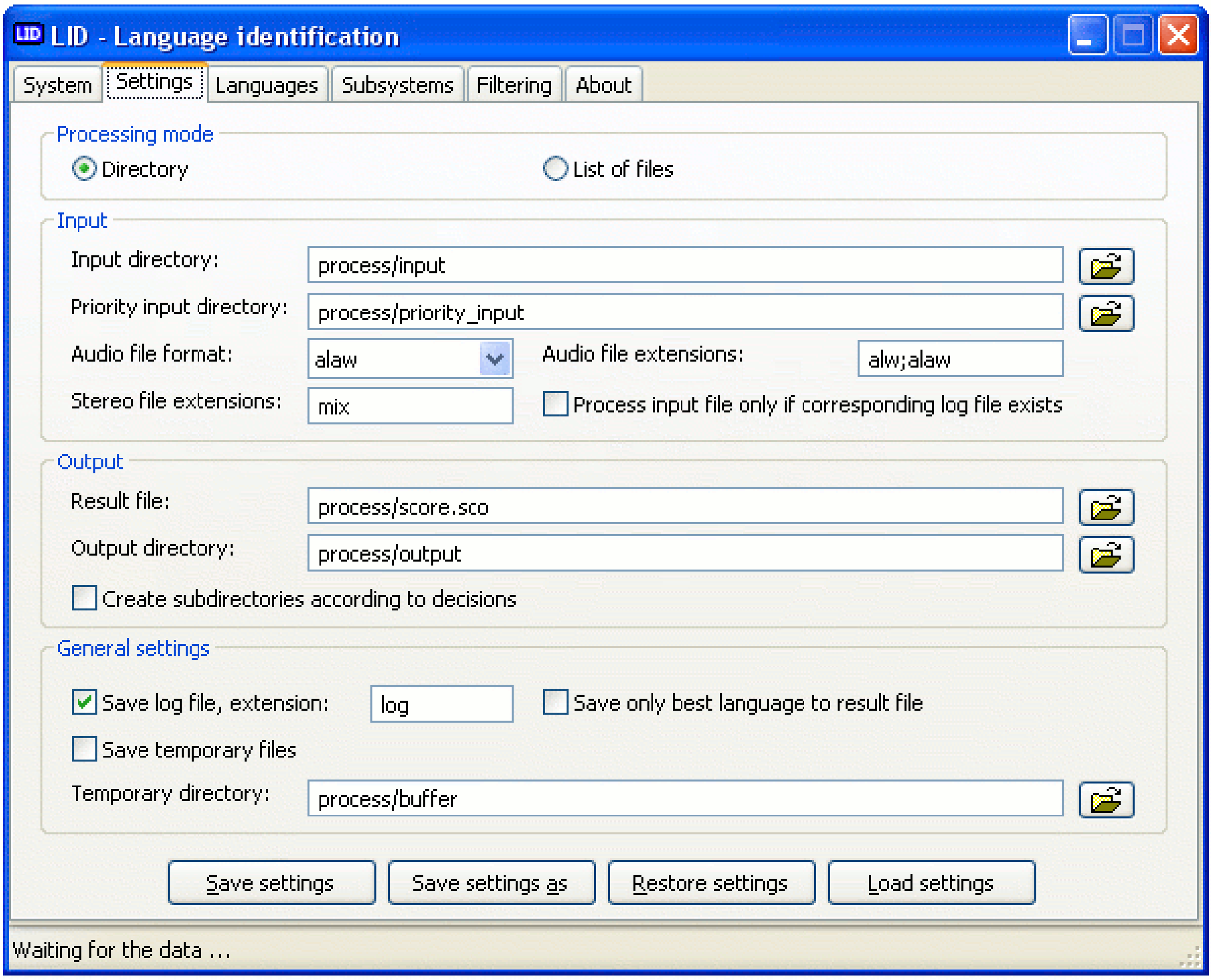


The application can run in two modes:

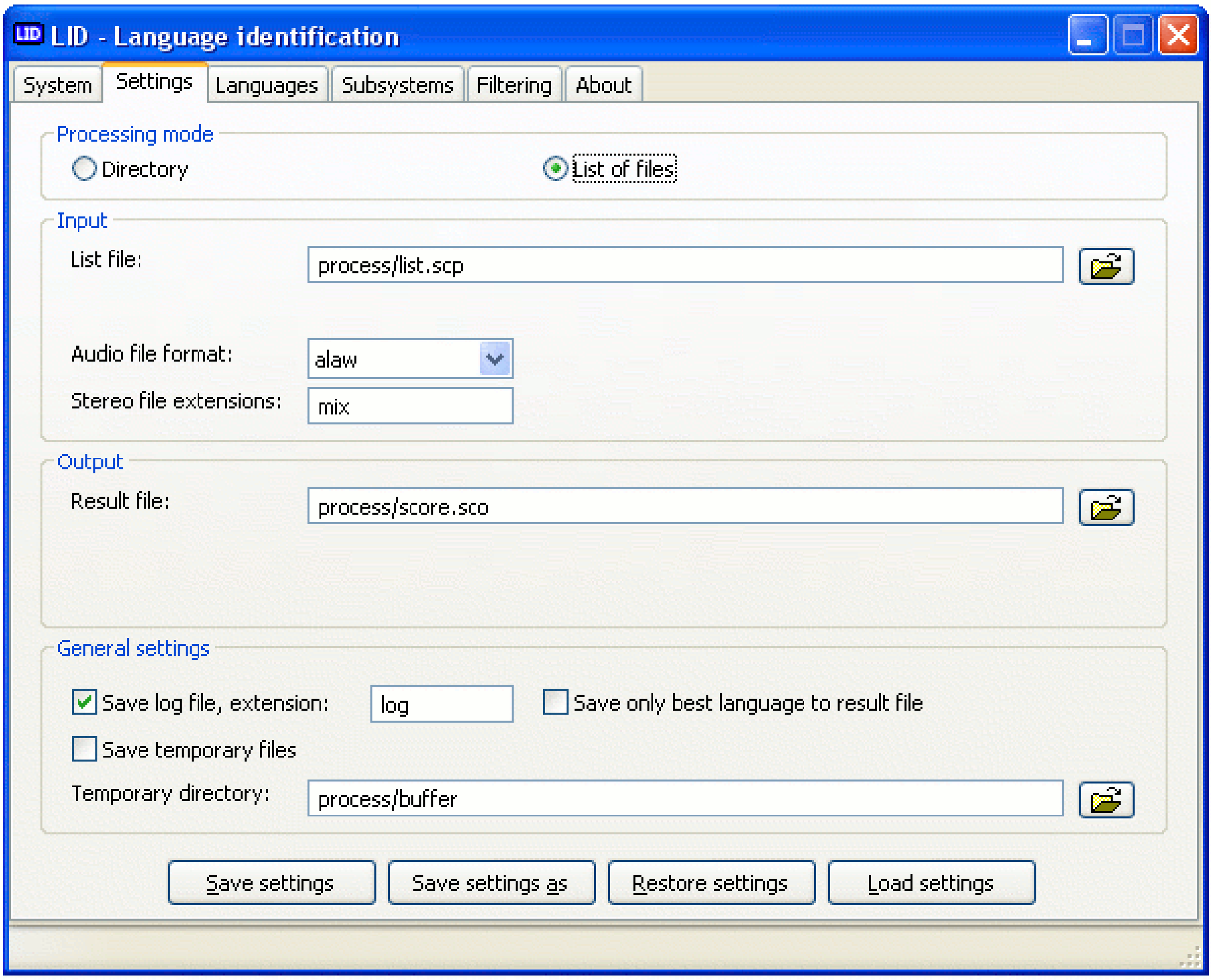
1. In the first mode, it waits for incoming files in a directory and when a new file appears, it is processed.
2. The second mode allows to process files according to a file list.

Processing files from a directory

The field "Processing mode" on page "Settings" must be set to "Directory". In this case, an input directory must be specified. A priority input directory can be specified too. If files are waiting in the input directory and a user would like to quickly process some files, they can be copied to the priority directory. The processed files are moved to an output directory. Optionally, some subdirectories according to the winning languages can be automatically created and the speech files can be sorted to these subdirectories. The output scores/decision can be also saved to a result file. The program can wait for a log file to exist before processing of a speech file. This ensures a better collaboration with another software or hardware. After processing, the log file is moved to the output directory too.
The application uses a locking mechanism (opening of input files for an exclusive access and lock files), therefore it is safe to run more instances of the program on more cores/CPUs for speeding-up the processing.



Processing of files according to a file list
The field “Processing mode” on page “Settings” must be set to “List of files”. In this case, the output can be saved to a result file only.

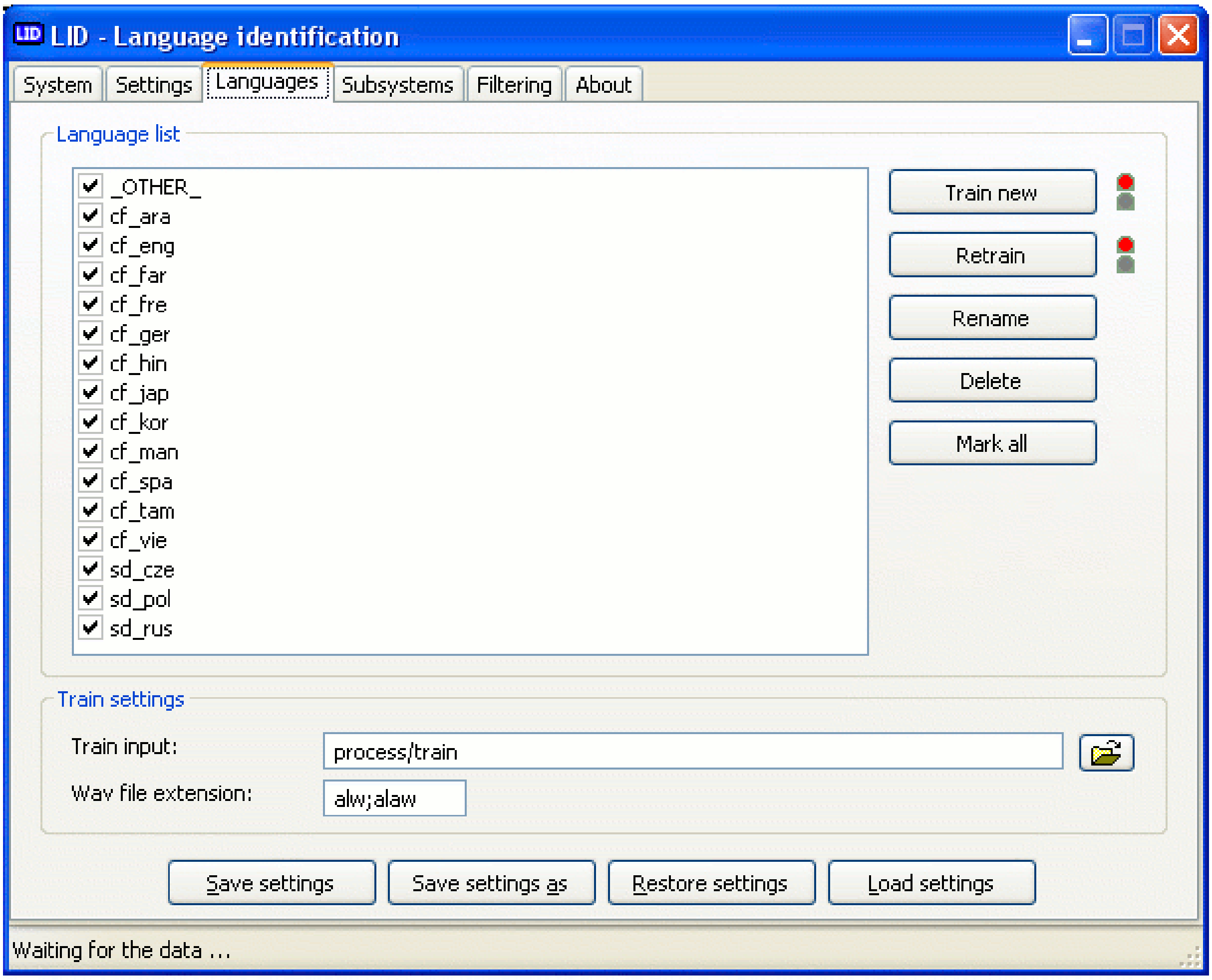


Input file format
Input file format can be specified in the “Input” section. Currently, wav files and raw files with linear coding (16 bits / 8 bits), A-law or Mu-law are supported. The program supports stereo files. In this case, recordings from both channels are concatenated. Format is detected automatically from the file header if the input file is wav, no matter what is set in the “Audio file format” box.

General settings
It is possible to save a log file about processing. This can be enabled in section “General settings”, field “Save log file”. The result file can contain scores for all languages or just the winning language. This is specified by field “Save only best language to result file”. If some files are processed many times (for example with different language model sets), the processing can be speeded-up by saving temporary files. In this case, the application creates some additional temporary files during the processing. These temporary files are stored in a directory specified in “Temporary directory”.

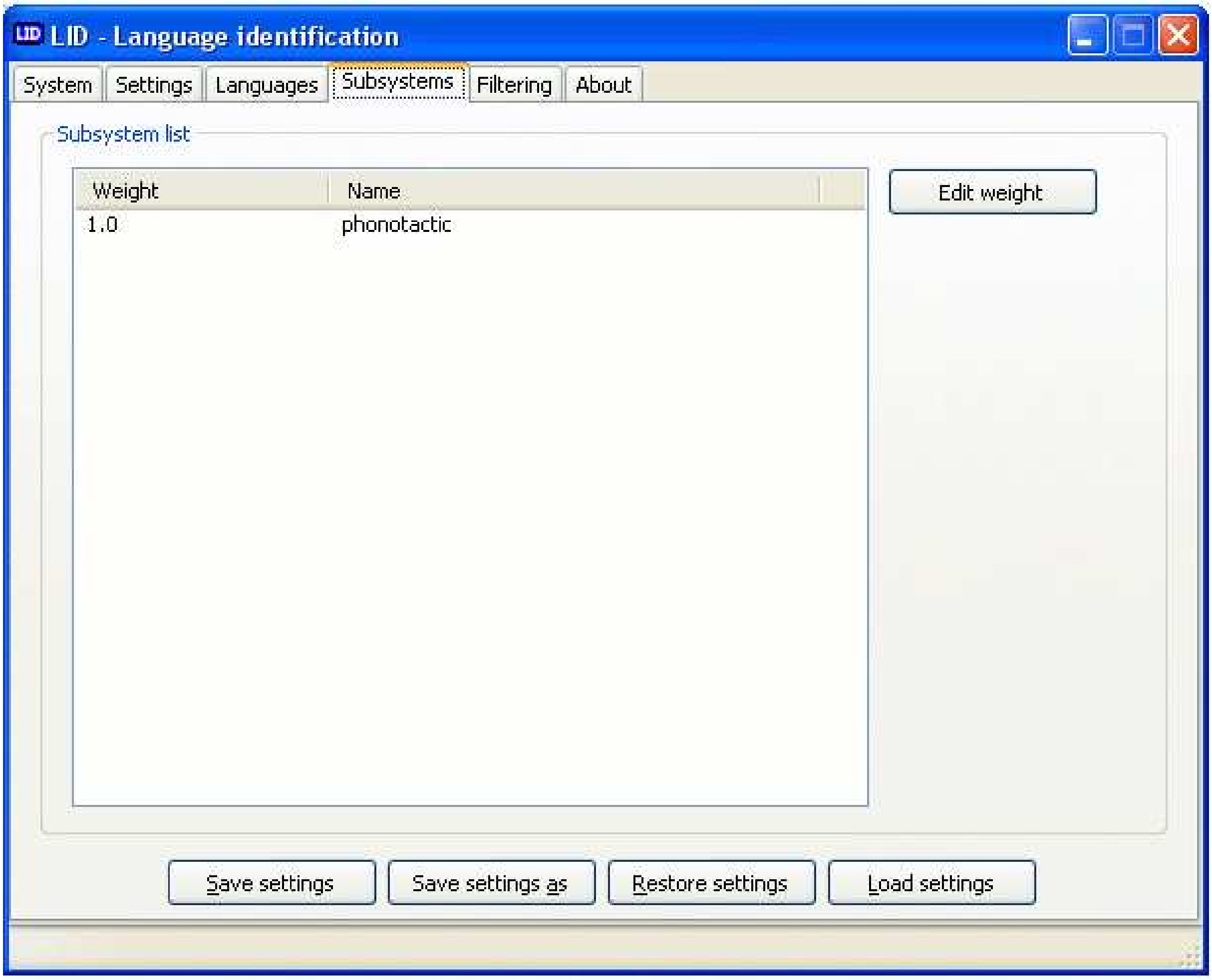
Languages

The application is distributed with a set of several languages. Very often, it is not necessary to identify all of them, for example if we have a closed set of possible languages or if there is also another source of information about the language. Languages that are used during processing are specified on page “Languages”. This page can be also used for language set maintenance. A language can be renamed, deleted, a new language can be trained on user data, or a language can be retrained.



Subsystems

More subsystems can be used for language identification, for example a phonotactic subsystem or an acoustic subsystem. The phonotactic subsystem models order of phonemes (sounds/ letters) in speech. The acoustic subsystem models how the speech sounds. The phonotactic system is more robust to non-native speakers. And the acoustic subsystem works better with different dialects. The page “Subsystems” allows to adjust a weight of each subsystem. Currently only the phonotactic subsystem is supported.



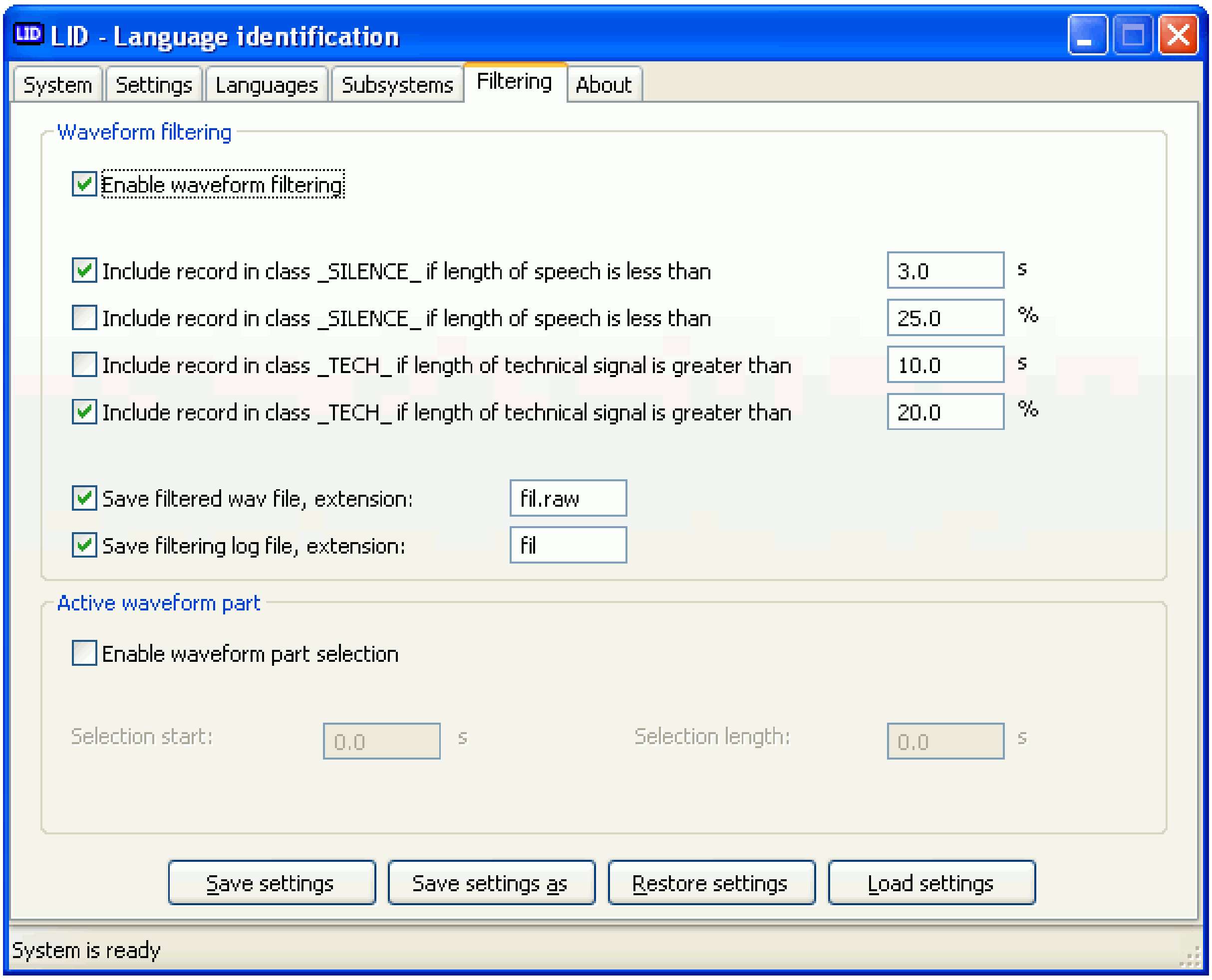
Filtering

This page is for the users that are not able to ensure quality of the input signal (security domain) or for users that would like to speed up the processing. Field "Enable waveform filtering" enables a filter that:

- cuts off technical signals (tones, fax, modem)
- removes long segments of silence
- filters out pulse noise

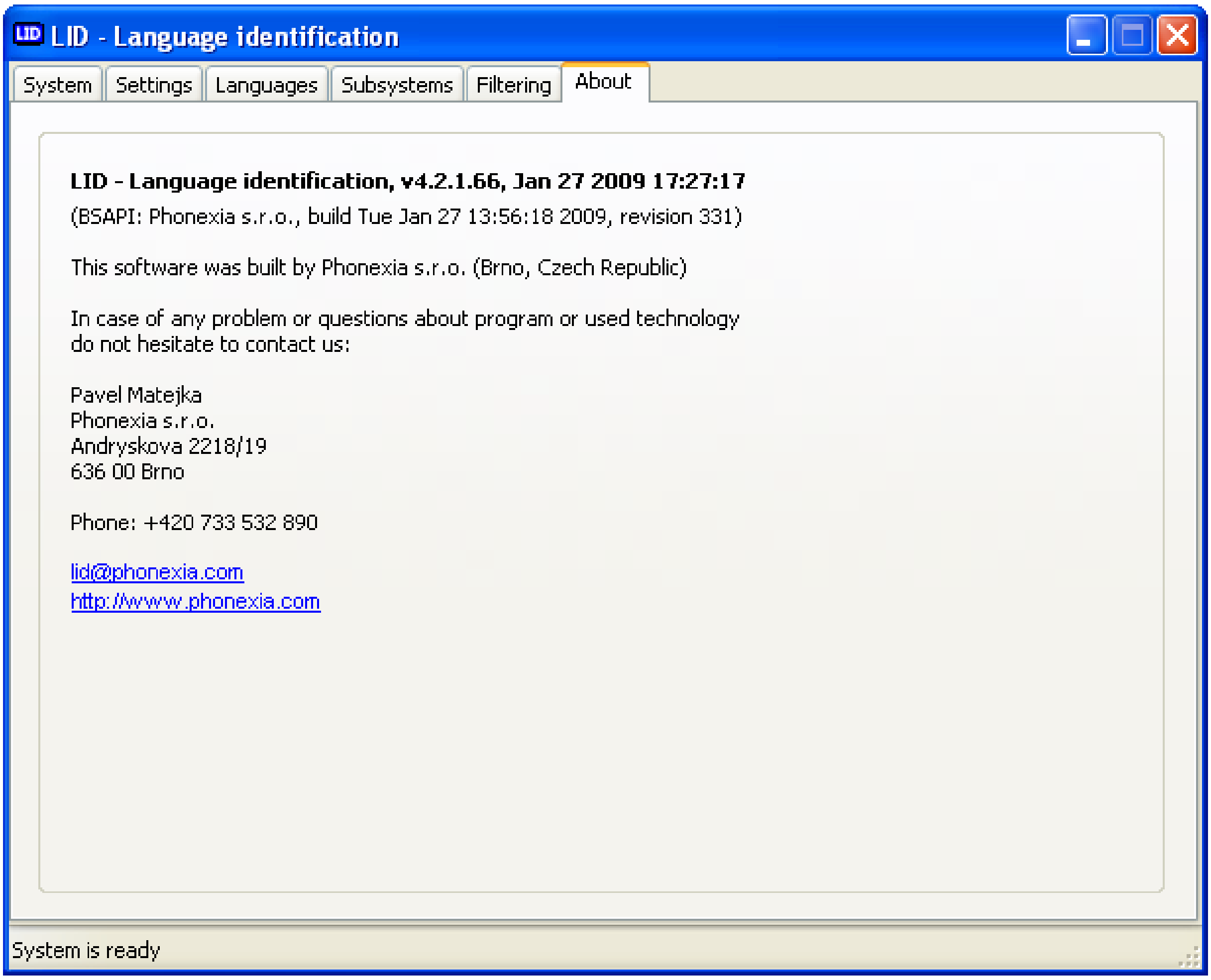
According to statistics from the filter, a decision whether to process the record or not can be done. Records passed through the filter can be saved for a later use. The coding is always linear 16 bits.

An active waveform part can be also specified for input records. This can significantly speed up the processing if the records are very long. A sophisticated algorithm is used for selection, so if there is no speech in the specified area, the area is adjusted.



About

Page “About” gives you a contact to our technical support. Please use this contact in case of any questions. Please do not forget to add GUI version and BSAPI revision to your question.


LID - Language identification
System
Settings
Languages
Subsystems
Filtering
About
LID - Language identification, v4.2.1.66, Jan 27 2009 17:27:17
(BSAPI: Phonexia s.r.o., build Tue Jan 27 13:56:18 2009, revision 331)
This software was built by Phonexia s.r.o. (Brno, Czech Republic)
In case of any problem or questions about program or used technology do not hesitate to contact us:
Pavel Matejka
Phonexia s.r.o.
Andryskova 2218/19
636 00 Brno
Phone: +420 733 532 890
lid@phonexia.com
http://www.phonexia.com
System is ready