Department of Signal Processing and Acoustics

Aalto Aparat

Manual v2.0

_____

Hilla Pohjalainen, Manu Airaksinen,
Matti Airas and Paavo Alku
2015

A? **Aalto University**
School of Electrical
Engineering

Contents

1   About Aalto Aparat

Aalto Aparat (later Aparat) is a voice inverse filtering and parametrisation software for estimating and analysing the glottal flow. MATLAB-based Aparat is available under an open-source licence and can be used in fundamental research of speech, fonetics and study of occupational voice. The programme provides a user-friendly graphical interface from where the results can be easily saved and re-used in other projects.

The original Aparat was created and designed by Matti Airas [1]. The first version of the software was published in 2005 under the name TKK Aparat, referring to the former Helsinki University of Technology (which was merged into Aalto University in 2010). The updated version 2.0 includes a new glottal inverse filtering (GIF) method and several minor new features, such as more versatile saving options.

1.1   Glottal inverse filtering methods

There are two different GIF methods available in Aparat: quasi closed phase analysis (QCP) [2] glottal inverse filtering and iterative adaptive inverse filtering (IAIF) [3]. The methods share three general parameters: the number of formants, lip radiation coefficient, and low-frequency noise cutoff (in Hz).

QCP is the default GIF method in Aparat since the version 2.0 update. QCP is based on the principles of closed phase (CP) analysis [4], that is the estimation of the vocal tract during the glottal closed phase. In contrast with the basic CP method, the estimation is applied with the autocorrelation criterion using a long analysis window. This is enabled by weighted linear prediction (WLP) [5] with an attenuated main excitation (AME) weight function [6].

The algorithm consists of four stages. First, glottal closure instants (GCI) are detected from the speech signal by using electroglottography or specific GCI detection algorithms. To perform this, Aparat uses an automatic algorithm called SEDREAMS [7]. After that, the AME function is constructed based on the estimated GCIs (Fig. 1). From the Aparat users' perspective, this is the most interactive part of the computation as there are three AME function parameters that the user can manipulate on Aparat control panel: position quotient (PQ), duration quotient (DQ), and the number of linear ramp samples (Nramp). The AME function is utilised when creating the WLP-based vocal tract model, and the default Aparat parameter values of PQ, DQ and Nramp have been selected in order to minimise the estimation error on a large data set of synthetic sustained vowels. In most cases, the modification of these parameters can be viewed as fine tuning, and commendable results can be obtained even without modifying them at all.
Finally, the glottal flow derivative estimate is obtained by inverse filtering the input speech signal with the vocal tract model.

IAIF is familiar to Aparat users since the first version of the software, and it was the default GIF method until QCP was implemented. While QCP has been shown
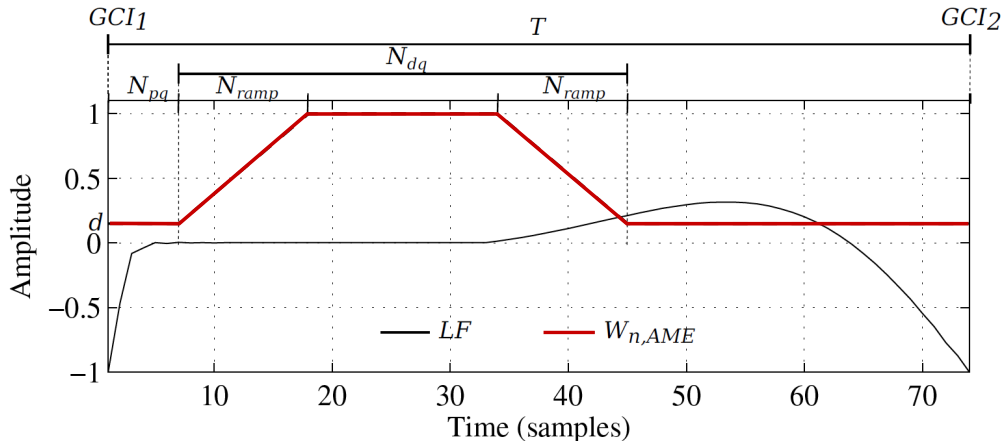
Figure 1: A glottal flow derivative waveform (LF) and the AME weight function ($W_{n,AME}$) between two consecutive GCIs. Glottal flow derivative was computed from the LF model [8]. The AME function was computed with PQ = 0.1, DQ = 0.5, and RQ = 10 [2].

to improve the quality of GIF when compared to results obtained with IAIF, it is not as robust as IAIF when reliable GCI estimates are unavailable: for example for signals with low SNR.

In IAIF method the glottal flow and its contribution to the speech spectrum is estimated with an iterative structure, which is repeated twice during the computation. The final result (as well as the first estimate) is obtained by cancelling the effects of the vocal tract and lip radiation from the original speech signal by inverse filtering.

The IAIF method presented in most scientific articles utilizes linear predictive analysis (LPC) to estimate the vocal tract. However, in Aparat the default autoregressive model in use is a discrete all-pole (DAP) model [9], as LPC and minimum variance distortionless response (MVDR) model [10] can be found on advanced settings panel.

Former GIF methods in Aparat, removed from an updated version 2.0:
The two deleted methods are direct inverse filtering (DIF) method (simpler version of an IAIF method) and experimental inverse filtering method. The DIF method is basically a simple version of IAIF method, as the DIF method follows the same filtering process as IAIF but finishes after the first integration step [1]. The experimental method was never fully finished and it did not run properly.

2   Installing Aparat

Aparat is a Matlab-based programme that can be run with or without MATLAB software. The installation instructions for both cases are explained.

2.1   When having MATLAB licence

For MATLAB users installing Aparat is really straightforward: MATLAB users' installation package is selected and uncompressed to a suitable location (for example, under the users personal MATLAB working directory). Aparat is opened by invoking its main function "aparat.m".

Aparat uses Matsig, an object-oriented signal processing library for MATLAB. Previous releases of TKK Aparat required Matsig to be separately installed, but the current package includes also Matsig.

2.2   Without MATLAB

Installation and the usage of the stand-alone version of Aparat requires the Matlab Component Runtime to be downloaded and installed. It is, however, available together with Aparat files on the website of the author.

The Matlab Component Runtime can be installed by double-clicking the file and following the instructions on the screen.

Currently, no installation program exists for Aalto Aparat. Instead, it can be installed by following the steps below:

1. Load the right installation package according to your operating system.
2. Create a folder for the programme
3. Uncompress all files from the installation package into that folder.
4. (Optional) Create a shortcut to Aparat.exe on the desktop by right-clicking the desktop, selecting New -> Shortcut, and then selecting Aparat.exe.

Now, Aalto Aparat may be run by double-clicking the Aparat.exe file.

3   Work with Aparat

3.1   File types

Aparat processes .wav files and save the processed data to .wav, .mat, and .tab files. The .wav files in process may be in stereo or mono, and the optimal length of the sample is from 0.5 to 10 seconds.

The obtained waveforms can be saved to .wav or .mat files. The saving process is presented in detail in the section Saving data.

3.2   The main windows and their features

Aalto Aparat has two main windows: control window on the left and signal view window on the right (Fig. 2) . They both appear automatically when opening the programme.
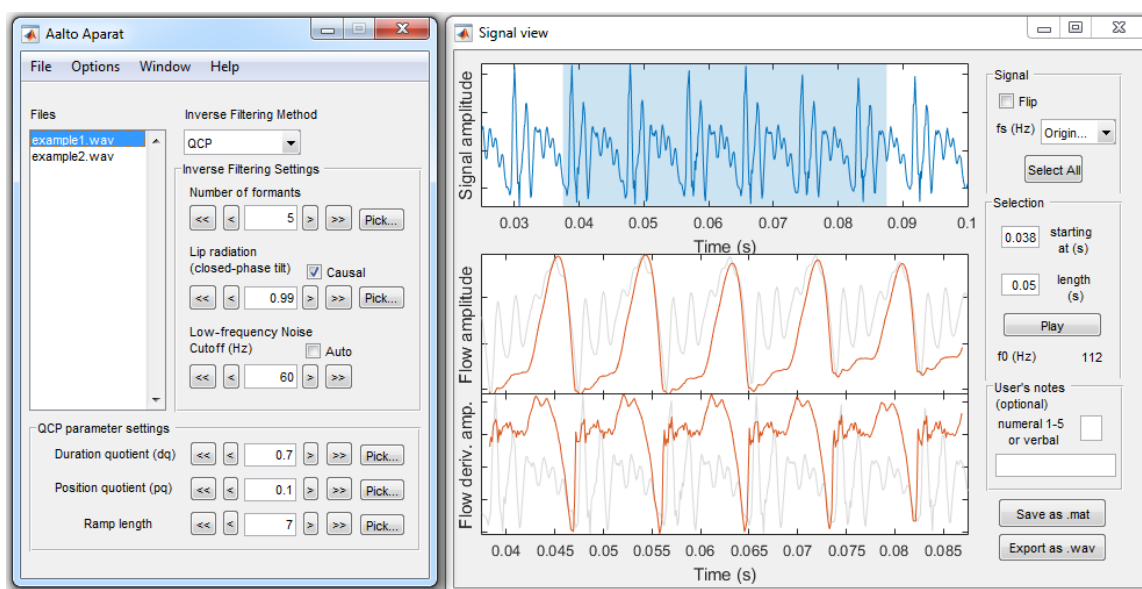

Figure 2: Aalto Aparat main windows: Control window on the left and signal view on the right

In Aparat control window there is a list of all the .wav files that exist in the same folder as the programme (Fig. 3). If some of the files also has a corresponding .mat file, there is an asterisk (*) in the beginning of the file name. More about files and saving: Saving data.

Despite different operating systems Aparat should always look the same, with slightly varying window frames.

Aparat can be closed by clicking the close icon on the top right corner of the control window or via File menu -> Exit.
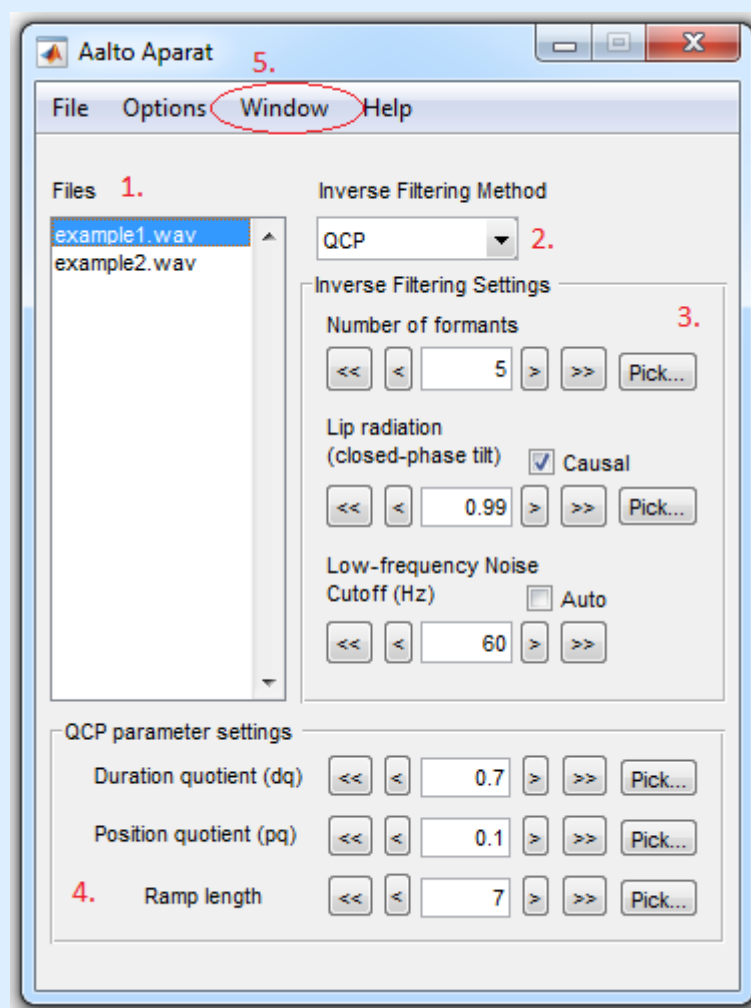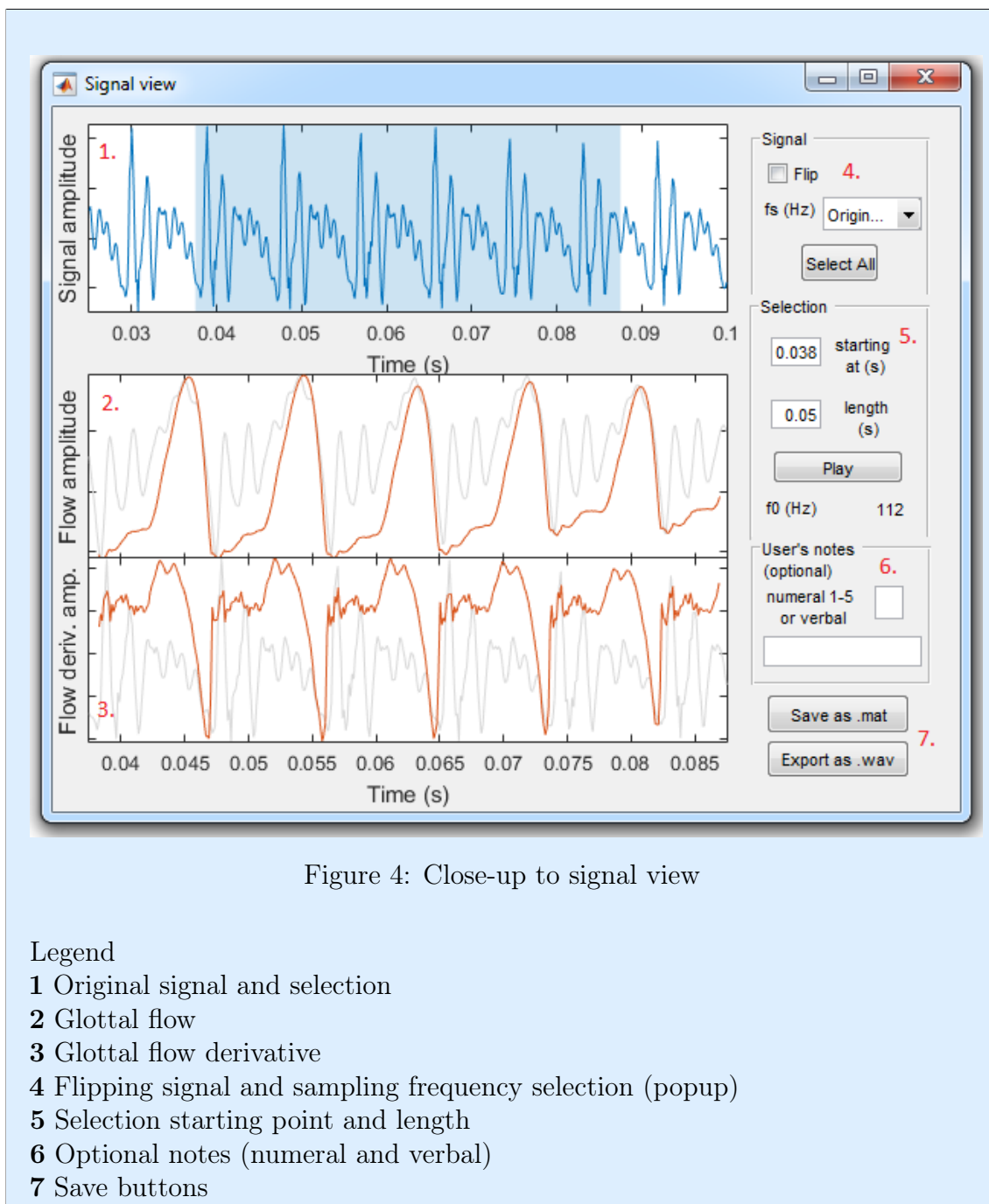
Figure 3: Close-up to control window

Legend
1 List of .wav files in the programme folder
2 Glottal inverse filtering (GIF) method popup-selection
3 General filtering settings for all GIF methods
4 Method dependent parameter settings
5 Special windows

Figure 4: Close-up to signal view

Legend
1 Original signal and selection
2 Glottal flow
3 Glottal flow derivative
4 Flipping signal and sampling frequency selection (popup)
5 Selection starting point and length
6 Optional notes (numeral and verbal)
7 Save buttons

3.3 Parameter settings

The user is able to manipulate a wide selection of parameters in updated Aparat. Not all filtering methods use exactly the same parameters but the logic is the same behind all the similar handles.

The handles of the following three general parameters are always visible on the control window: the number of formants, lip radiation coefficient, and low-frequency noise cutoff (in Hz) (Fig. 5). Additionally, there are QCP parameter and IAIF parameter panels under the filtering setting panel: they will become visible when corresponding filtering method is selected (Fig. 6). The QCP method has three and IAIF two extra parametes. All the parameters and their influence on the filtering process can be seen on a Table 1.

Table 1: The GIF parameter settings on the Aparat control window

| Parameter	Effect	Valid values	Other
Number of formants	Defines the order of the vocal tract filter	2 - 30	
Lip radiation effect	Affect on the integration coefficient of the flow waveform	0.9 - 1	
Low-frequency noise cutoff (Hz)	(filter)	depends on f0	
Duration quotient (DQ)	Relative length of the non-attenuated section of AME window	0.1 - 1	For QCP only
Position quotient (PQ)	Relative starting point of the non-attenuated section of AME window	0 - 0.15	For QCP only
Ramp length	Number of transitional samples of the AME window	1 - 20	For QCP only
AR modeling method	Selection of the vocal tract modelling method	DAP, LPC, and MVDR	For IAIF only
Glottis filter order	Prediction order of the glottal flow filter	2 - 8	For IAIF only
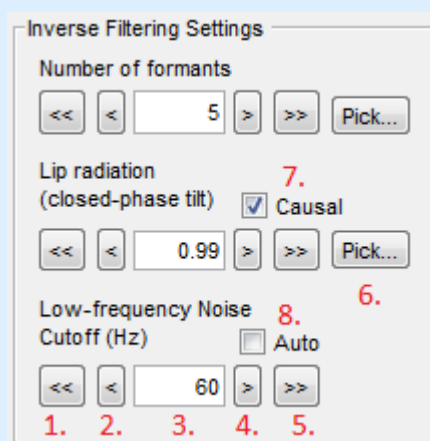
Figure 5: Close-up to general parameter settings

Legend
1 Decrease the value by two steps
2 Decrease the value by one step
3 Open box to show the value and edit it manually
4 Increase the value by one step
5 Increase the value by two steps
6 Pick button (see more at Fig. 7)
7 Causality option (on / off)
8 Set low-freq noise cut automatically (selecting "auto" disables other low-freq handles)

The best parameter values can be found by iterating step by step or writing the exact numbers in the open box. However, as this is not always the fastest way to work there is also a "Pick" button in Aparat. "Pick" allows the user to hand-pick the best-looking waveform.
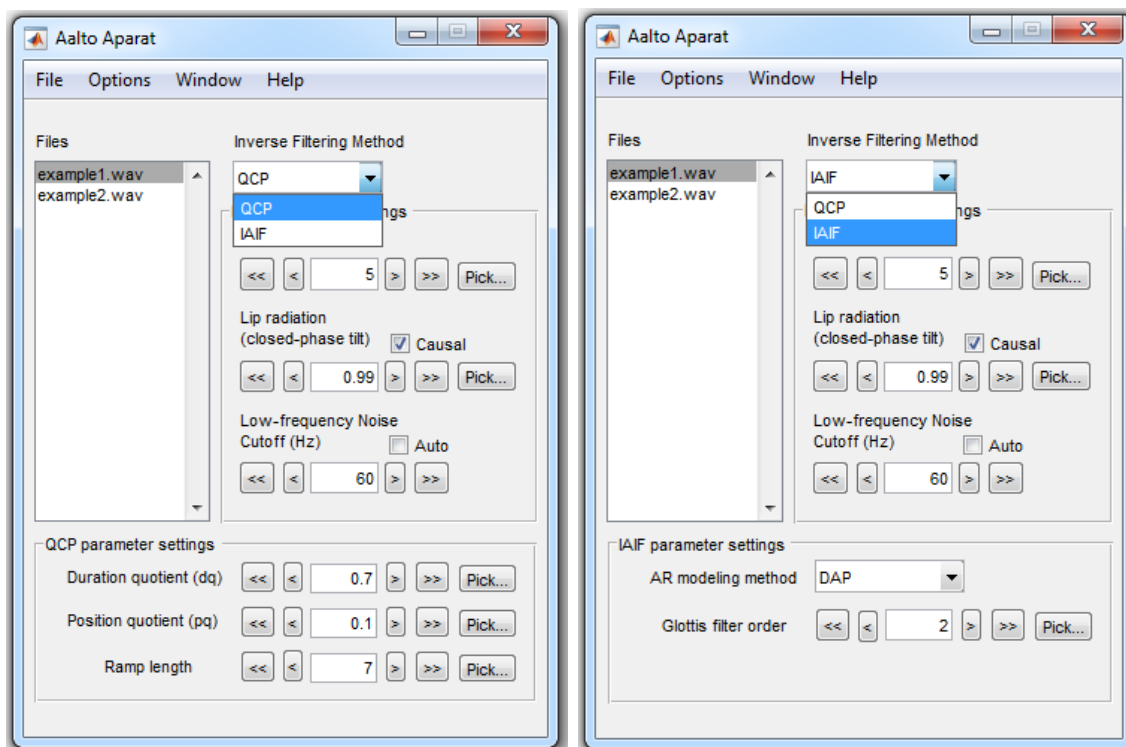
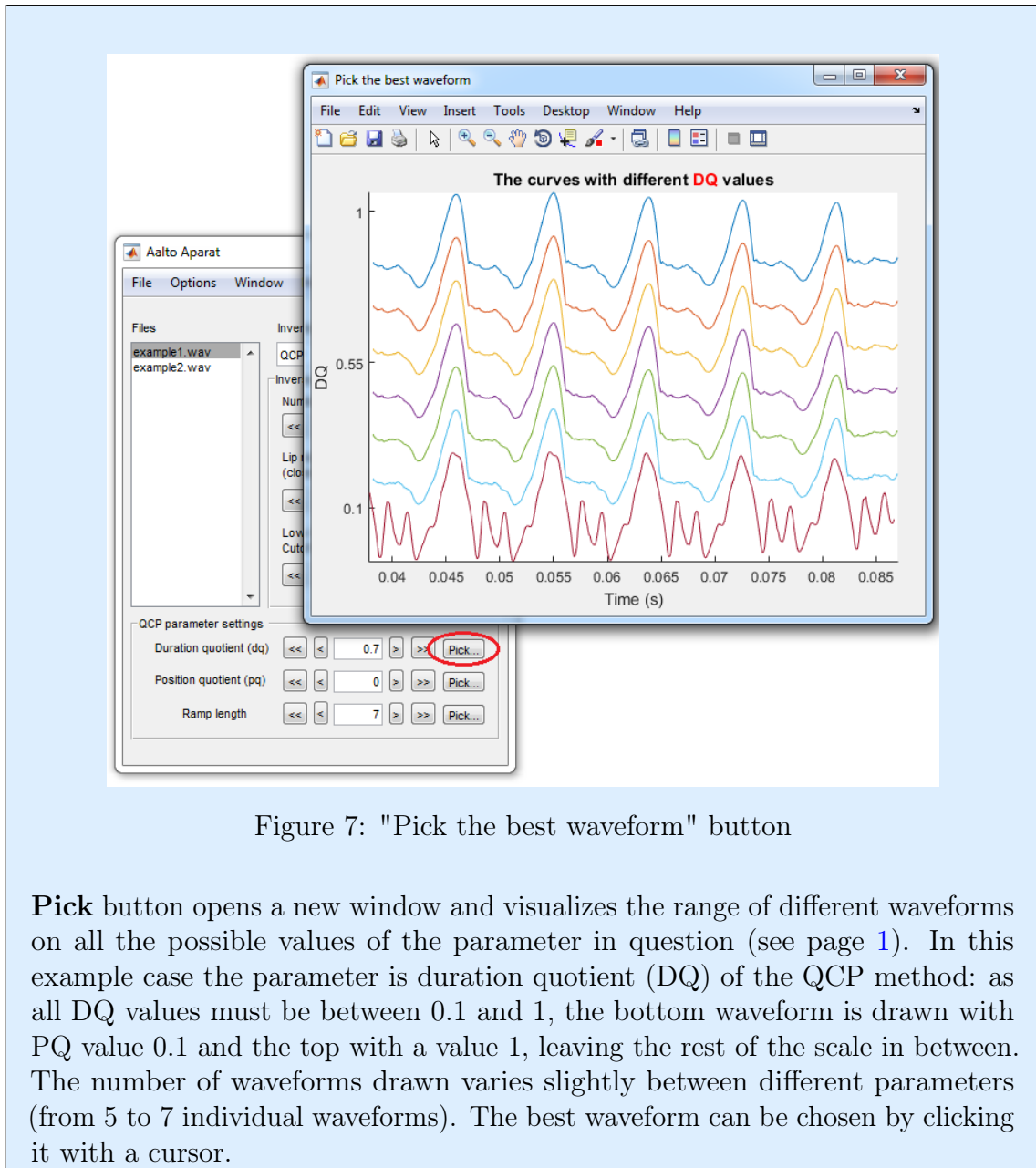Figure 6: The appearance of the control window changes according to method in use

Figure 7: "Pick the best waveform" button

Pick button opens a new window and visualizes the range of different waveforms on all the possible values of the parameter in question (see page 1). In this example case the parameter is duration quotient (DQ) of the QCP method: as all DQ values must be between 0.1 and 1, the bottom waveform is drawn with PQ value 0.1 and the top with a value 1, leaving the rest of the scale in between. The number of waveforms drawn varies slightly between different parameters (from 5 to 7 individual waveforms). The best waveform can be chosen by clicking it with a cursor.

3.4   Other options and tools

In addition to parameter settings Aparat has several different tools to analyse the inverse filtered signal. These tools can be found on a menu bar of the control window (see Fig. 8).
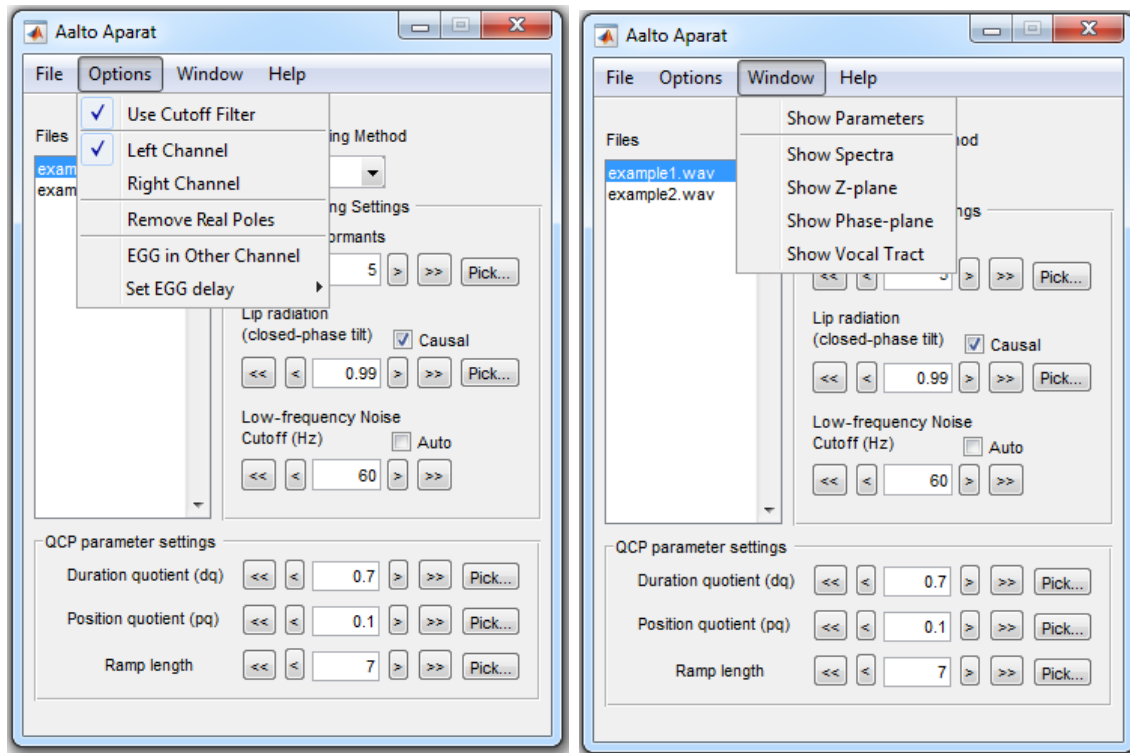

Figure 8: Advanced options and tools can be found on the menu bar

In **Options** menu there are special options in connection to the input signal: channel election, cutoff filter on/off, real poles removal from the vocal tract model, and settings for EGG analysis. In **Window** menu the glottal source parametrisation tool and graphic presentations of spectra, z-plane, phase-plane, and vocal tract can be viewed (Fig. 9-13).
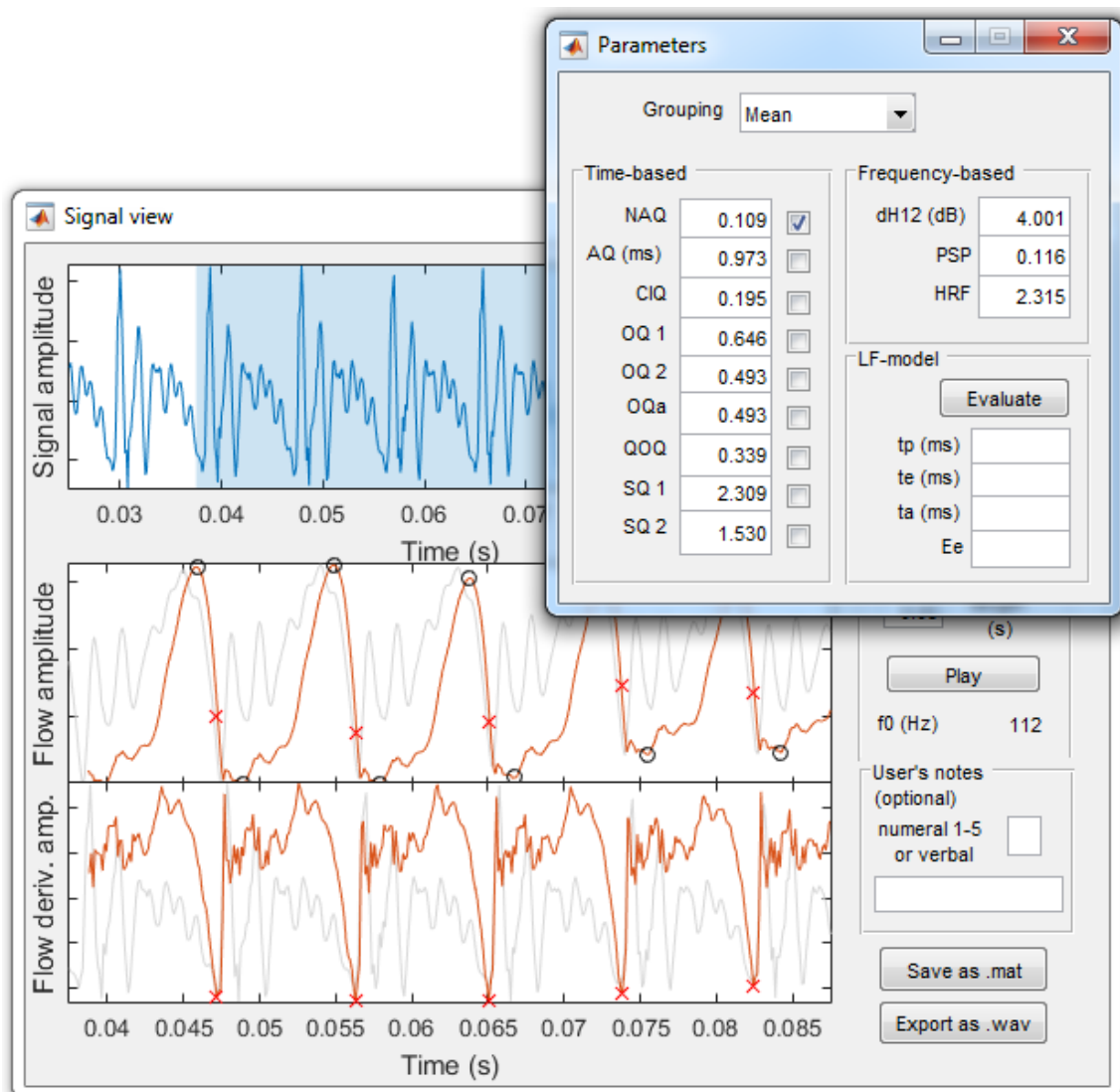
Figure 9: The parameter window shows the parameters computed from a glottal flow estimate. By clicking the checkbox next to the value (i.e. NAQ as in the figure) the time-domain instants related to the computation of the parameter can be seen on the signal view window [1].
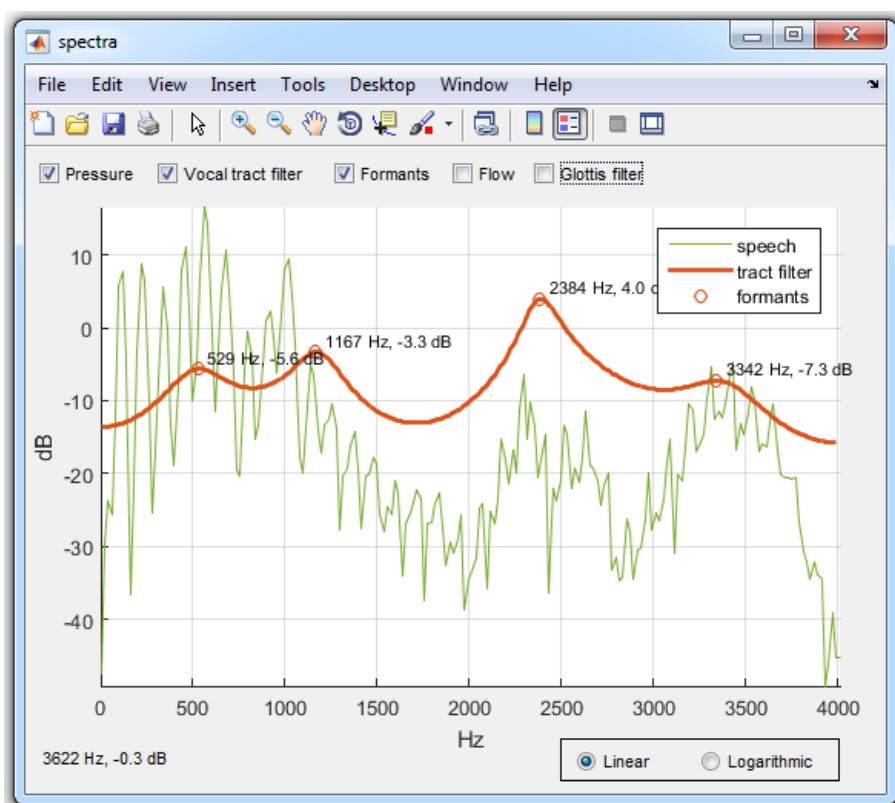
Figure 10: The spectra window is able to illustrate signal pressure, vocal tract filter, formants, glottal flow and glottis filter all in the same plot, using either the linear or logarithmic scale. In this example, only pressure, vocal tract filter and formants are selected: the selection is made by clicking corresponding check boxes.
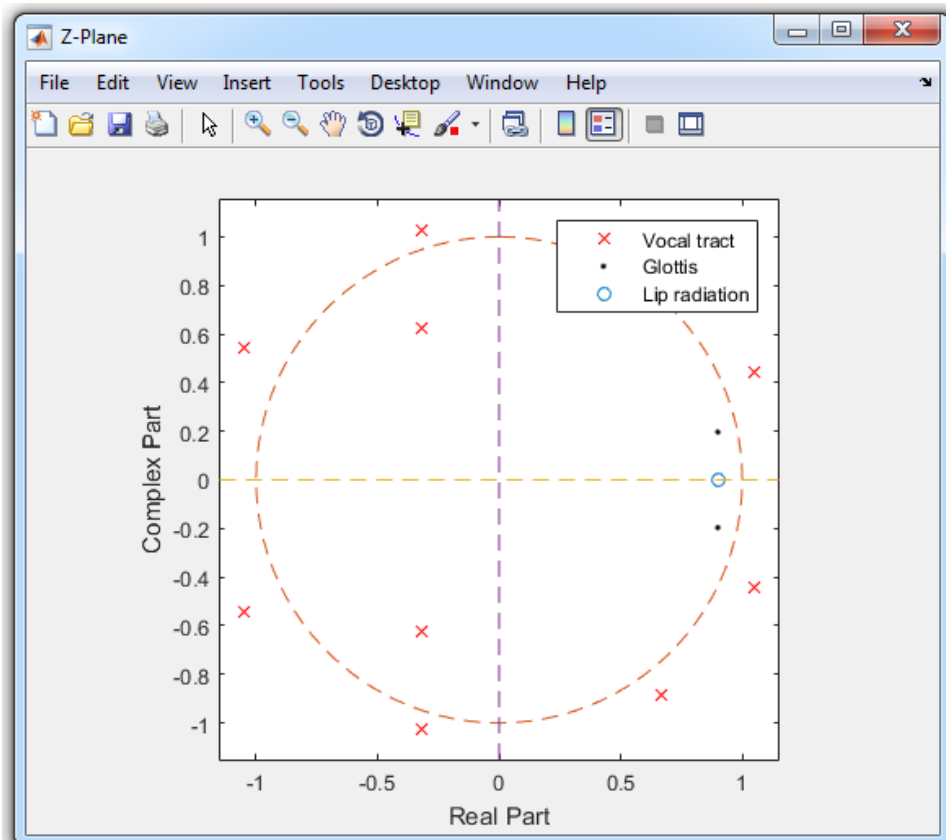
Figure 11: Z-plane figure plots the current vocal tract filter estimate on the z-plane. The resonances of different stages are indicated differently.
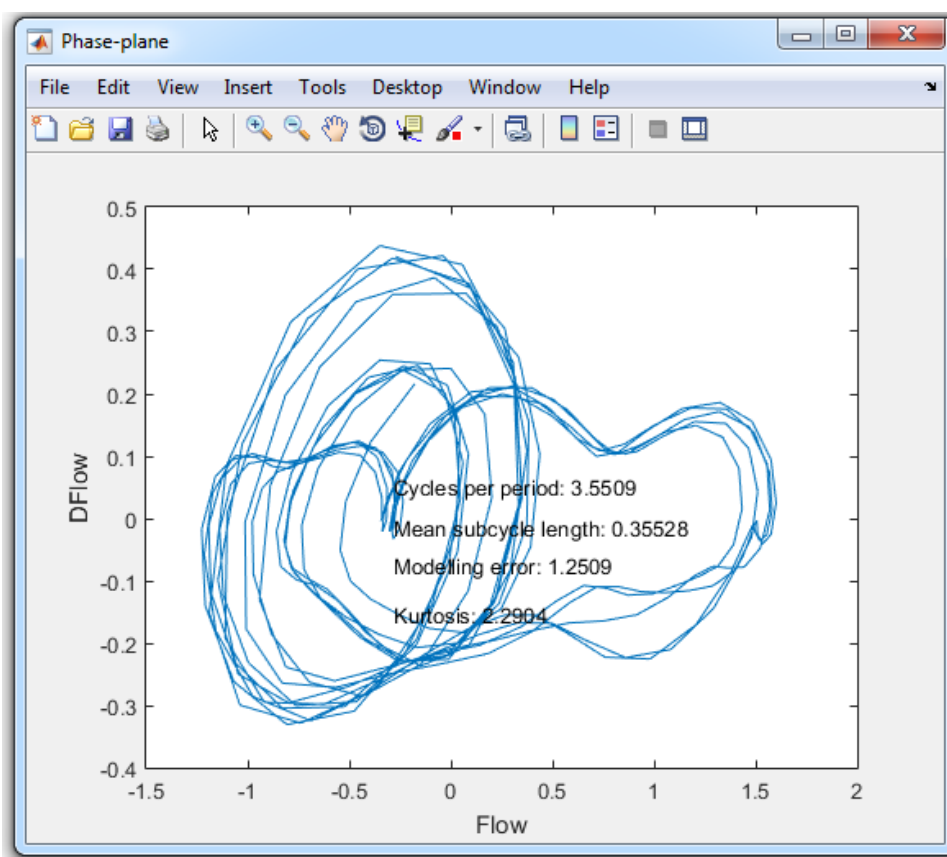
Figure 12: Phase-plane figure shows an xy-plot with the glottal flow samples on the x-axis and the corresponding samples of the flow derivative on the y-axis. It also announces the cycles per period, mean sub-cycle length, modelling error and kurtosis.
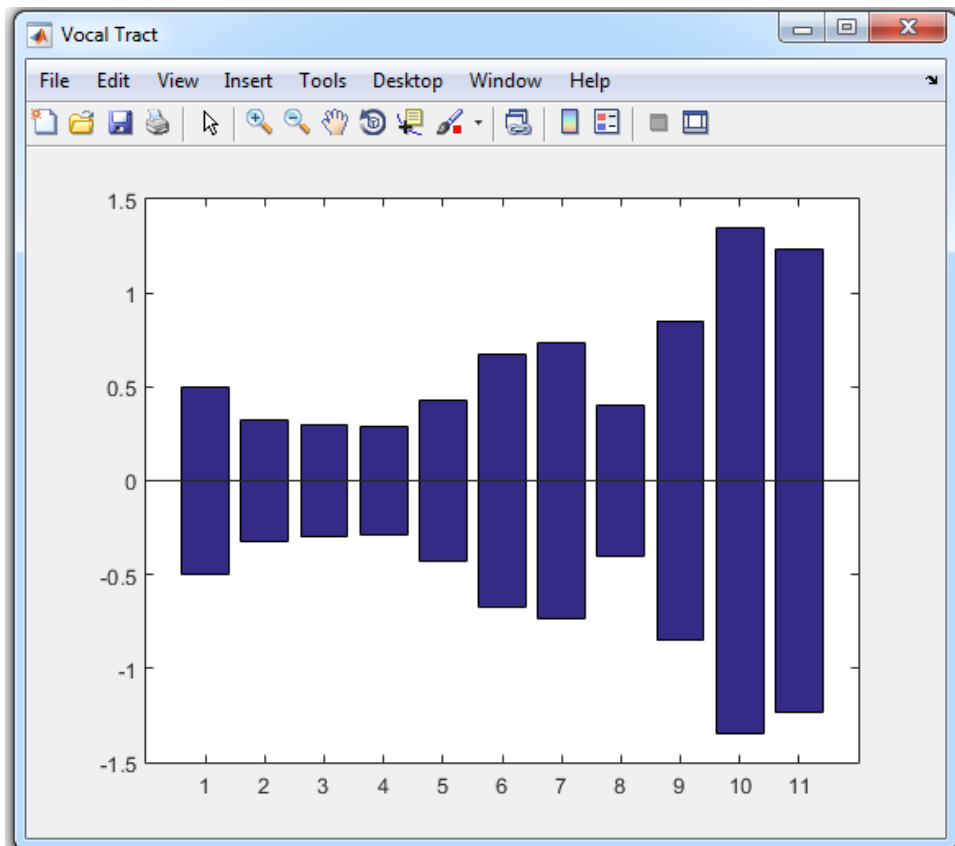
Figure 13: Vocal tract view shows a plot of the cross-sectional diameter of the tube model derived from the vocal tract filter.

3.5 Saving data

Aparat provides two ways to save your work: saving all current parameter values and selection as a .mat file and saving the selected sound pressure and glottal flow as .wav files. The save buttons can be found on both the control window (under the File menu, see Fig. 14) and the bottom right corner of the signal view window.

The **Export as .mat** command names the new .mat file after the original wav. file in process, saves all current values and places the file itself to the work space folder of the programme. The existence of the .mat file can be viewed on the Aparat control window as there is an asterisk (*) in the beginning of the name of the original signal (Fig. 3), and the saved parameter values are now default settings for this particular .wav file . By contrast, the **Save as .wav** command creates a new folder (exportWavs) and saves the samples there, and they are not shown in the Aparat control window without changing the work space folder. The names of the .wav samples include the name of the original signal followed by "Flow" (for glottal flow) or "Pressure" (for selected sound pressure).

Note: to make several different samples of the same signal there must be several differently named copies of the original as well, as Aparat overwrites the existing sample(s) with a new one if the names are identical.
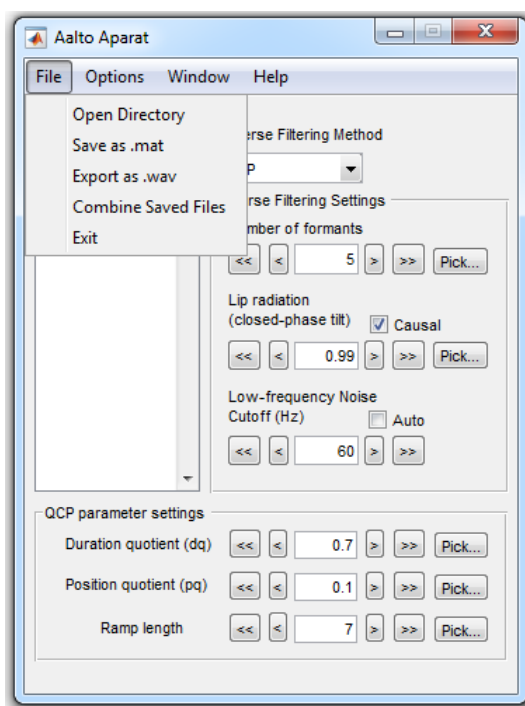

Figure 14: Saving options as well as exit can be found under File menu

After making various .mat files there may be need to compare them and process them as a one file. For that purpose there is a **Combine Saved Files** command on the File menu (Fig. 14). The command collects all the .mat files from the work space and packs them to a file named "combined.tab".

References

[1] M. Airas. TKK Aparat: An environment for voice inverse filtering and parameterization. *Logopedics Phoniatrics Vocology*, 33(1):49–64, 2008.

[2] M. Airaksinen, T. Raitio, B. Story, and P. Alku. Quasi closed phase glottal inverse filtering analysis with weighted linear prediction. *IEEE Transactions in Audio, Speech and Language processing*, 22(3):596–607, 2014.

[3] P. Alku. Glottal wave analysis with pitch synchronous iterative adaptive inverse filtering. *Speech Communication*, 11(2-3):109–118, 1992.

[4] D. Wong, J. Markel, and A. Gray Jr. Least squares glottal inverse filtering from the acoustic speech waveform. *Audio, Speech, and Language Processing, IEEE Transactions on*, 27(4):350–355, 1979.

[5] C. Ma, Y. Kamp, and L. Willems. Robust signal selection for linear prediction analysis of voiced speech. *Speech Communication*, 12(1):66–81, 1993.

[6] P. Alku, J. Pohjalainen, M. Vainio, A.-M. Laukkanen, and B. Story. Formant frequency estimation of high-pitched vowels using weighted linear prediction. *The Journal of the Acoustical society of America*, 134(2):1295–1313, 2013.

[7] T. Drugman, M. Thomas, J. Gudnason, P. Naylor, and T. Dutoit. Detection of glottal closure instants from speech signals: A quantitative review. *Audio, Speech, and Language Processing, IEEE Transactions on*, 20(3):994–1006, 2012.

[8] G. Fant, J. Liljencrants, and Q. Lin. A four-parameter model of glottal flow. *STL-QPSR*, 26(4):1–13, 1985.

[9] A. El-Jaroudi and J. Makhoul. Discrete all-pole modeling. *Signal processing, IEEE Transactions on*, 39(2):411–423, 1991.

[10] M. Wölfel and J. McDonough. *Front Matter, in Distant Speech Recognition*. John Wiley & Sons, Ltd, 2009.