Sevana AQuA – Audio Quality Analyzer 8.2 User Manual



Contents

Introduction....4
Functionality....5
Requirements....5
Compare wav files....5
Testing parameters....5
Scientific Background....6
Perceptual model and audio synchronization....7
AQuA Command Line parameters....9
AQuA Usage....9
Print AQuA command line help....9
Define program mode: -mode <mod>....11

Copyright © Sevana Oü, 2026

Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Command line argument: -clibf <file>..... 12
Set source of reference sound: -src <file <fname> | gen <mode> | folder <fname> <ext>>..... 12
Set type of weight coefficients: -ct <ctype>..... 13
Set name of the file under test: -tstf <fname>..... 13
Set name of report file for folder processing mode: -frep <fname>..... 13
Set name of the file generated by the speech model: -dst <fname>..... 13
Generate full speech sounds distribution or synthesized sound: -sn <all-speech | short-noise | normal-noise | long-noise | <sname>>..... 13
Set voice type: -voit <female | male>..... 13
Set duration: -slen <num>..... 13
Set quality loss or naturallness: -qt <quality | naturalness>..... 14
Enable indication of codec speed performance: -power <on | off>..... 14
Enable energy normalization -enorm <on | off | rms>..... 14
Turns compatibility with G.711 codec on or off: -g711 <on | off>..... 14
Turns compatibility with G.729 codec on or off: -g729 <on | off>..... 14
Turns compatibility with GSM FR codec on or off: -gsm-fr <on | off>..... 14
Turns compatibility with G.723.1 codec on or off: -g723.1 <on | off>..... 14
Turns compatibility with G.726 codec on or off: -g726 <on | off>..... 14
Turns compatibility with G.728 codec on or off: -g728 <on | off>..... 14
Turns compatibility with iLBC codec on or off: -ilbc <on | off>..... 14
Turns compatibility with EVS super-wideband on or off: -evs-swb <on | off>..... 15
Set number of link points: -npnt <num | auto>..... 15
Set envelope smoothing level: -miter <num>..... 15
Turn on “waiting for key press” after showing voice quality output: -gch..... 15
Do not print the copyright banner: -nologo..... 15
Print reason of quality loss: -fau <fname>..... 15
Set voice quality output type: -ratem <% | m | p | q>..... 15
Set spectral analysis precision: -acr <num | auto>..... 16
Set delta correction mode: -decor <on | off>..... 16
Set spectrums integrating mode: -emode <normal | log | 10log>..... 16
Set signals type: -mprio <on | off>..... 16
Set initial delay: -tdel <num>..... 16
Enable perception correction: -spfrcor <on | off>..... 16

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Enable processing speech related frequency bands only: -voip <on | off>.....17
Set psychoacoustics: -psyf <on | off>.....17
Set psychoacoustics: -psyn <on | off>.....17
Set level gradation: -grad <on | off>.....17
AQuA performance calculation: -tmc <on | off>.....17
Set average levels correction: -avlp <on | off>.....17
Smart energy normalization: -smtnrm <on | off>.....17
Export spectral pairs into CSV file: -specp <num> <fname>.....17
Set program performance speed: -fst <num>.....17
Set silence trimming: -trim[-src|-tst] <a | r> <level>.....18
Set short file duration: -short <sec | -1>.....18
Calculate pitch statistics: -hist-pitch <on | off > <on | off>.....18
Calculate signal levels: -hist-levels <on | off > <on | off> <on | off>.....18
Remove audio from beginning or end of reference file: -cut-src <start-time> <end-time>.....19
Remove audio from beginning or end of test file: -cut-tst <start-time> <end-time>.....19
Set the output format for the report file: -output <txt | json>.....19
Calculate echo on reference file: -echo-src <on | off>.....20
Calculate echo on test file: -echo-tst <on | off>.....20
Set echo detection frame length: -echo-interval <milliseconds>.....20
Set minimal delay for echo detection: -echo-min-delay <milliseconds>.....20
Set max length for echo detection: -echo-max-length <milliseconds>.....20
Set channel for analysis in tested audio file: -channel-tst <index>.....20
Set channel for analysis in reference audio file: -channel-src <index>.....20
Use audio impairments analysis: -new-impairments <on | off>.....20
Set output path for impairments report: -save-impairments-report <report_path>.....21
Predict MOS with the AQuA MOS Model: -normalize-mos <off | on>.....21
PVQA configuration: -pvqa-config <path | this>.....21
Add the model inputs to the report: -report-features <on | off>.....21
Config file path: -config <path_to_config_file>.....21
Log file path: -log <path_to_log_file>.....22
Verbose logging: -verbose.....22
MCP server mode: -mcp-server.....22
AQuA Command Line usage.....23

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Compare two audio files and learn about reasons for voice quality loss....24
Test two audio files and receive audio quality score....25
Adapting AQuA to actual environment....26
Synchronizing reference and test files....27
Analysis of possible reasons for voice and audio quality loss....30
Pitch and signal level statistics....30
ASR (Active Speech Ratio)....30
RMS classic and RMS bounded....31
Trimming levels....31
Sample value (MaxSample / AvgSample / MinSample)....31
Energy (MaxEnergy / AvgEnergy / MinEnergy)....31
PowerByVAD (AvgPowerByVAD / AvgActPowerByVAD / AvgPassPowerByVAD)....32
Signal/Noise Ratio (SNR)....32
Duration distortion....32
Amplitude clipping....32
Delay/Advancing of audio signal activity....33
Audio signal activity mistiming....33
Corrupted signal spectrum....33
Visualizing signals spectrum for analysis....34
AQuA Benefits....36
AQuA Error Messages....36
Command line parameters errors....37
Runtime errors....39

Introduction

AQuA is a simple but powerful tool to provide intrusive perceptual voice quality analysis. This is the easiest way to compare two audio files and test voice quality loss between reference and degraded files. Besides this the software can also test audio codecs and generate audio signals for voice quality testing.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

AQuA gives a unique opportunity to design your own voice quality testing solution not being dependent on particular hardware and software. It is available as a library for Windows and Linux, portable to Java and mobile devices.

Functionality

Requirements

AQuA is capable to work with audio files represented in .wav or .pcm formats. Audio files should have the following parameters depending on the version one uses:

AQuA	Sample rate	Bits per sample	Compressions	Channels
Voice	8kHz	8, 16, 24, 32	Uncompressed, a-law, u-law	Mono
HD Voice	8kHz – 22.05kHz	8, 16, 24, 32	Uncompressed, a-law, u-law	Not limited (16 channels max)
WB	8kHz – 192kHz	8, 16, 24, 32	Uncompressed, a-law, u-law	Not limited (16 channels max)

Compare wav files

- Allows intrusive testing of reference wav file against degraded wav file;
- Allows measuring voice quality for any language;
- Reports percentage of quality similarity and MOS score according to ITU-T P.800.

Testing parameters

AQuA supports the following parameters for voice quality testing:

- Choosing type of quality measurement: overall quality loss or voice naturalness;
- Choosing filenames for original, test and generated audio files;
- Define type of weight coefficients: uniform, linear or logarithmic;
- Allows energy normalization;
- Allows setting envelope smoothing level from 1 to 10;
- Allows choosing source for original sound (external file or generated internally);
- Contains audio synchronization and voice activity detection;
- Provides reasons for voice quality loss (quite unique feature on the market):

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

- Duration distortion;
- Changes in signal spectrum;
- Distortion detection in low, medium and high frequency bands;
- ASR;
- Amplitude clipping;
- Signal energy analysis;

- Provides voice quality feedback in:
  - Percentage of similarity;
  - MOS value;
- Enabling advanced psycho-acoustic model:
  - Psycho-acoustic filter;
  - Normalization to loudness level at 1kHz;
  - Spectrum transform into detectable range of loudness;
- Audio synchronization – trimming silence in the beginning and end of the test file;
- Adjusting ratio between calculation performance and quality score forecast accuracy.

Scientific Background

The human ear is a non-linear system, which produces an effect named masking. Masking occurs on hearing a message against a noisy background or masking sounds.

As result of the research of the harmonic signal masking by narrow-band noise Zwiker has determined that the entire spectrum of audible frequencies could be divided into frequency groups or bands, recognizable by the human ear. Before Zwiker, Fletcher, who had named the selected frequency groups as critical bands of hearing, had drawn a similar conclusion.

Critical bands determined by Fletcher and Zwiker differ since the former has defined bands by means of masking with noise and the latter – from the relations of perceived loudness.

Sapozhkov has determined a critical band as “a band of frequency speech range, perceptible as a single whole”. In his earlier researches he even suggested that sound signals in a band could be substituted by an equivalent

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

tone signal, but experiments did not confirm this assumption. Critical bands determined by Sapozhkov differ from those determined by Fletcher and Zwiker since Sapozhkov proceeded from the properties of speech signal.

Pokrovskij has also determined critical bands on the basis of speech signal properties. According to his definition the bands provide equal probability of finding formants in them.

The value of spectrum energy in bands can be used for different purposes; one of which is the sound signal quality estimation. However, using only one author's critical bands (for example, Zwiker's critical bands are used in prototype) does not allow getting an estimation objective enough, since they show only one of the aspects of perception or speech production. AQuA can determine energy in various critical bands as well as in logarithmic and resonator bands, allowing taking into consideration more properties of hearing and speech processing.

Taking into account that the bands determined by Pokrovskij and Sapozhkov are better for speech signal and not for sound signal, in general allows increasing the accuracy of estimation depending on its purpose.

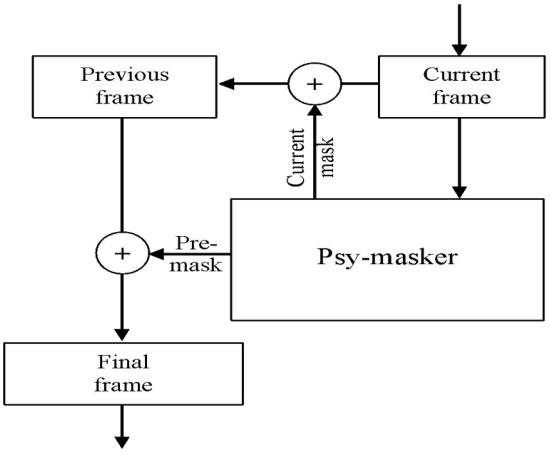
Perceptual model and audio synchronization

AQuA utilized research results of the above mentioned scientists implementing different algorithms in one software solution. AQuA also has several advantages compared to other existing voice quality measurement software:

Besides critical bands new AquA implements a more advanced psycho-acoustic model, which consists of three layers:

- psy-filtering;
- level normalization;
- transform into detectable range.

Psycho-acoustic model is based on dependencies obtained during experiments. The most complex phase is psy-filtering represented at Fig. 1.



Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Fig. 1. General scheme of psy-filtering 1

Masking procedure includes the following sequence of actions:

1. hearing threshold processing;
2. fluid level masking;
3. spectrum separation into tones and noises;
4. creating masks from tone components;
5. creating masks from noise components;
6. joining tone and noise mask components;
7. joining current mask with post-mask;
8. preparing post-mask for the next frame;
9. creating a mask for the previous frame.

Hearing threshold corresponds to ear sensitivity towards intensity of sound energy, and minimal sound pressure that produces feeling of hearing is called hearing threshold. Threshold level depends on type of sound fluctuations and measuring conditions. One of possible options to detect hearing threshold (implemented in AQuA 7.x) is standardized in ISU/R-226.

Psycho-acoustic model implemented in AQuA 7.3 introduces the so-called range of detectable loudness, which is a minimal change of signal amplitude detectable by a human ear. It's a well-known fact that depending on signal loudness level and frequency human perception varies from 2 up to 40%.

AQuA algorithms have certain advantages:

- it is universal since it allows measuring signals quality from various sources and processed in different ways;
- one can optimize quality estimation depending on the purpose:
  - for speed (for example, it is possible to receive rough estimation quickly);
  - by signal type (using different bands for speech signals and sound signals in general);
- resulting estimations correlate well with that of MOS;
- quality estimations received for speech signals can be translated in values of various scores of intelligibility.

Original and degraded signals are sent to the sound quality measurement system input. The quality of reference signal is considered 100%, the system treats it as "ideal signal", the one without distortions and degradation.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Test signal is a result of the reference transmission over the communication channel, which may contain different impairments.

AQuA Command Line parameters

AQuA Usage

`AquA-XX <license file> [options]`

<license file> is the path to the AQuA license file. If the file is not found, AQuA also looks for a file with the same name next to the executable. To use a network license, give the license server address instead: license://<host>:<port> (or lic://<host>:<port>).

Options start with a single or a double dash (-config and --config are the same option). Options can also be put into a configuration file (see -config); command line options and config files are applied in the order they are given.

Print AQuA command line help

-h prints AQuA command line help;

AQuA 8.2 prints the following help (the line after the copyright shows the license owner):

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.2.0.
Copyright (c) 2006 - 2025 by Sevana Estonia. All rights reserved.
----------------------------------------------------------------
<license owner information>
----------------------------------------------------------------
----------------------------------------------------------------
-h [sndn | exam]
              - prints this help;
                 sndn - prints list of sounds names;
                 exam - prints samples of program usage;
-mode <mod>   - defines program mode; The following modes are available
                <mod> :
                   codec    - codec testing mode;
                   files    - audio file comparison mode;
                   folder   - folder comparison mode;
                   generate - test signals generation mode;
-clibf <file> - codec library file name;
-src <file <fname> | gen <mode> | folder <fname> <ext>>
                determines source of initial sound: <file> - external sound
                file, <gen> - internal signal generator or <folder> external
                sound files from folder. In file mode one should specify name
                of audio file. In generator mode - one of signal generation
                modes: short, normal or long. In folder mode - path to audio
                files and their extensions;
-srcf <fname>
                synonym of -src file <fname>
-ct <ctype>   - sets type of weight coefficients: uniform, linear, logarithmic
                or htrsdelta;
-tstf <fname> - sets name of the file being tested; in <folder> mode sets
                name of folder, with the files being tested;
-frep <fname> - sets name of report file for folder processing mode;
-dst <fname>  - sets name of the file generated by the speech model;
-sn <all-speech | short-noise | normal-noise | long-noise | <sname>>
              - sets the name of the synthesized speech sound or runs
```

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

```
                the generation of a sound signal corresponding to one
                of the models:
                    all-speech - full distribution of speech sounds;
                    short-noise - short noise model;
                    normal-noise - the average noise model;
                    long-noise - the large noise model;
-voit <female | male>
              - sets voice type for synthesized sound;
-slen <num>   - sets duration for synthesized sound equal to <num> samples.
                num = 1..960000;
-qt <quality | naturalness>
              - sets type of quality measurement overall quality loss
                or voice naturality;
-power <on | off>
              - enable/disable output of codec speed performance indicator;
-enorm <on | off | rms>
              - enable/disable energy normalization;
-g711 <on | off>
              - enable/disable normalization to G.711 quality scale;
-g729 <on | off>
              - enable/disable normalization to G.729 quality scale;
-g723.1 <on | off>
              - enable/disable normalization to G.723.1 quality scale;
-g726 <on | off>
              - enable/disable normalization to G.726 quality scale;
-g728 <on | off>
              - enable/disable normalization to G.728 quality scale;
-ilbc <on | off>
              - enable/disable normalization to iLBC quality scale;
-gsm-fr <on | off>
              - enable/disable normalization to GSM FullRate quality scale;
-evs-swb <on | off>
              - enable/disable normalization to EVS SuperWideband quality scale (usually this is EVS codec);
-npnt <<num> | auto>
              - sets numbers of links points. num = 1..10;
                auto - enables detection of optimal amount of linking points;
-miter <num>  - sets envelope smoothing level. num = 1..10;
-gch          - turns waiting for key pressed after voice quality output on;
-fau <fname>  - prints reasons for quality loss;
-ratem <%|m|p|q> - sets output estimations: %) voice quality in percentage,
                m) MOS-like estimation, p) PESQ-like estimation q) POLQA-like estimation;
-acr <<num> | auto>
              - sets spectral analysis precision. num = 7..16,
                auto - enables automated analysis  precision detection
                according to sampling frequency;
-decor <on | off>
              - enable/disable delta correction;
-emode <normal | log | 10log>
              - integrating mode: normal - linear, [10]log - logarithmic;
-mprio <on | off>
              - sets signal type: on - music, off - voice;
-tdel <num>   - delay from start of test file;
-spfrcor <on | off>
              - On/off different frequencies perceptions correction;
-psyf <on | off>
              - On/off psy-filter;
-psyn <on | off>
              - On/off psy-normalyzer;
-smtnrm <on | off>
              - enable/disable smart normalization of energy;
-avlp <on | off>
              - enable/disable average levels correction;
-tmc <on | off>
              - allow/forbid calculation time measure;
-voip <on | off>
              - turns on/off processing of only speech related
                and specific frequency bands. In particular
                this parameter forces AQuA to consider signals
                only in the range between 300Hz and 3.4kHz
                (telephone frequency band);
-grad <on | off>
```

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

```
              - allow/forbid amplitude gradation using;
-specp <num> <fname>
              - out <num> spectral pair values to file <fname>. num = 8,16 or 32;
-fst <num>    - sets program performance speed. Increasing speed decreases accuracy;
                num = 0.0 (slow) ... 1.0 (fast);
-trim[-src|-tst] <a | r> <level>
                sets trimming type ((a)bsolute or (r)elative threshold) and level
                (0.0 <= level < 120.0). -trim-src run trimming for source file;
                -trim-tst run trimming for degraded file; -trim run trimming for
                both files;
-short <<sec> | -1>
                sets the maximal value of the short file duration (2.0 <= sec <= 180.0);
                -1 Disables operation with short files;
-hist-pitch <on | off>  <on | off>
                Allows calculating pitch statistics; First flag enables or disables
                metrics calculation. The second flag enables or disables metrics usage
                for quality evaluation;
-hist-levels <on | off> <on | off> <on | off>
                Allows calculating signal levels and quantization step histogram.
                First flag enables or disables metrics calculation. The second flag allows
                signal levels histogram usage for quality evaluation, the third flag
                allows quantization step histogram usage for quality evaluation;
-cut-src <start-time> <end-time>
                This option removes audio from the beginning or end of reference file before
                processing. <start-time> - remove <start-time> milliseconds in the beginning
                of the correspondent audio file; <end-time> - remove
                <end-time> milliseconds in the end of the correspondent audio file;
-cut-tst <start-time> <end-time>
                This option removes audio from the beginning or end of test file before
                processing. <start-time> - remove <start-time> milliseconds in the beginning
                of the correspondent audio file; <end-time> - remove <end-time> milliseconds
                <end-time> milliseconds in the end of the correspondent audio file;
-output <txt | json>
                Sets the output format for the report file.
                txt - the report file is displayed in a simple text form;
                json - the report file is output in the json format;
-echo-src <on | off>
              - allow/forbid calculation of the echo parameters on source file;
-echo-tst <on | off>
              - allow/forbid calculation of the echo parameters on test file;
-echo-interval <milliseconds>
              - specifies echo detection frame length (in milliseconds); default is 680ms;
-echo-min-delay <milliseconds>
              - specifies echo detection minimal delay (in milliseconds); default is 50ms
-echo-max-length <milliseconds>
              - specifies echo detection max length (in milliseconds); default is 2800ms
-new-impairments <on | off>
              - search for impairments in both reference & test audio.
-save-impairments-report <difference_report_path>
              - save detectors' difference report
-normalize-mos <off | on>
              - <on> reports the MOS predicted by the AQuA MOS Model, a learned model that
               combines AQuA's analysis with PVQA reports of both files. Its MOS scale is the
               one it was trained for (the distributed aqua_human.cfg: listener MOS).
               The model is read from the 'AQuA MOS Model' section of the PVQA config.
               <off> (default) reports the AQuA MOS. <on> forces -new-impairments on.
               The former <all> and <diff> values and -mos-weight were removed.
-report-features <on | off>
              - add the AQuA MOS Model inputs to the JSON report ("features").
-config <path to config file>
              - use config file to avoid too much command line parameters.
```

Define program mode: -mode <mod>

Defines AQuA mode of operation. The following modes are available:

mod:

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

codec	codec testing mode;
files	audio file comparison mode;
folder	folder comparison mode;
generate	test signals generation mode.

For example: *-mode codec*

```
aqua-v tst.lic -mode files -src file ORIGINAL_FILE -tstf REFERENCE_FILE
aqua-v.exe tst.lic -mode codec -clibf <DLL_LIBRARY_NAME> -src file <TEST_AUDIO_FILE>
aqua-v.exe tst.lic -mode codec -clibf GSM610.dll -src file short.pcm
```

Command line argument: -clibf <file>

Codec library file name.

Set source of reference sound: -src <file <fname> | gen <mode> | folder <fname> <ext>>

Determines source of initial sound:

file	external sound file
gen	internal signal generator;
folder	external sound files from folder.

In file mode one should specify name of audio file.

In generator mode - one of signal generation modes:

mode	short, normal or long.

In folder mode - path to audio files and their extensions:

ext	.wav or .pcm.

Another form of this parameter is -srcf <fname>. It is literally the same as -src file <fname>.

Determines source of initial sound as file; fname is the path to file.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Set type of weight coefficients: -ct <ctype>

ctype	uniform, linear, logarithmic or htrsdelta.

Set name of the file under test: -tstf <fname>

<folder> mode sets name of folder, with the files to test.

Set name of report file for folder processing mode: -frep <fname>

Set name of the file generated by the speech model: -dst <fname>

Generate full speech sounds distribution or synthesized sound: -sn <all-speech | short-noise | normal-noise | long-noise | <sname>>

Sets the name of the synthesized speech sound or runs the generation of a sound signal corresponding to one of the models:

all-speech	full distribution of speech sounds;
short-noise	short noise model;
normal-noise	the average noise model;
long-noise	the large noise model.

Examples:

```
aqua-v tst.lic -mode generate -sn MODEL -dst SPEECH_MODEL_FILE
```

Here is example for generating **full distribution of speech sounds model** audio signal:

```
aqua-v tst.lic -mode generate -sn all-speech -dst generated_01.pcm
```

also can specify separate sounds from the table of sounds you can see in the manual.

```
aqua-v tst.lic -mode generate -sn a0 -voit female -slen 8000 -dst generated_02.pcm
aqua-v tst.lic -mode generate -sn i0 -voit male -slen 8000 -dst generated_03.pcm
```

For separate sounds one can also set type of voice "-voit male/female" and duration of the sound to be generated "-slen 8000".

Set voice type: -voit <female | male>

Sets voice type for synthesized sound.

Set duration: -slen <num>

Sets duration for synthesized sound equal to <num> samples.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

num	1 .. 960000.

Set quality loss or naturallness: -qt <quality | naturalness>

Sets type of quality measurement **overall quality loss** or **voice naturalness**.

Enable indication of codec speed performance: -power <on | off>

Enables output of codec speed performance indicator.

Enable energy normalization -enorm <on | off | rms>

Enables energy normalization.

on/off	these parameters manage amplitude normalization;
rms	this parameter turns RMS normalization on.

Turns compatibility with G.711 codec on or off: -g711 <on | off>

Enable/disable normalization to G.711 quality scale.

Turns compatibility with G.729 codec on or off: -g729 <on | off>

Enable/disable normalization to G.729 quality scale.

Turns compatibility with GSM FR codec on or off: -gsm-fr <on | off>

Enable/disable normalization to GSM FR quality scale.

Turns compatibility with G.723.1 codec on or off: -g723.1 <on | off>

Enable/disable normalization to G.723.1 quality scale.

Turns compatibility with G.726 codec on or off: -g726 <on | off>

Enable/disable normalization to G.726 quality scale.

Turns compatibility with G.728 codec on or off: -g728 <on | off>

Enable/disable normalization to G.728 quality scale.

Turns compatibility with iLBC codec on or off: -ilbc <on | off>

Enable/disable normalization to iLBC quality scale.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Turns compatibility with EVS super-wideband on or off: -evs-swb <on | off>

Enable/disable normalization to EVS super-wideband quality scale.

Set number of link points: -npnt <num | auto>

Sets number of link points:

num 1 .. 10;

auto enables detection of optimal amount of linking points (recommended). This is the default.

Set envelope smoothing level: -miter <num>

Smoothing level is in the range of [1..10].

Turn on “waiting for key press” after showing voice quality output: -gch

Turns on “waiting for a key press” after output of voice quality.

Do not print the copyright banner: -nologo

Does not print the AQuA banner (product name, version, copyright and license owner) before the results.

Print reason of quality loss: -fau <fname>

Prints reasons for quality loss to the file specified.

Set voice quality output type: -ratem <% | m | p | q>

% voice/audio quality in percentage;

m MOS score prediction (objective score of P.800 MOS prediction). With -normalize-mos on it is the MOS of the AQuA MOS Model;

p PESQ-like estimation;

q POLQA-like estimation. With -normalize-mos on it is the MOS of the AQuA MOS Model as is.

Letters can be combined, e.g. -ratem %mq prints the percentage, the MOS and the POLQA-like value.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Set spectral analysis precision: -acr <num | auto>

Sets spectral analysis precision.

num	7 .. 16;
auto	enables automated analysis precision detection according to sampling frequency.

Set delta correction mode: -decor <on | off>

Enable/disable delta correction.

Set spectrums integrating mode: -emode <normal | log | 10log>

Sets one of the integration modes:

normal	linear;
[10]log	logarithmic.

Set signals type: -mprio <on | off>

Sets signal type:

on	music;
off	voice.

Set initial delay: -tdel <num>

Sets delay in samples <num> from the beginning of test file. In order to obtain correct number of samples for certain period in milliseconds please use this formula:

$$num = \frac{delay\,(ms) \cdot Sampling\ frequency\,(Hz)}{1000}$$

and vice versa:

$$delay\,(ms) = \frac{num \cdot 1000}{Sampling\ frequency\,(Hz)}$$

Enable perception correction: -spfrcor <on | off>

Turns on/off perception correction. This option introduces additional coefficients to specific frequencies is preferred for VoIP or G.729 signal only (8kHz only).

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Enable processing speech related frequency bands only: -voip <on | off>

Turns on/off processing of only speech related and specific frequency bands. In particular, this parameter forces AQuA to consider signals only in the range between 300Hz and 3.4kHz (telephone frequency band). When the option is turned on differences in signals spectrum outside of the range above is not considered. This option is recommended for VoIP, mobile, PSTN and converged networks transmitting telephone-like speech signals.

Set psychoacoustics: -psyf <on | off>

Sets psycho-acoustic filter on/off.

Set psychoacoustics: -psyn <on | off>

Sets psycho-acoustic normalizer on/off.

Set level gradation: -grad <on | off>

Allows/forbids amplitude gradation.

AQuA performance calculation: -tmc <on | off>

Allows/forbids quality score calculation time measurement.

Set average levels correction: -avlp <on | off>

Enables/disables average levels correction.

Smart energy normalization: -smtnrm <on | off>

Enables/disables smart energy normalization. Performs energy normalization according to energy levels in integral spectrums of the most significant frequency band.

Export spectral pairs into CSV file: -specp <num> <fname>

Exports specified amount (<num>) of spectral pairs into the file specified (<fname>).

num	this parameter may be equal to 8, 16 or 32. This is important for visualizing differences in original and degraded signals' spectrums.

Set program performance speed: -fst <num>

Sets program performance speed. Increasing the speed decreases score accuracy.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

num this parameter should be in the range between 0.0 (slow) and up to 1.0 (fast).

Set silence trimming: -trim[-src|-tst] <a | r> <level>

This option trims silence in the beginning and end of the file(s) up to the predefined silence level.

Sets silence trimming type: absolute (a) (should be below average signal level), or relative (r) threshold (should be below SNR level), the <level> parameter is set in dB and varies from 0.0 up to 120.0.

-trim runs trimming for both files (synchronizes both audio files in time domain);

-trim-src runs trimming for source file only;

-trim-tst runs trimming for degraded file only;

a absolute threshold, audio below <level> set in dB will be removed in the beginning and end of the audio files;

r relative threshold, audio with SNR below <level> set in dB will be removed in the beginning and end of the audio files;

level signal trimming level is set in dB and varies from 0.0 up to 120.0

Set short file duration: -short <sec | -1>

Sets the maximal value of the short file duration.

-1 this value disables operation with short files. This option guarantees quality measurement for the files with *duration* of the *audio* from 2 seconds to defined value <sec> regardless of the amount of active sound in the recording;

sec 2 .. 180.

Calculate pitch statistics: -hist-pitch <on | off > <on | off>

Allows calculating pitch statistics.

First flag enables or disables metrics calculation;

The second flag enables or disables metrics usage for quality evaluation.

Calculate signal levels: -hist-levels <on | off > <on | off> <on | off>

Allows calculating signal levels and quantization step histogram.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

First flag enables or disables metrics calculation;

The second flag allows signal levels histogram usage for quality evaluation;

The third flag allows quantization step histogram usage for quality evaluation.

Remove audio from beginning or end of reference file: -cut-src <start-time> <end-time>

This option removes audio from the beginning or end of the reference file before processing.

start-time	remove <start-time> milliseconds in the beginning of the corresponding audio file;
end-time	remove <end-time> milliseconds in the end of the corresponding audio file.

Remove audio from beginning or end of test file: -cut-tst <start-time> <end-time>

This option removes audio from the beginning or end of the test file before processing.

start-time	remove <start-time> milliseconds in the beginning of the corresponding audio file;
end-time	remove <end-time> milliseconds <end-time> milliseconds in the end of the corresponding audio file.

Set the output format for the report file: -output <txt | json>

This option sets the output format for the report file.

txt	the report file is displayed in a simple text form;
json	the report file is output in the json format.

The JSON report (written to the -fau file) is an AQuAReport object with these members:

- Version – the AQuA version;
- QualityResults – the scores selected with -ratem: Percent, MOS, PESQ, POLQA;
- Reasons – the reasons for quality loss, as in the text report;
- ValuesArray – the reference and test signal metrics (SNR, ASR, Raw ASR, RMS, energies, sample values, amplitude clipping, trimming), described in “Analysis of possible reasons for voice and audio quality loss”;
- SpectrumPairs – the spectral pairs, when -specp is set;
- impairments and impairments_report_diff – the PVQA impairments of both files and their difference, when the impairments analysis is on and -save-impairments-report is embed;
- mos_model – the AQuA MOS Model that produced the MOS, with -normalize-mos on;
- features – the AQuA MOS Model inputs, with -report-features on.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Calculate echo on reference file: -echo-src <on | off>

This option allows/forbids calculation of the echo parameters on source file.

Calculate echo on test file: -echo-tst <on | off>

This option allows/forbids calculation of the echo parameters on test file.

Set echo detection frame length: -echo-interval <milliseconds>

Specifies echo detection frame length (in milliseconds).

milliseconds 5 .. 120000. Default is 680 ms.

Set minimal delay for echo detection: -echo-min-delay <milliseconds>

Specifies echo detection minimal delay (in milliseconds).

milliseconds 5 .. 120000. Default value is 50 ms.

Set max length for echo detection: -echo-max-length <milliseconds>

Specifies echo detection max length (in milliseconds).

milliseconds 5 .. 120000. Default value is 2800 ms.

Set channel for analysis in tested audio file: -channel-tst <index>

Selects the audio channel of the test file to analyze: <index> is 0..9 (0 is the first channel). -1 or no value mixes all channels together to a single mono signal; this is the default.

Set channel for analysis in reference audio file: -channel-src <index>

Selects the audio channel of the reference file to analyze: <index> is 0..9 (0 is the first channel). -1 or no value mixes all channels together to a single mono signal; this is the default.

Use audio impairments analysis: -new-impairments <on | off>

Enables the search for audio impairments (PVQA detectors) in both the reference and the test file. Default is off.

The PVQA detector configuration is set with -pvqa-config; usually it is part of the AQuA config file (pvqa-config: this).

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Older versions required a separate pvqa.cfg file; such files are still supported. -normalize-mos on and -save-impairments-report turn this option on.

Set output path for impairments report: -save-impairments-report <report_path>

Specifies output file to save impairments report. Implicitly enables audio impairments collecting (-new-impairments on)

report_path Path to output file where impairments saved. embed puts the report into the JSON report instead.

Predict MOS with the AQuA MOS Model: -normalize-mos <off | on>

Selects which MOS AQuA reports.

- on – reports the MOS predicted by the AQuA MOS Model: a learned model that combines the AQuA analysis with the PVQA impairment reports of the reference and the test file. The model is the “AQuA MOS Model” section of the PVQA configuration (see -pvqa-config). Its MOS scale is the one it was trained for: listener MOS for the distributed aqua_human.cfg. Implies -new-impairments on.
- off – reports the AQuA MOS. Default behavior.

The values all and diff and the option -mos-weight were removed in AQuA 8.2; using them is an error. With on, the PVQA configuration must contain the AQuA MOS Model section, otherwise AQuA stops with “Failed to configure PVQA analyzer”.

PVQA configuration: -pvqa-config <path | this>

Path to the PVQA configuration used by the impairments analysis and the AQuA MOS Model. In a YAML AQuA config file the value this means that file itself, so one file holds the AQuA options, the PVQA detector sections and the AQuA MOS Model section (as the distributed aqua_human.cfg does).

Add the model inputs to the report: -report-features <on | off>

Adds the inputs of the AQuA MOS Model to the JSON report as “features”. Default is off.

Config file path: -config <path_to_config_file>

Path to AQuA config file. The file is either YAML with a top-level AQuA: map, or plain text with one <option name> = <option value> per line (lines starting with # or ; are comments). Option names are the command line option names without the dash. If the file is not found, AQuA also looks for a file with the same name next to the executable. Example (YAML):

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

```
AQuA:
    mode:                   files
    acr:                    auto
    npnt:                   auto
    voip:                   on
    enorm:                  rms
    ratem:                  "%m"
    fau:                    faults.json
    output:                 json
```

and run it with the files to compare:

```
aqua-wb aqua-wb.lic -config aqua.cfg -src file reference.wav -tstf test.wav
```

To report the MOS of the AQuA MOS Model, start from the distributed aqua_human.cfg instead. Its AQuA map adds pvqa-config: this, save-impairments-report: embed and normalize-mos: on, and the file goes on with the PVQA sections and the AQuA MOS Model section the model was trained with. Keep its AQuA options unchanged: the model expects them.

Log file path: -log <path_to_log_file>

Path to optional AQuA log file. By default the AQuA logs to console Quality & MOS values with some informational and error messages. In MCP server mode the console output is suppressed however the log output may be needed; otherwise usually this option is not needed.

Verbose logging: -verbose

Writes additional diagnostic messages to the log.

MCP server mode: -mcp-server

Runs the AQuA as MCP server. As usually the license file should be specified and -config option used to specify needed options. It runs in stdio mode. The server reads JSON-RPC messages line-by-line from stdin and writes responses to stdout.

The server provides two tools:

- aqua_compare – compares reference_path with test_path (both required) and returns percent, mos and faults (the reasons for quality loss). Optional: config_path (otherwise the server's -config file, or aqua-wb.cfg in the current directory), channel_src and channel_tst (default -1: all channels mixed);
- aqua_info – returns the AQuA version and product name.

For example:

```
aqua-wb aqua-wb.lic -mcp-server -config aqua.cfg -log aqua-mcp.log
```

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

AQuA Command Line usage

Most of our customers represent the following business segments:

- VoIP service providers
- Mobile service providers
- Satellite service providers
- Audio and web conferencing providers
- Radio communications
- Unified communications
- Telecom solution providers

AQuA helps telecom business to solve a wide range of tasks:

- test conference bridges quality when dialing from different locations
- monitor quality on a conference bridge
- monitor quality to certain destinations
- monitor quality at different terminations by end-to-end testing with termination's echo server
- test quality in converged networks (f.e. Mobile-VoIP)
- IVR system tests
- device testing in various environments
- audio improvement algorithms development

In all cases AQuA is the means for intrusive (active) end-to-end testing, which involves a reference audio file compared to the test one passed through a network, device or any other environment that may introduce degradation (f.e. a voice codec).

In order to show how AQuA performs perceptual voice quality assessment we are going to use the sample files from the AQuA package: jane_8k.wav is the reference recording, jane_8k_10.wav and jane_8k_40.wav are degraded copies of it, the first with mild and the second with strong degradation. The synchronization examples use variants of jane_8k_10.wav padded with silence or noise. All of them are recorded at 8 kHz sampling and are 16 bit mono, so they can be used with AQuA Voice as well as AQuA Wideband; any other audio files can be used the same way.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Note: the examples and outputs in this chapter were recorded with AQuA 8.2 (AQuA-Wideband v.8.2.0). The license line of the banner is shown as <license owner information>; scores can differ between versions.

Note: we use Linux to run the examples, so the sample scripts are .sh files and long commands are split into several lines with the shell line continuation character \. On Windows use the ^ instead of \.

Compare two audio files and learn about reasons for voice quality loss

To compare two audio files in AQuA Command Line version when one is interested to get extensive feedback from the software we suggest invoke AQuA in the following manner (sample-01.sh):

```
aqua-wb ./lic/aqua-wb.lic -mode files -src file ./wavs/jane_8k.wav \
      -tstf ./wavs/jane_8k_40.wav -acr auto -npnt auto -miter 1 -ratem %m \
      -fau sample-01-log.txt
```

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.2.0.
Copyright (c) 2006 - 2025 by Sevana Estonia. All rights reserved.
---------------------------------------------------------------
<license owner information>
---------------------------------------------------------------
---------------------------------------------------------------
File Quality is
        Percent value   66.53
        MOS value       3.55
```

Thus one can see that file comparison gives only 66.53% of similarity what corresponds to 3.55 MOS.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

After test was executed `sample-01-log.txt` file contains quantitative reasons for voice quality loss:

```
Advancing of audio signal activity.
Signal advances the original by   -160 ms.

Audio signal activity mistiming (unsynchronization) is 11.15 percent.

Value Name              Source  Degraded        Units
SNR                :    63.81   55.70           dB
ASR                :    56.30   36.70           %
Raw ASR            :    48.11   30.21           %
RMS classic        :    0.0005  0.0002
RMS bounded        :    0.0022  0.0017
Is RMSbounded valid:    true    true
AvgEnergy          :    31.00   35.16           dB
MinEnergy          :    5.03    9.06            dB
MaxEnergy          :    68.84   64.76           dB
AvgSample          :    0       -6
MinSample          :    -14128  -8928
MaxSample          :    6376    5984
AvgPowerByVAD      :    31.11   15.27           dB
AvgActPowerByVAD   :    48.53   47.65           dB
AvgPassPowerByVAD  :    14.95   1.25            dB
```

Test two audio files and receive audio quality score

In case we like to simply compare two audio files and get feedback on how similar the quality of the one under test is towards the reference audio we suggest invoking AQuA in the following manner (sample-02.sh):

```
aqua-wb ./lic/aqua-wb.lic -mode files -src file ./wavs/jane_8k.wav \
            -tstf ./wavs/jane_8k_10.wav -acr auto -npnt auto -miter 1 -ratem %m
```

Result will be:

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.2.0.
Copyright (c) 2006 - 2025 by Sevana Estonia. All rights reserved.
---------------------------------------------------------------
<license owner information>
---------------------------------------------------------------
---------------------------------------------------------------
File Quality is
        Percent value   90.96
        MOS value       4.64
```

or invoking it for the other degraded file (sample-03.sh):

```
aqua-wb ./lic/aqua-wb.lic -mode files -src file ./wavs/jane_8k.wav \
            -tstf ./wavs/jane_8k_40.wav -acr auto -npnt auto -miter 1 -ratem %m
```

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.2.0.
Copyright (c) 2006 - 2025 by Sevana Estonia. All rights reserved.
---------------------------------------------------------------
<license owner information>
---------------------------------------------------------------
---------------------------------------------------------------
File Quality is
        Percent value   66.53
        MOS value       3.55
```

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Adapting AQuA to actual environment

AQuA parameters have pre-set values by default, however, in some cases it is required to adapt the algorithm to actual environment, which is network, device, or specific codec. Majority of our customers don't require adjusting AQuA parameters, but in some cases software tuning makes test results more consistent. There is no common case when it's 100% required, but some of our customers mentioned that when doing tests in VoLTE networks, or VoIP-mobile this tuning gives better scores.

In case your tests show unexpected results means that AQuA engine or VAD may need tuning. We suggest to start with these parameters first:

- -npnt — This parameter sets the amount of linking points required to catch different "holes" inside the signal. By default it is auto.
- -miter — Sets amount of voice activity detector frames that are used during smoothing. By default it's 5. This is required to smooth the detector's vibration.

For example (sample-04.sh):

```
aqua-wb ./lic/aqua-wb.lic -mode files -src file ./wavs/jane_8k.wav \
            -tstf ./wavs/jane_8k_40.wav -acr auto -npnt 1 -miter 5 -ratem %m
```

Result is:

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.2.0.
Copyright (c) 2006 - 2025 by Sevana Estonia. All rights reserved.
---------------------------------------------------------------
<license owner information>
---------------------------------------------------------------
---------------------------------------------------------------
File Quality is
        Percent value   60.10
        MOS value       3.21
```

or invoking it for the other degraded file (sample-05.sh):

```
aqua-wb ./lic/aqua-wb.lic -mode files -src file ./wavs/jane_8k.wav \
      -tstf ./wavs/jane_8k_10.wav -acr auto -npnt 1 -miter 5 -ratem %m \
      -fau sample-05-log.txt
```

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.2.0.
Copyright (c) 2006 - 2025 by Sevana Estonia. All rights reserved.
---------------------------------------------------------------
<license owner information>
---------------------------------------------------------------
---------------------------------------------------------------
File Quality is
        Percent value   89.84
        MOS value       4.59
```

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

With these settings AQuA rates the strongly degraded file lower than with the defaults (3.21 instead of 3.55 MOS), while the mildly degraded file keeps almost the same score (4.59 instead of 4.64 MOS). One can find the reasons for voice quality loss in the sample-05-log.txt file:

```
Audio signal activity mistiming (unsynchronization) is 4.14 percent.

Value Name              Source  Degraded        Units
SNR                 :   63.81   67.26           dB
ASR                 :   58.06   50.24           %
Raw ASR             :   49.22   44.04           %
RMS classic         :   0.0005  0.0004
RMS bounded         :   0.0022  0.0024
Is RMSbounded valid:    true    true
AvgEnergy           :   31.00   35.61           dB
MinEnergy           :   5.03    1.59            dB
MaxEnergy           :   68.84   68.85           dB
AvgSample           :   0       -10
MinSample           :   -14128  -14048
MaxSample           :   6376    6496
AvgPowerByVAD       :   31.11   31.36           dB
AvgActPowerByVAD    :   48.53   48.36           dB
AvgPassPowerByVAD   :   14.95   18.38           dB
```

Synchronizing reference and test files

In many cases when monitoring voice quality in real life one receives degraded file from the network containing pauses (silence) before and/or after the actual audio. Let's consider such an example. The reference audio jane_8k.wav is a female voice pronouncing a phrase in English with the following wave form:

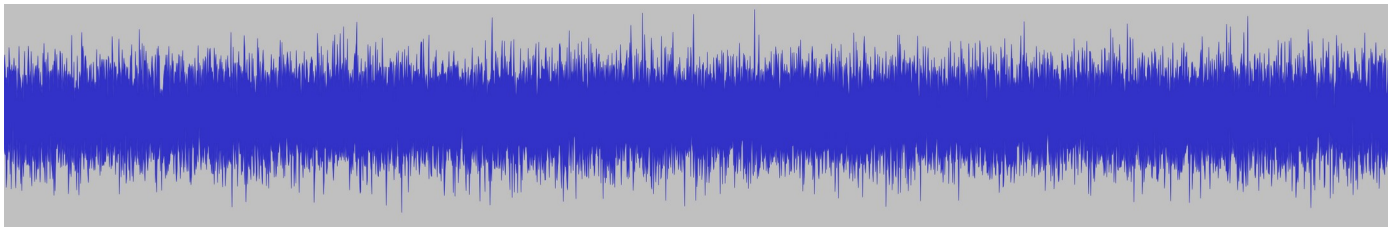


This audio is sent over a mobile network and then recorded back, but due to delays before the call is established and after hang-up degraded file has delays in the beginning and end of the audio (jane_8k_10_5s_both_-50db_noise.wav imitates such a recording):



Furthermore, if one zooms into the “silence” one will see that it contains noise:

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598



According to AQuA algorithm introduction of silence or noise into audio signal leads to quality degradation, and taking into account that establishing a test call as well as then detecting disconnect tone may take even a couple of seconds, this may significantly decrease the final quality score.

In order to trim irrelevant parts of the test signal in the beginning and end of the degraded file one just needs to invoke AQuA with a -trim parameter (sample-06.sh):

```
aqua-wb ./lic/aqua-wb.lic -mode files -src file ./wavs/jane_8k.wav \
            -tstf ./wavs/jane_8k_10_5s_both_-50db_noise.wav -acr auto -npnt auto \
            -miter 1 -trim r 5 -ratem %m -fau sample-06-log.txt
```

AQuA output will be:

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.2.0.
Copyright (c) 2006 - 2025 by Sevana Estonia. All rights reserved.
---------------------------------------------------------------
<license owner information>
---------------------------------------------------------------
---------------------------------------------------------------
File Quality is
        Percent value   82.49
        MOS value       4.30
```

The relative threshold follows the level of the signal itself, so the noise is not trimmed: sample-06-log.txt still reports the signal delayed by 6720 ms and the audio stretched by 49.80 percent.

An absolute threshold does trim the noise (sample-07.sh):

```
aqua-wb ./lic/aqua-wb.lic -mode files -src file ./wavs/jane_8k.wav \
            -tstf ./wavs/jane_8k_10_5s_both_-50db_noise.wav -acr auto -npnt auto \
            -miter 1 -trim a 30 -ratem %m -fau sample-07-log.txt
```

AQuA output will be:

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.2.0.
Copyright (c) 2006 - 2025 by Sevana Estonia. All rights reserved.
---------------------------------------------------------------
<license owner information>
---------------------------------------------------------------
---------------------------------------------------------------
File Quality is
        Percent value   90.16
        MOS value       4.61
```

Now AQuA trims 6795 ms from the head and 4548 ms from the tail of the test file (sample-07-log.txt), and the score is close to the 4.64 MOS of jane_8k_10.wav without padding.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

However, in order to be absolutely sure that the trimming works properly let's test it with an artificially created file containing silence (sample-08.sh):



```
aqua-wb ./lic/aqua-wb.lic -mode files -src file ./wavs/jane_8k.wav \
            -tstf ./wavs/jane_8k_10_5s_beginning.wav -acr auto -npnt auto -miter 1 \
            -trim a 30 -ratem %m -fau sample-08-log.txt
```

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.2.0.
Copyright (c) 2006 - 2025 by Sevana Estonia. All rights reserved.
-----------------------------------------------------------------
<license owner information>
-----------------------------------------------------------------
-----------------------------------------------------------------
File Quality is
        Percent value   90.16
        MOS value       4.61
```

and another file with silence in the beginning and the end of the file (sample-09.sh):



```
aqua-wb ./lic/aqua-wb.lic -mode files -src file ./wavs/jane_8k.wav \
            -tstf ./wavs/jane_8k_10_5s_both.wav -acr auto -npnt auto -miter 1 \
            -trim a 30 -ratem %m -fau sample-09-log.txt
```

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.2.0.
Copyright (c) 2006 - 2025 by Sevana Estonia. All rights reserved.
-----------------------------------------------------------------
<license owner information>
-----------------------------------------------------------------
-----------------------------------------------------------------
File Quality is
        Percent value   90.16
        MOS value       4.61
```

Both files get the same score as the noisy one: the trimming removed the padding completely.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Analysis of possible reasons for voice and audio quality loss

Besides audio quality score AQuA gives a possibility to determine and analyze possible reasons that caused audio signal degradation. Software automatically prepares analysis results that are stored in a log file.

Additional audio quality metrics returned by the system may not look trivial to understand and this chapter is devoted to the main principles of how these metrics are built and how one can interpret them.

AQuA returns additional metrics only in the case when they are out of range for their “typical values” (exception Signal/Noise Ratio (SNR) that is always present in the report). In case the metrics are within the range the system returns “Cannot determine the major reason for audio quality loss”.

Pitch and signal level statistics

Pitch and signal level statistics AQuA builds according to level histograms and rate of sample values and pitch change, e.g.:

```
Value Name              Source  Degraded        Units
SNR                :    63.81   55.70           dB
Average pitch      :    223.75  204.23          Hz
Pitch delta        :    2.02    2.06            Hz
ASR                :    56.30   36.70           %
Raw ASR            :    48.11   30.21           %
RMS classic        :    0.0005  0.0002
RMS bounded        :    0.0022  0.0017
Is RMSbounded valid:    true    true
AvgEnergy          :    31.00   35.16           dB
MinEnergy          :    5.03    9.06            dB
MaxEnergy          :    68.84   64.76           dB
AvgSample          :    0       -6
MinSample          :    -14128  -8928
MaxSample          :    6376    5984
AvgPowerByVAD      :    31.11   15.27           dB
AvgActPowerByVAD   :    48.53   47.65           dB
AvgPassPowerByVAD  :    14.95   1.25            dB

Pitch frequencies distribution distortion  : 40.77.
Samples values distribution distortion     : 64.76.
Quantization steps distribution distortion : 54.30.
```

ASR (Active Speech Ratio)

ASR (Active Speech Ratio) calculation is based on VAD algorithm as division of the number of active speech frames by the overall number of frames in the signal. ASR is represented in percentage, e.g.:

```
Value Name              Source  Degraded        Units
ASR                :    57.41   51.49           %
Raw ASR            :    48.11   22.87           %
```

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Raw ASR (which is present in latest AQuA version) is the same ASR but calculated before the source and/or degraded signal is cutted due to trim settings.

RMS classic and RMS bounded

RMS “classic” (Root Mean Square) for both signals is calculated at the first step of processing if RMS normalization is turned on (-enorm rms). The system normalizes RMS of degraded signal to match with RMS of the reference signal.

```
Value Name              Source  Degraded        Units
RMS classic         :   0.0005  0.0002
RMS bounded         :   0.0022  0.0023
Is RMSbounded valid:    true    true
```

RMS “bounded” excludes low level part of the signal and peak values of the audio samples that occur due to degradation. This parameter is used to optimize normalization by RMS. In case there is enough data in the signal to calculate this parameter it will be used for normalization, otherwise classic RMS is applied. Validity of RMSbound (meaning that there was enough data to calculate it) is present in AQuA report file, e.g. “Is RMSbounded valid” can be true of false.

Trimming levels.

Depending on input trim parameter audio trimming from the head and/or tail may happen. “Trimming level” refers to set parameter, “Trimming in head” and “Trimming in tail” means how much time was cut before the analysis.

```
Value Name              Source  Degraded        Units
Trimmed             :   true    true
Trimming Level      :   30.00   30.00           dB
Trimming in head    :   1796    6795            ms
Trimming in tail    :   0       4548            ms
```

Sample value (MaxSample / AvgSample / MinSample)

Just a max / average / min sample values in audio converted to PCM16.

Energy (MaxEnergy / AvgEnergy / MinEnergy)

These values are energy found during the VAD analyzer work on its work segment; SNR result depends on it.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

PowerByVAD (AvgPowerByVAD / AvgActPowerByVAD / AvgPassPowerByVAD)

These values are similar to Energy values; but they are computed on the 3x smaller segments and also differentiate on all (AvgPowerByVAD), voice (AvgActPowerByVAD) and silence (AvgPassPowerByVAD) segments.

Signal/Noise Ratio (SNR)

These metrics represent SNR both in the original and degraded files.

```
Value Name              Source      Degraded    Units
SNR                :    XX.XX       XX.XX       dB
```

These metrics show the signal/noise ratio of the original and degraded signals. Typically signal quality gets lower when SNR value decreases (or significantly increases) in the test audio.

Duration distortion

This metric represents continuity of compared audio files. Ideally amount of audio data in the original signal and file under test should be the same. During audio processing or transfer over communication channels audio fragments may be lost as well as inserted into the audio. If such audio degradation took place then value of this metric is lower than 100. The bigger the difference the stronger the degradation, however, this metric does not consider possible starting pauses.

When the value is less than 100% this means that audio data was lost and analysis result will be:

```
Audio shrinking corresponds to XX.XX percent.
```

where XX.XX corresponds to deviation from 100%.

When the actual value is more than 100% this means that data was inserted and analysis result will be:

```
Audio stretching corresponds to XX.XX percent.
```

where XX.XX corresponds to deviation from 100%.

Tolerance range for this value is set to 100% ± 1%.

Amplitude clipping

Amplitude clipping impairment or the so called “buzziness” is related to the fact if the signal amplitude is too high at some point along the analog voice path, when the voice signal is converted to a digital form amplitude clipping can occur. Users report that speech may seem excessively loud and potentially "buzzy" or "fuzzy". In case amount of clipped samples is higher than 2% the audio quality gets considerably low. One often reason for having amplitude clipping impairment in network is gateway amplification settings on the voice path, e.g.:

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

```
Value Name              Source  Degraded        Units
MinSample          :    -14128  -32768
MaxSample          :    6376    32557
Amplitude Clipping :    0.00    0.03            %
```

Delay/Advancing of audio signal activity

This metric represents signal shift in test file compared to the original and determines how much active level of the test signal delays/advances active level of the reference (original) signal. When it is delayed analysis returns the following

`Signal delayed by XX.XX ms.`

where XX.XX is delay time in milliseconds. Correspondingly, when the signal advances the original the return string is

`Signal advances the original by -XX.XX ms.`

where XX.XX is advancing time.

Tolerance range for this value is interval of ±50 ms.

Audio signal activity mistiming

This metric represents unsynchronization of active levels in reference and test signals. Original (reference) audio signal and test signal are merged to determine characteristics of audio activity, and when the characteristics of audio activity do not match system increases unsynchronization counter. After processing the final unsynchronization value is presented as percentage of cases when unsynchronization was detected.

If the metric value is not zero analysis result represents it as:

`Audio signal activity mistiming (unsynchronization) is XX.XX percent.`

where XX.XX is percentage of unsynchronization. The value is not considered if it is less than 1%.

Corrupted signal spectrum

This represents a set of metrics reflecting differences in integral energy spectrums of the original signal and audio under test. If overall spectrums difference is more than 15% than analysis returns the following string:

`Corrupted signal spectrum.`

If difference in spectrums is multidirectional (goes both into positive and negative zones) analysis returns the following string:

`Vibration along the whole spectrum [-XX.XX, YY.YY] %`

where XX.XX and YY.YY are deviations to negative and positive zones correspondingly. Tolerance range of the deviation is ±5%.

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

If spectrum distortions are unidirectional (only negative or only positive) analysis returns this string:

```
Amplification approaches YY.YY %
```

When distortions are positive, or

```
Attenuation approaches XX.XX %
```

when distortions are negative.

Other metrics returned by analysis correspond to distortions occurred in different frequency groups. Analysis of different frequency bands performs in a similar manner to spectrum analysis. When talking about frequency bands in question we consider:

Low frequencies	below 1000 Hz;
Medium frequencies	from 1000 Hz to 3000 Hz;
High frequencies	greater than 3000 Hz.

When analyzing frequency bands we use different tolerance range for different bands. Distortion in low frequencies is considered when they are greater than 5%, in medium frequencies – 10% and in high frequencies – 30%.

Multidirectional spectrum changes (vibration) are considered when they are greater than 2.5% in low frequencies, 7% in medium frequencies and 15% in high frequencies.

Unidirectional distortions (no matter positive or negative) are considered when they are greater than 5% in low frequencies, 10% in medium frequencies and 25% in high frequencies.

Visualizing signals spectrum for analysis

AQuA has a special parameter to store pairs of spectrum energy in critical bands of original and degraded audio to a .csv file (sample-10.sh):

```
aqua-wb ./lic/aqua-wb.lic -mode files -src file ./wavs/jane_8k.wav \
            -tstf ./wavs/jane_8k_10_5s_both.wav -npnt auto -miter 1 -ratem %p \
            -fau sample-10-report.txt -tmc on -gch -psyn off -psyf off -smtnrm on \
            -enorm on -grad on -specp 32 sample-10-spect.csv
```

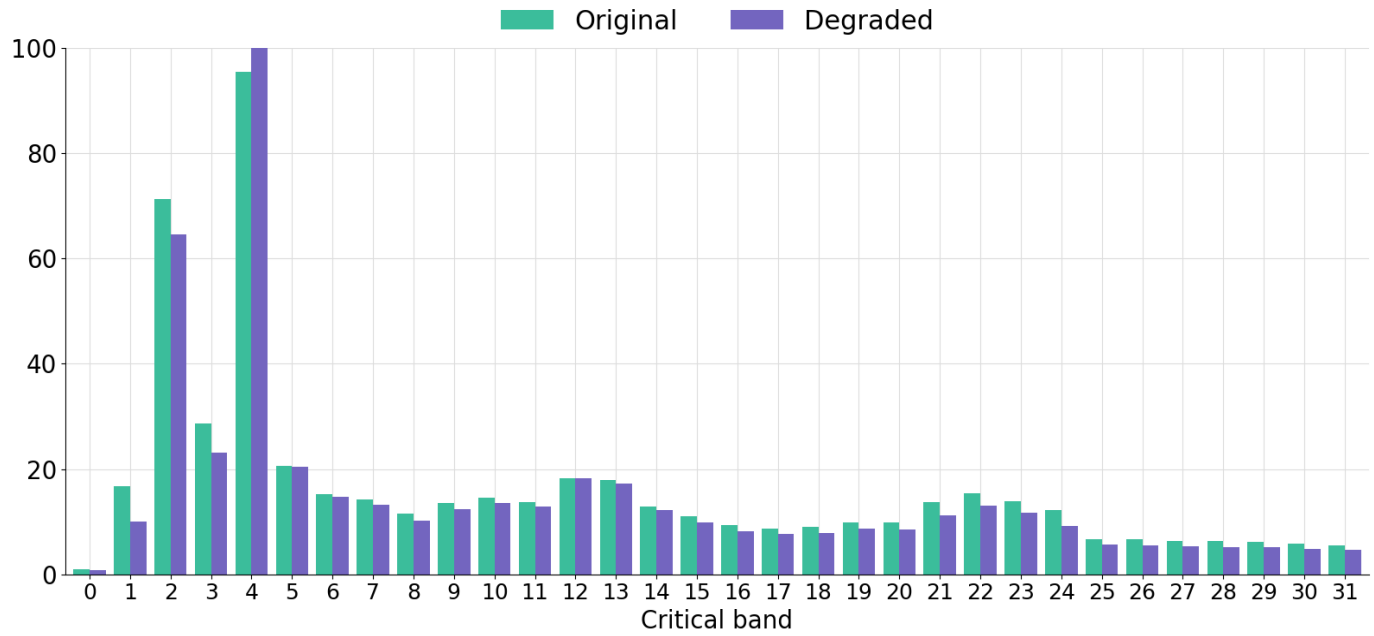
This command produces the following output:

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.2.0.
Copyright (c) 2006 - 2025 by Sevana Estonia. All rights reserved.
---------------------------------------------------------------
<license owner information>
---------------------------------------------------------------
---------------------------------------------------------------
File Quality is
        Percent value   78.79
        PESQ value      3.02
Calculating time 0.094364 sec.
```

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

```
Press any key to continue....
```

File sample-10-spect.csv contains 32 pairs related to spectrum energies of both files, so after importing the file into electronic spreadsheet we can plot a diagram visualizing differences in signals' spectrum:



As one can see the difference is not big and the reasons for received MOS score are stored in sample-10-report.txt:

```
Duration distortion.
Audio stretching corresponds to 52.36 percent.

Delay of audio signal activity.
Signal delayed by   5000 ms.

Audio signal activity mistiming (unsynchronization) is 3.71 percent.

Value Name              Source  Degraded        Units
SNR                 :   63.81   58.94           dB
ASR                 :   56.30   57.08           %
Raw ASR             :   48.11   22.87           %
RMS classic         :   0.0005  0.0002
RMS bounded         :   0.0022  0.0024
Is RMSbounded valid:    true    true
AvgEnergy           :   31.00   35.36           dB
MinEnergy           :   5.03    9.25            dB
MaxEnergy           :   68.84   68.19           dB
AvgSample           :   0       -5
MinSample           :   -14128  -14048
MaxSample           :   6376    6496
AvgPowerByVAD       :   31.11   7.49            dB
AvgActPowerByVAD    :   48.53   48.23           dB
AvgPassPowerByVAD   :   14.95   -4.58           dB
```

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

As one can see the main reasons are the delay and the duration distortion caused by the silence we have not removed. If we remove it as described in the previous chapter (sample-11.sh):

```
aqua-wb ./lic/aqua-wb.lic -mode files -src file ./wavs/jane_8k.wav \
            -tstf ./wavs/jane_8k_10_5s_both.wav -npnt auto -miter 1 -ratem p \
            -fau sample-11-report.txt -tmc on -gch -psyn off -psyf off -smtnrm on \
            -enorm on -grad on -trim r 5 -specp 32 sample-11-spect.csv
```

the delay and the duration distortion are gone and the score rises from 3.02 to 3.20:

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.2.0.
Copyright (c) 2006 - 2025 by Sevana Estonia. All rights reserved.
---------------------------------------------------------------
<license owner information>
---------------------------------------------------------------
---------------------------------------------------------------
File Quality is
        PESQ value      3.20
Calculating time 0.088639 sec.

Press any key to continue....
```

AQuA Benefits

Among AQuA benefits one will definitely appreciate that:

- AQuA is available for Windows, Linux and MAC OS operating systems;
- AQuA is available for both 32 and 64 bit systems;
- AQuA is easy to deploy and use in any production development;
- AQuA provides perceptual estimation of audio quality and can be utilized in VoIP, PSTN, ISDN, GSM, CDMA, LTE/4G, VoLTE, satellite and radio networks and combinations of those;
- AQuA perceptual model is language independent;
- AQuA can work with audio starting from 2 seconds duration;
- AQuA can work with audio longer than 30 seconds;
- AQuA is available as C/C++ native library for custom software development;
- AQuA can work with stereo audio.

AQuA Error Messages

The following error messages may appear during AQuA run time:

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

SEVANA

Command line parameters errors

Parameters listing error.
Mode error (-mode).
Data source error (-src).
Sound data generator error (-src).
Weights of sound bands error (-ct).
Synthesized voice type error (-voit).
Incorrect sound duration to synthesize (-slen).
Unknown type of quality estimation defined (-qt).
Incorrect numbers of links points (-npnt).
Incorrect envelope smoothing level (-miter).
Incorrect out quality mode (-ratem).
Incorrect accuracy value (-acr).
Test file name is missing (-tstf).
Speech model file name is missing (-dst).
Name of synthethized sound is missing (-sn).
Codec performance measurement is not set (-power on|off).
Unknown parameter in codec performance measurement option (-power).
Parameter is missing in energy normalization management option (-enorm).
Unknown parameter in in energy normalization management option (-enorm).
File name to store reasons for audio quality loss is not specified.
Parameter is missing in delta-correction management option (-decor).
Unknown parameter in delta-correction management option (-decor).
Parameter is missing in integration mode option (-emode).
Unknown parameter in integration mode option (-emode).
Parameter is missing in signal type selection option (-mprio).
Unknown argument in signal type selection option (-mprio).
Value of initial delay is missing (-tdel).
Value of initial delay is negative or integer value set is incorrect (-tdel).
Parameter is missing in option (-spfrcor).
Unknown parameter in option (-spfrcor).
Parameter is missing in psycho-acoustic model option (-psyf).
Unknown parameter in option setting psycho-acoustic filter on (-psyf).
Parameter in time measurement option is missing (-tmc).
Unknown parameter in time measurement option (-tmc).

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Parameter in smart energy normalization option is missing (-smtnrm).

Unknown parameter in smart energy normalization option (-smtnrm).

Parameter is missing in option (-avlp).

Unknown parameter in option (-avlp).

Parameter in psy-normalizer management option is missing (-psyn).

Unknown parameter in psy-normalizer management option (-psyn).

Incomplete parameters in spectrum pairs output option (-specp).

Incorrect amount of spectrum pairs (-specp).

Parameter is missing in filter management option (-voip).

Unknown parameter in filter management option (-voip).

Incompatible set of parameters of psy-normalization and amplitude ranges.

Parameter is missing in amplitude ranges management option (-grad).

Unknown parameter in amplitude ranges management option (-grad).

Parameter is missing in program performance speed option (-fst).

Program performance parameter is out of range (-fst).

Parameter(s) is missing in silence trimming option (-trim).

Unknown type of silence level detection (-trim).

Incorrect silence level (-trim).

Source audio files folder or extension are not defined (-src).

Data source type is not defined (-tstf).

Test files folder is not defined (-tstf).

Report file name is missing (-frep).

No value of the flag, enabling the shift range of estimates.

Unknown value of the flag, enabling the shift range of estimates.

No value of the parameter - the short file duration.

Unknown value of the parameter - the short file duration.

Unknown value of the parameter. Setting options (-short) must be a number [2; 180], -1.

Unknown value of the parameter. Setting options (-trim-tst) must be a number [0.00; 120.00].

Unknown value of the parameter. Setting options (-trim-src) must be a number [0.00; 120.00].

Unknown value of the parameter. Setting options (-trim) must be a number [0.00; 120.00].

Unknown value of the parameter. Setting options (-fst) must be a number [0.00; 1.00].

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Unknown value of the parameter. Setting options (-specp) must be a number 8, 16 or 32.

Unknown value of the parameter. Setting options (-tdel) must be a positive number.

Unknown value of the parameter. Setting options (-acr) must be a number [7; 16].

Unknown value of the parameter. Setting options (-miter) must be "auto" or a number [1; 10].

Unknown value of the parameter. Setting options (-npnt) must be "auto" or a number [1; 10].

Unknown value of the parameter. Setting options (-slen) must be a positive number.

Histograms management options (-hist) format error! Use help.

Unknown switch ID in histograms management options (-hist)!

Unknown histogram ID in histograms management options (-hist)!

Management trimming reference file option (-cut-src) format error!

Option (-cut-src) format error! Time must be set a number greater or equal.

Management trimming degraded file option (-cut-tst) format error!

Option (-cut-tst) format error! Time must be set a number greater or equal 0.

Option (-output) format error!

Option (-echo) format error!

Incorrect value of echo-tst parameter. on/off is expected.

Incorrect value of echo-src parameter. on/off is expected.

Incorrect value of echo-interval parameter. Value from 5 to 120000 (milliseconds) is expected.

Incorrect value of echo-min-delay parameter. Value from 5 to 120000 (milliseconds) is expected.

Incorrect value of echo-max-length parameter. Value from 5 to 120000 (milliseconds) is expected.

Not recognized <option> / Unknown option.

Got exception: Bad option '<option>' value. (Some options add the valid range.)

'normalize-mos all' (or diff) was removed: 'on' now predicts the MOS with the AQuA MOS Model from the PVQA configuration.

Problem when processing AQuA config file.

Unknown parameter: '<name>' (in a config file).

Runtime errors

Error opening source file!

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598

Error opening file under test (degraded)!

Error: files have different sampling frequencies!

Error: sampling frequency is not supported.

Error: files have different channels!

Error: sampling frequency (in source file) is not supported.

Error: sampling frequency (in degraded file) is not supported.

Error: Source sound signal duration is less than 4096 samples.

Error: Test sound signal duration is less than 4096 samples.

Error: After signal alignment source file duration became less than required.

Error: After signal alignment test file duration became less than required.

Error: Source file sound data is too short.

Error: Test file sound data is too short.

Error: Source file duration is too short.

Error: Test file duration is too short.

Error: Source file duration is too short (shorter than 2 seconds).

Error: Test file duration is too short (shorter than 2 seconds).

Error: Source file manual cut settings failed.

Error: Test file manual cut settings failed.

Error: Can't calculate Quality! Failed to configure PVQA analyzer (most probably smth bad with configuration file).

Copyright © Sevana Oü, 2026
Tulika tn 17, 10613,
Kristiine linnaosa, Tallinn,
Harju maakond, Estonia
Phone: +37253686598