Sevana AQuA – Audio Quality Analyzer 8.x User Manual

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

Contents

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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 test files. Besides this the software can also test audio codecs and generate audio signals for voice quality testing.

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
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 test 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 reference, 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 reference sound (external file or generated internally);

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

- Contains audio synchronization and voice activity detection;
- Provides reasons for voice quality loss (quite unique feature on the market):
  - 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.

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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 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;

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

- 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.

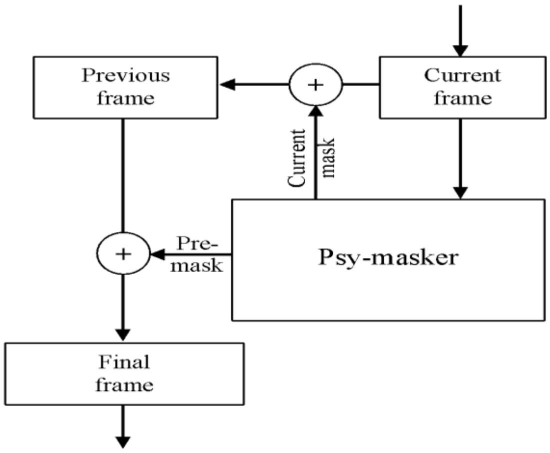


Fig. 1. General scheme of psy-filtering 1

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.

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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 intelligibilities.

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

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

AQuA Command Line parameters

AQuA Usage

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

Sevana Audio Quality Analyzer - AQuA-Wideband v.8.0.5.1279.
Copyright (c) 2021 by Sevana Estonia. All rights reserved.
----------------------------------------------------------------
Internal Sevana Ou license
----------------------------------------------------------------
Today is 22/06/19. This program is licensed to work until 22/12/31
----------------------------------------------------------------
-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;
-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;
```


Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

```
-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;
-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;
-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;
```


Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

```
-gch            - turns waiting for key pressed after voice quality output on;
-fau <fname>    - prints reasons for quality loss;
-ratem %mp      - sets output estimations: %) voice quality in percentage,
                  m) MOS-like estimation, p) PESQ-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>
                - allow/forbid amplitude gradation using;
-specp <num> <fname>
                - out <num> spectral pair values to file <fname>. num = 8,16 or 32;
```


Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

```
-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;
```


Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

```
-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
            500ms;
-echo-min-delay <milliseconds>
              - specifies echo detection minimal delay (in milliseconds);
-echo-max-length <milliseconds>
              - specifies echo detection max length (in milliseconds);
-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 | all | diff>
              - normalize MOS with audio impairments information.
               <on> option normalizes MOS for tests when audio impairments analysis
            results are worser than AQuA.
               <all> option normalizes MOS for all tests.
               Enabled option forces -new-impairments option to be set.
-mos-weight <fraction>
              - fraction is in range [0.0, 1.0].
               It is fraction of AQuA MOS taken for MOS normalization.
               Default value is 0.5
```

Print AQuA command line help

-h prints AQuA command line help;

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.0.5.1279.
Copyright (c) 2021 by Sevana Estonia. All rights reserved.
------------------------------------------------------------
Internal Sevana Ou license
------------------------------------------------------------
```

```
Today is 22/06/19. This program is licensed to work until 22/12/31
---------------------------------------------------------------
-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;
-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
                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>
```


Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

```
              - 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;
-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 %mp    - sets output estimations: %) voice quality in percentage,
                m) MOS-like estimation, p) PESQ-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;
```


Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

-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>

- 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);

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

```
-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
```


Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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 500ms;

-echo-min-delay <milliseconds>

- specifies echo detection minimal delay (in milliseconds);

-echo-max-length <milliseconds>

- specifies echo detection max length (in milliseconds);

-new-impairments <on | off>

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

```
                - search for impairments in both reference & test audio.
-save-impairments-report <difference_report_path>
                - save detectors' difference report
-normalize-mos <off | on | all | diff>
                - normalize MOS with audio impairments information.
                 <on> option normalizes MOS for tests when audio
impairments analysis results are worser than AQuA.
                 <all> option normalizes MOS for all tests.
                 Enabled option forces -new-impairments option to be set.
-mos-weight <fraction>
                - fraction is in range [0.0, 1.0].
                 It is fraction of AQuA MOS taken for MOS normalization.
                 Default value is 0.5
```

Define program mode: -mode <mod>

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

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

For example: *-mode codec*

```
aqua-v.exe 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.

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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.

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.

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

Examples:

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

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.

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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.

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).

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.

Print reason of quality loss: -fau <fname>

Prints reasons for quality loss to the file specified.

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

%	voice/audio quality in percentage;

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

m	MOS score prediction (objective score of P.800 MOS prediction).

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:

$$< \mathrm{num} >= \frac{delay(ms)\ *\ Sampling\ frequency(Hz)}{1000}$$

and vice versa:

$$\mathrm{delay(ms)} = \frac{< num > * 1000}{Sampling\ frequency(Hz)}$$

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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).

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 reference and test signals' spectrums.

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

Set program performance speed: -fst <num>

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

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 reference file only;

-trim-tst runs trimming for test 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.

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

Calculate signal levels: -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.

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.

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

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

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

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

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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

Selects audio channel in test file. Default value is -1, which means “mix all channels together to single mono”.

Set channel for analysis in reference audio file: -channel-src

Selects audio channel in reference file. Default value is -1, which means “mix all channels together to single mono”.

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

Enables audio impairments collecting. Default behavior is off (no audio impairments collecting).

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.

Enable MOS normalization: -normalize-mos <on | off | all | diff>

Forces MOS normalization depending on audio impairments analysis.

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

- on – MOS is normalized on files which audio impairments quality estimation is worse than AQuA analysis.
- off – MOS is not normalized. Default behavior.
- all – MOS normalization runs on all audio files.
- diff – MOS normalization uses information about differences in audio impairments (i.e. changes in audio impairments between reference and test files).

MOS weight coefficient: -mos-weight <weight>

Sets AQuA non-normalized MOS value weight coefficient so, that final MOS is calculated according to the following formula MOS = AQuA MOS (non normalized) * weight + audio impairments MOS * (1.0 – weight). Default value is 0.5.

Config file path: -config <path_to_config_file>

Path to AQuA config file. Config file is ф typical text file containing command line options in the following format <option name> = <option value>.

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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 WAV files one can download from Microtronix web site (http://www.microtronix.ca/pesq.html). However, one can use any audio files within AQuA Wideband or those that are recorded at 8kHz sampling and are 16 bit mono (in case of AQuA Voice).

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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.bat):

```
aqua-wb.exe ./lic/aqua-wb.lic -mode files -src file ./wavs/Or272.wav  -tstf
      ./wavs/Dg002.wav  -acr auto -npnt auto -miter 1 -ratem %m -fau sample-01-
      log.txt
```

Thus one can see that file comparison gives only 39.31% of similarity what corresponds to 2.23 MOS. By the way, this is an example of when AQuA does detect voice quality loss and PESQ does not (please read more details about this test case on Microtronix page).

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

```
Duration distortion.

Audio stretching corresponds to 1.13 percent.


Delay of audio signal activity.

Signal delayed by    120 ms.


Audio signal activity mistiming (unsynchronization) is 1.29 percent.


Corrupted signal spectrum.

Overall spectral energy distortion approaches 49.24 %

Vibration along the whole spectrum [-19.68, 29.56] %


Significant distortion in low frequencies band.

Energy distortion approaches 29.82 %

Spectrum vibration in low frequency band [-16.92, 12.90] %


Significant distortion in medium frequencies band.
```


Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

```
Energy distortion approaches 18.48 %

Amplification approaches 15.72 %


Value Name                Source    Degraded  Units

SNR                 :     63.21     71.19          dB

ASR                 :     44.02     48.96          %

RMS classic         :     0.0010    0.0014

RMS bounded         :     0.0040    0.0068

Is RMSbounded valid:      true true

AvgEnergy           :     29.67     27.28          dB

MinEnergy           :     6.89 1.66      dB

MaxEnergy           :     70.10     72.84          dB

AvgSample           :     79   13

MinSample           :     -11461    -13221

MaxSample           :     11254     14211

AvgPowerByVAD       :     29.73     26.56          dB

AvgActPowerByVAD    :     52.00     48.66          dB

AvgPassPowerByVAD   :     9.73 5.30      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.bat):

```
aqua-wb.exe ./lic/aqua-wb.lic -mode files -src file ./wavs/Or272.wav -tstf
            ./wavs/Dg001.wav  -acr auto -npnt auto -miter 1 -ratem %m
```

Result will be:

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.0.5.1279.
      Copyright (c) 2021 by Sevana Estonia. All rights reserved.
      ----------------------------------------------------------------------
```


Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

```
Internal Sevana Ou license
----------------------------------------------------------------
Today is 22/06/19. This program is licensed to work until 22/12/31
----------------------------------------------------------------
File Quality is
Percent value
97.34
MOS value       4.89
```

or invoking it for the other test file (sample-03.bat):

```
aqua-wb.exe ./lic/aqua-wb.lic -mode files -src file ./wavs/Or272.wav -tstf
        ./wavs/Dg002.wav  -acr auto -npnt auto -miter 1 -ratem %m


Sevana Audio Quality Analyzer - AQuA-Wideband v.8.0.5.1279.
        Copyright (c) 2021 by Sevana Estonia. All rights reserved.
        ---------------------------------------------------------------
        Internal Sevana Ou
        license
        ---------------------------------------------------------------
        Today is 22/06/19. This program is licensed to work until
        22/12/31
        ---------------------------------------------------------------
        File Quality is
        Percent value   51.36
        MOS value       2.80
```

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 the value is 5.
- -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.bat):

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

```
aqua-wb.exe ./lic/aqua-wb.lic -mode files -src file ./wavs/Or272.wav  -tstf
          ./wavs/Dg002.wav -acr auto -npnt 1 -miter 5 -ratem %m
```

Result is:

Sevana Audio Quality Analyzer - AQuA-Wideband v.8.0.5.1279.
Copyright (c) 2021 by Sevana Estonia. All rights reserved.
--
Internal Sevana Ou license
--
Today is 22/06/19. This program is licensed to work until 22/12/31
--
File Quality is
Percent value   53.44
MOS value   2.89

or invoking it for the other test file (sample-05.bat):

```
aqua-wb.exe ./lic/aqua-wb.lic -mode files -src file ./wavs/Or272.wav  -tstf
     ./wavs/Dg001.wav  -acr auto -npnt 1 -miter 5 -ratem %m -fau sample-05-
     log.txt

Sevana Audio Quality Analyzer - AQuA-Wideband v.8.0.5.1279.
     Copyright (c) 2021 by Sevana Estonia. All rights reserved.
     ---------------------------------------------------------------
     Internal Sevana Ou
     license
     ---------------------------------------------------------------
     Today is 22/06/19. This program is licensed to work until
     22/12/31
     ---------------------------------------------------------------
     File Quality is
     Percent value   98.13
     MOS value       4.93
```

In fact this result is much closer to what one would hear, however, the file was degraded. One can find the reasons for voice quality loss in the log.txt file, e.g.:

Delay of audio signal activity.

Signal delayed by   150 ms.

Audio signal activity mistiming (unsynchronization) is 1.02 percent.

Corrupted signal spectrum.

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

Overall spectral energy distortion approaches 48.20 %

Vibration along the whole spectrum [-20.46, 27.74] %

Significant distortion in low frequencies band.

Energy distortion approaches 30.04 %

Spectrum vibration in low frequency band [-17.66, 12.38] %

Significant distortion in medium frequencies band.

Energy distortion approaches 17.37 %

Amplification approaches 14.56 %

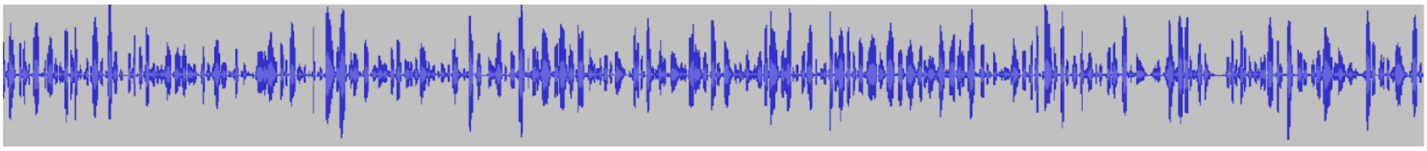
Value Name		Source	Degraded	Units
SNR	:	63.21	71.19	dB
ASR	:	44.19	48.56	%
RMS classic	:	0.0010	0.0014	
RMS bounded	:	0.0040	0.0068	
Is RMSbounded valid	:	true	true	
AvgEnergy	:	29.67	27.28	dB
MinEnergy	:	6.89	1.66	dB
MaxEnergy	:	70.10	72.84	dB
AvgSample	:	79	13	
MinSample	:	-11461	-13221	
MaxSample	:	11254	14211	
AvgPowerByVAD	:	29.73	26.56	dB
AvgActPowerByVAD	:	52.00	48.66	dB
AvgPassPowerByVAD	:	9.73	5.30	dB

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

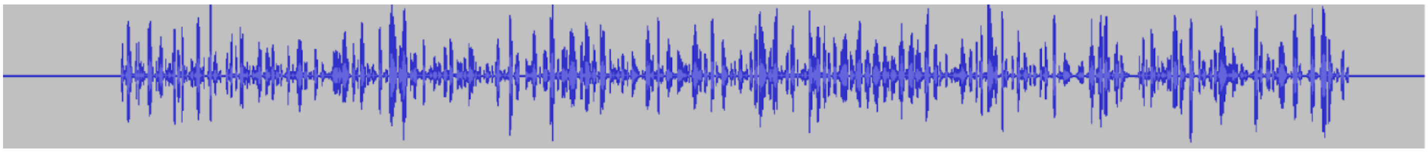
Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

Synchronizing reference and test files using AQuA 8.x

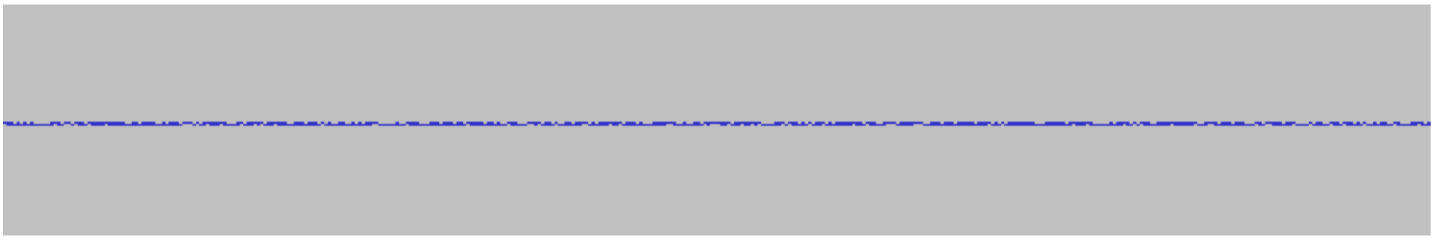
In many cases when monitoring voice quality in real life one receives degraded file from the network containing pauses (silence) before and/or after the actual audio. Let's consider an example received from one of our customers while doing voice quality monitoring in a mobile network. Initial audio is a male voice pronouncing a phrase in English language 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:



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



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 test file one just needs to invoke AQuA with a -trim parameter (sample-06.bat):

```
aqua-wb.exe ./lic/aqua-wb.lic -mode files -src file ./wavs/male.wav  -tstf
            ./wavs/male_5s_delay_5s_end_-36db_whitenoise.wav  -acr auto -npnt auto -
            miter 1 -trim r 5 -ratem %m -fau sample-06-log.txt
```

AQuA output will be:

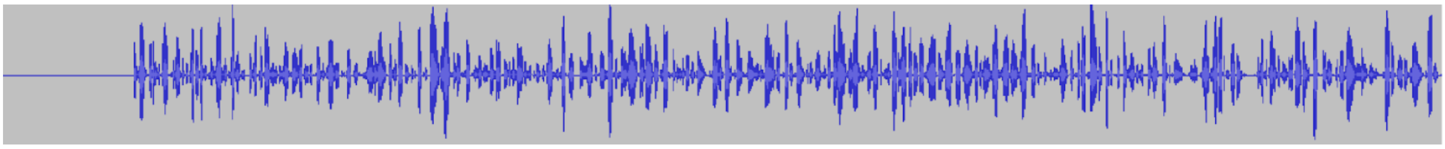
or one can use another option as described above (sample-07.bat):

```
aqua-wb.exe ./lic/aqua-wb.lic -mode files -src file ./wavs/male.wav   -tstf
            ./wavs/male_5s_delay_5s_end_-36db_whitenoise.wav   -acr auto -npnt auto -
            miter 1 -trim a 45 -ratem %m -fau sample-07-log.txt
```

AQuA output will be:

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.0.5.1279.
      Copyright (c) 2021 by Sevana Estonia. All rights reserved.
      --------------------------------------------------------------
      Internal Sevana Ou license
      --------------------------------------------------------------
      Today is 22/06/19. This program is licensed to work until 22/12/31
      --------------------------------------------------------------
      File Quality is
      Percent value   73.23
      MOS value       3.91
```

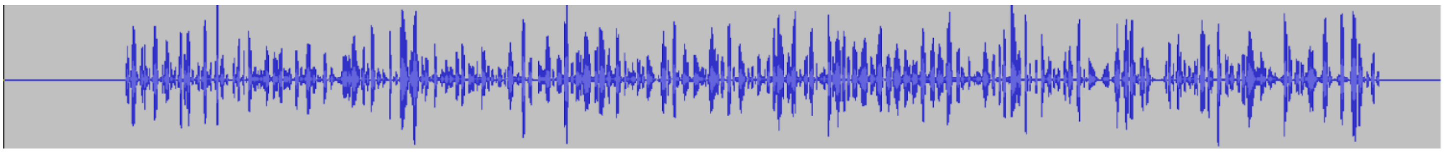
However, in order to be absolutely sure that the trimming works properly let's test it with an artificially created file containing silence (sample-08.bat):



```
aqua-wb.exe ./lic/aqua-wb.lic -mode files -src file ./wavs/male.wav   -tstf
            ./wavs/male_5s_delay_5s_beginning.wav  -acr auto -npnt auto -miter 1 -trim
            a 45 -ratem %m -fau sample-08-log.txt
```

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.0.5.1279.
      Copyright (c) 2021 by Sevana Estonia. All rights reserved.
      --------------------------------------------------------------
      Internal Sevana Ou license
      --------------------------------------------------------------
      Today is 22/06/19. This program is licensed to work until
      22/12/31
      --------------------------------------------------------------
      File Quality is
      Percent value   100.00
      MOS value       5.00
```

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



```
aqua-wb.exe ./lic/aqua-wb.lic -mode files -src file ./wavs/male.wav   -tstf
            ./wavs/male_5s_delay_5s_end.wav  -acr auto -npnt auto –miter 1 -trim a 45 -
            ratem %m -fau sample-09-log.txt
```


Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.0.5.1279.
Copyright (c) 2021 by Sevana Estonia. All rights
reserved.
---------------------------------------------------------------
Internal Sevana Ou
license
---------------------------------------------------------------
Today is 22/06/19. This program is licensed to work until 22/12/31
---------------------------------------------------------------
File Quality is
Percent value   100.00
MOS value       5.00
```


Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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
Average pitch       :     145.12    155.04    Hz
Pitch delta         :     2.44      2.96      Hz
Pitch frequencies distribution distortion  : 19.52.
```

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               :       44.19     48.56     %
```

Newer AQuA versions reports Raw ASR as well. The difference is Raw ASR calculation takes into account all the audio recording; ASR works after audio recording was trimmed.

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 test signal to match with RMS of the reference signal.

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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.:

```
Value Name                Source     Degraded   Units

RMS classic         :     0.0010     0.0014

RMS bounded         :     0.0040     0.0068

Is RMSbounded valid:      true true
```

Signal/Noise Ratio (SNR)

These metrics represent SNR both in the reference and test files.

```
Value Name                Source      Degraded     Units

SNR                  :    XX.XX       XX.XX        dB
```

These metrics show the signal/noise ratio of the reference and test 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 reference 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%.

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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.:

```
Value Name             Source      Degraded    Units

Amplitude Clipping :   0.04        0.20        %
```

Delay/Advancing of audio signal activity

This metric represents signal shift in test file compared to the reference 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 reference 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 reference signal and audio under test. If overall spectrums difference is more than 15% than analysis returns the following string:

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

```
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%.

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 7.x has a special parameter to store pairs of spectrum energy in critical bands of reference and test audio to a .csv file (sample-10.bat):

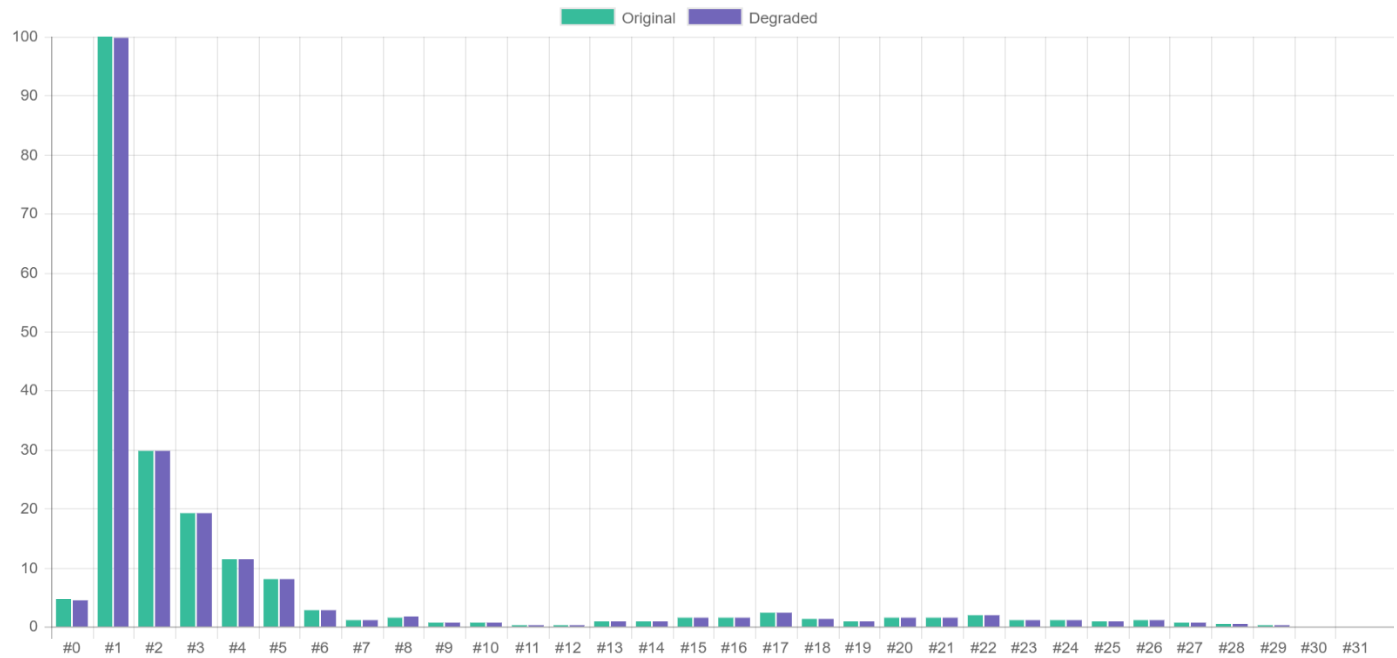
```
aqua-wb.exe ./lic/aqua-wb.lic -mode files -src file ./wavs/male.wav -tstf
            ./wavs/male_5s_delay_5s_end.wav -npnt auto -miter 1 -ratem %p -fau sample-
            10-report.txt -tmc on -gch -psyn off -psyf off -smtnrm on -enorm on -grad
            on -specp 32 sample-10-spect.csv
```

This command produces the following output:

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.0.5.1279.

Copyright (c) 2021 by Sevana Estonia. All rights reserved.

---------------------------------------------------------------

Internal Sevana Ou license

---------------------------------------------------------------

Today is 22/06/20. This program is licensed to work until 22/12/31

---------------------------------------------------------------

File Quality is

        Percent value   95.04

        PESQ value      3.57

Calculating time 0.284000 sec.
```

File 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:

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

As one can see the main reasons are mistiming and delay, which we have not removed, and if we remove it as described in previous chapter (sample-11.bat):

```
Duration distortion.

Audio stretching corresponds to 9.82 percent.


Delay of audio signal activity.

Signal delayed by   4990 ms.


Audio signal activity mistiming (unsynchronization) is 1.05 percent.
```

```
aqua-wb.exe ./lic/aqua-wb.lic -mode files -src file ./wavs/male.wav -tstf ./wavs/male_5s_delay_5s_end.wav -npnt auto -miter 1 -ratem m -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
```

we receive result showing that the files are of identical quality:

```
Sevana Audio Quality Analyzer - AQuA-Wideband v.8.0.5.1279.
      Copyright (c) 2021 by Sevana Estonia. All rights reserved.
      ---------------------------------------------------------------
      Internal Sevana Ou
      license
      ---------------------------------------------------------------
      Today is 22/06/22. This program is licensed to work until 22/12/31
      ---------------------------------------------------------------
      File Quality is
      MOS value       5.00
      Calculating time 0.276000 sec.
```

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 Android 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:

Command line parameters errors

```
Parameters listing error.
Mode error (-mode).
Data reference 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).
```


Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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).

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).

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

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

Data reference 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].

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 test 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!

Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331

```
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.
```

Runtime errors

```
Error opening reference file!
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 reference file) is not supported.
Error: sampling frequency (in test file) is not supported.
Error: Reference sound signal duration is less than 4096 samples.
Error: Test sound signal duration is less than 4096 samples.
Error: After signal alignment reference file duration became less than
          required.
Error: After signal alignment test file duration became less than required.
Error: Reference file sound data is too short.
Error: Test file sound data is too short.
Error: Reference file duration is too short.
Error: Test file duration is too short.
Error: Reference file duration is too short (shorter than 2 seconds).
Error: Test file duration is too short (shorter than 2 seconds).
Error: Reference file manual cut settings failed.
Error: Test file manual cut settings failed.
```


Sevana Oü
Roosikrantsi tn 2
Tallinna linn, Harju maakond, 10119
Estonia
Phone: +3726170331