**Source: Sony Mobile Communications**

**Title: ATeMPO\_SPINE round-robin tests conducted at Sony**

**Document for: Information**

## Agenda Item: 8.5

1. Background

Within the ongoing work item “Acoustic Test methods and Performance Objectives for Speakerphone Performance in Noisy Environments” (ATeMPO\_SPINE), a round-robin test is conducted. A number of commercial mobile phones are circulated between labs to study the suitability of different test setups. The round-robin test plan can be found in [1], from where we quote the objectives of the round-robin test:

“

Questions to answer:

How good is the reproducibility of the ETSI ES 202 396-1 and ETSI TS 103 224 background noise (BGN) simulation methods?

Find out whether the different BGN simulation methods lead to different S-MOS-LQO and N-MOS-LQO results when measuring the same UE.

”

The purpose of the present document is to document how the tests were conducted in the Sony lab and to report on some observations made.

1. Tests at Sony according to the test plan

Tests were conducted at the Sony premises in Lund, Sweden, during week 26, according to plan, with the following observations:

The basis was to measure S-MOS, N-MOS and G-MOS for the ETSI ES 103 106 quality predictor in various conditions.

Background noise:

- 8-speaker method (ETSI TS 103 224)

- 4.1 loudspeaker method (ETSI ES 202 396-1)

- 4.1 loudspeaker method using the same noise scenarios as in TS 103 224, for HHHF

DUT position:

- DUT in front of HATS “hand-held handsfree” (6 noise types plus silence)[[1]](#footnote-2)

- DUT laying on a table “desktop handsfree” (one noise type plus silence)

Bandwidths (UMTS CS call used):

- Narrowband

- Wideband

An “office-type” room (see ETSI ES 202 396-1) in the audio lab was used. It has the specified ceiling absorbers and thin carpet. The walls have diffusing and absorbing objects, resulting in a RT60 reverberation time of 139 ms and clarity C80 of 30 dB.

HEAD acoustics ACQUA with the HAE-BGN and 3PASS background noise systems were used. Loudspeakers were Numark delivered by HEAD acoustics. The test sequences were provided by HEAD acoustics.

A B&K 4128 HATS was used.

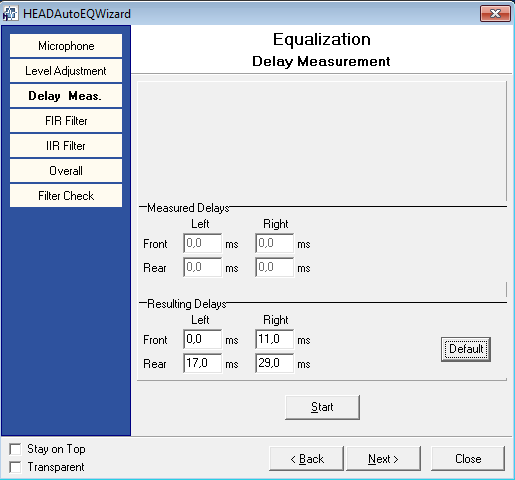
**Procedure**

The mouth simulator of the HATS was calibrated at MRP (90 to 11000 Hz) using a ¼-inch pressure-field microphone.

The HFRP calibration was performed for the two different measurement distances, 30 and 50 cm.

The HATS ears were calibrated.

Equalization of the 4.1 speaker system, standard delays were used:



Equalization of the 8 speaker system for hand-held handsfree.

Measurements of DUT:s in HHHF mode

Equalization of the 8 speaker system for desktop handsfree.

Measurements of DUT:s in DTHF mode

1. Additional tests (beyond the test plan in [1])

As programmed by HEAD acoustics, the noise spectrum was captured from the reference microphone and from the output of the system simulator. This makes it possible to assess what the actual input to and output from the DUT, for diagnostic purposes.

For desktop handsfree, an additional angle was tested (DUT rotated 90 degrees to simulate that more than one talker is of interest, especially when placing the DUT on a table).

For wideband speech:

* RLR for the hand-held handsfree mode was measured using the procedures of 26.132
* Recording of the receiving direction with applied background noise

For the latter test, the Dynastat speech file used the sending MOS-LQO tests was fed in the receiving direction at active speech level -16 dBm0. The receiving direction was recorded in hand-held handsfree mode using both ears of the HATS with free-field correction. The background noise was calibrated using TS 103 224 procedures, with the microphone array mounted on HATS.

The recordings are binaural recordings where the background noise and the mobile phone micro-speaker can be presented to subjects, providing realistic spatial cues.

The purpose was to give additional insight to which of the noise types that are relevant use cases for the hand-held handsfree mode, and which are too noisy to be used in practice.

1. Physical test setup for the tests in the sending direction

**HATS height**

HRP 95 cm above the floor

**Speaker height**

Speakers 1-4: top edge 122 cm, lower edge 98 cm. Speakers 5-8 : top edge 92 cm, lower edge 68 cm.

**Placement**

The loudspeakers were placed on a 1 m grid, forming a 2 x 2 m square. As seen above, speakers 5, 6, 7 and 8 were 30 cm lower in height. The HFRP of the HATS was placed approximately mid-way between the upper and lower speakers.

HATS was in most cases rotated ~22.5 deg, to avoid obstructing the sound from speaker 7, as indicated in the following pictures.

**Equalization of the 8-speaker method (TS 103 224) for the hand-held handsfree mode**

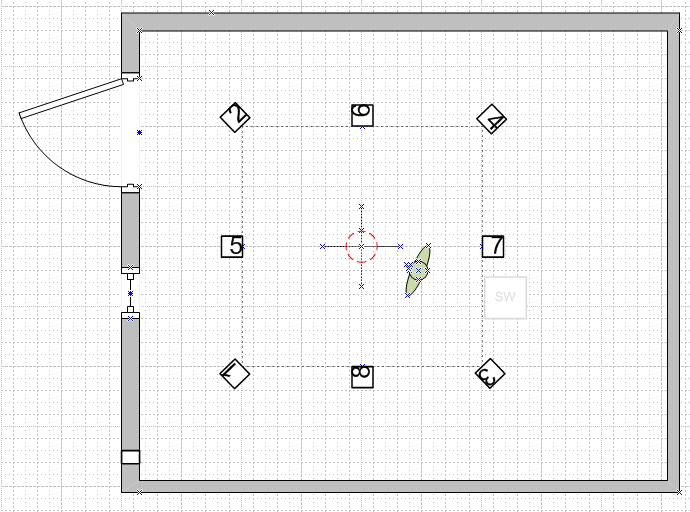


Figure 1 Setup for equalizing the 8-speaker setup for HHHF. The red circle indicates the 8-mic array. HATS is present but not active during this calibration.

**Equalization of the 8-speaker method (TS 103 244) for the desktop handsfree mode**

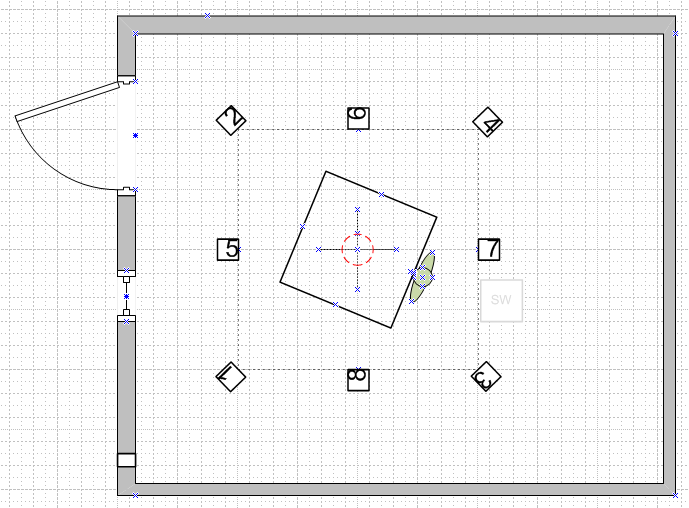


Figure 2 Setup for equalizing the 8-speaker setup for DTHF. The square indicates the 1x1 m table used. The red circle indicates the 8-mic array. HATS is present but not active during this calibration.

**Equalization of the 4.1 speaker method (ES 202 396-1)**

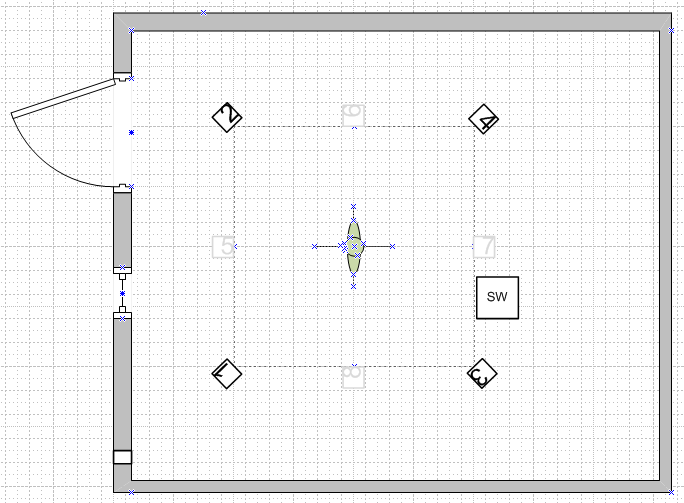


Figure 3 This equalization is used for both HHF and DTHF tests. The HATS ears are used for the equalization. After the equalization, the HATS is moved from the center of the setup to the “speaking position” shown in Figure 4

**Measurements in hand-held handsfree mode**

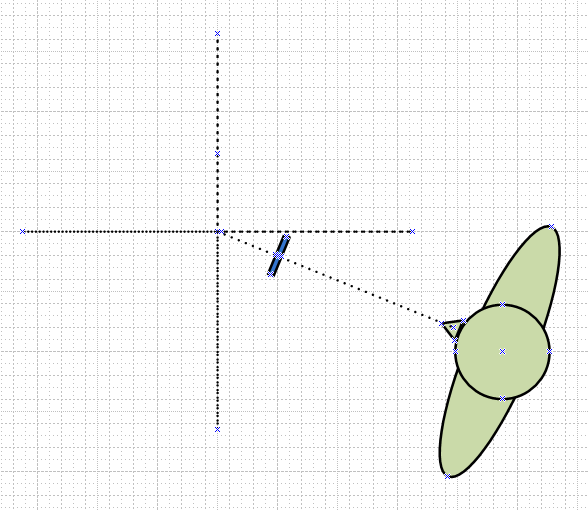


Figure 4 Measurements in HHHF mode using both background noise methods. The UE is shown in blue

**Measurements in desktop handsfree mode**

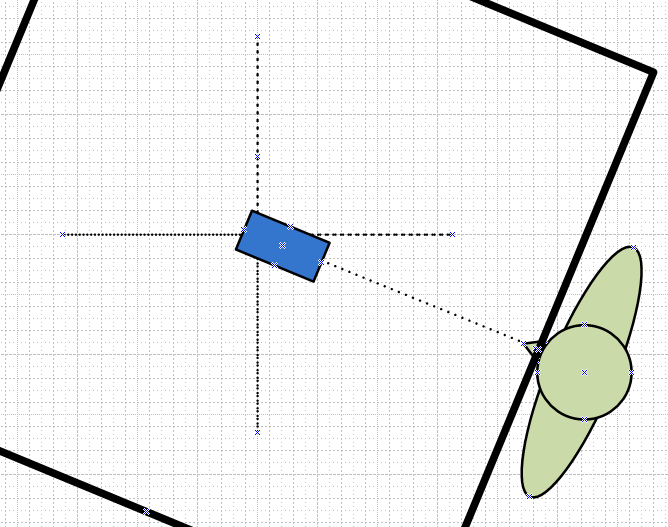


Figure 5 Measurements in DTHF mode using both background noise methods. The UE is shown in blue.

1. Equalization results
   1. Hand-held handsfree equalization results (8 speaker method)

Diagrams of validation results

|  |  |
| --- | --- |
| Calibration Position | Fine tuning position |
| Frequency Response I | 50 Hz to 10000 Hz |
|  |  |
| Frequency Response II | 10000 Hz to 16000 Hz |
|  |  |
| Average Frequency Response | 50 Hz to 20000 Hz |
|  |  |
| Mag. of Complex Coherence | 100 Hz to 1000 Hz |
|  |  |
| Phase of Complex Coherence I | 100 Hz to 1000 Hz |
|  |  |
| Phase of Complex Coherence II | 1000 Hz to 1500 Hz |
|  |  |

* 1. Desktop handsfree equalization results (8 speaker method)

Diagrams of validation results

|  |  |
| --- | --- |
| Calibration Position | Fine tuning position |
| Frequency Response I | 50 Hz to 10000 Hz |
|  |  |
| Frequency Response II | 10000 Hz to 16000 Hz |
|  |  |
| Average Frequency Response | 50 Hz to 20000 Hz |
|  |  |
| Mag. of Complex Coherence | 100 Hz to 1000 Hz |
|  |  |
| Phase of Complex Coherence I | 100 Hz to 1000 Hz |
|  |  |
| Phase of Complex Coherence II | 1000 Hz to 1500 Hz |
|  |  |

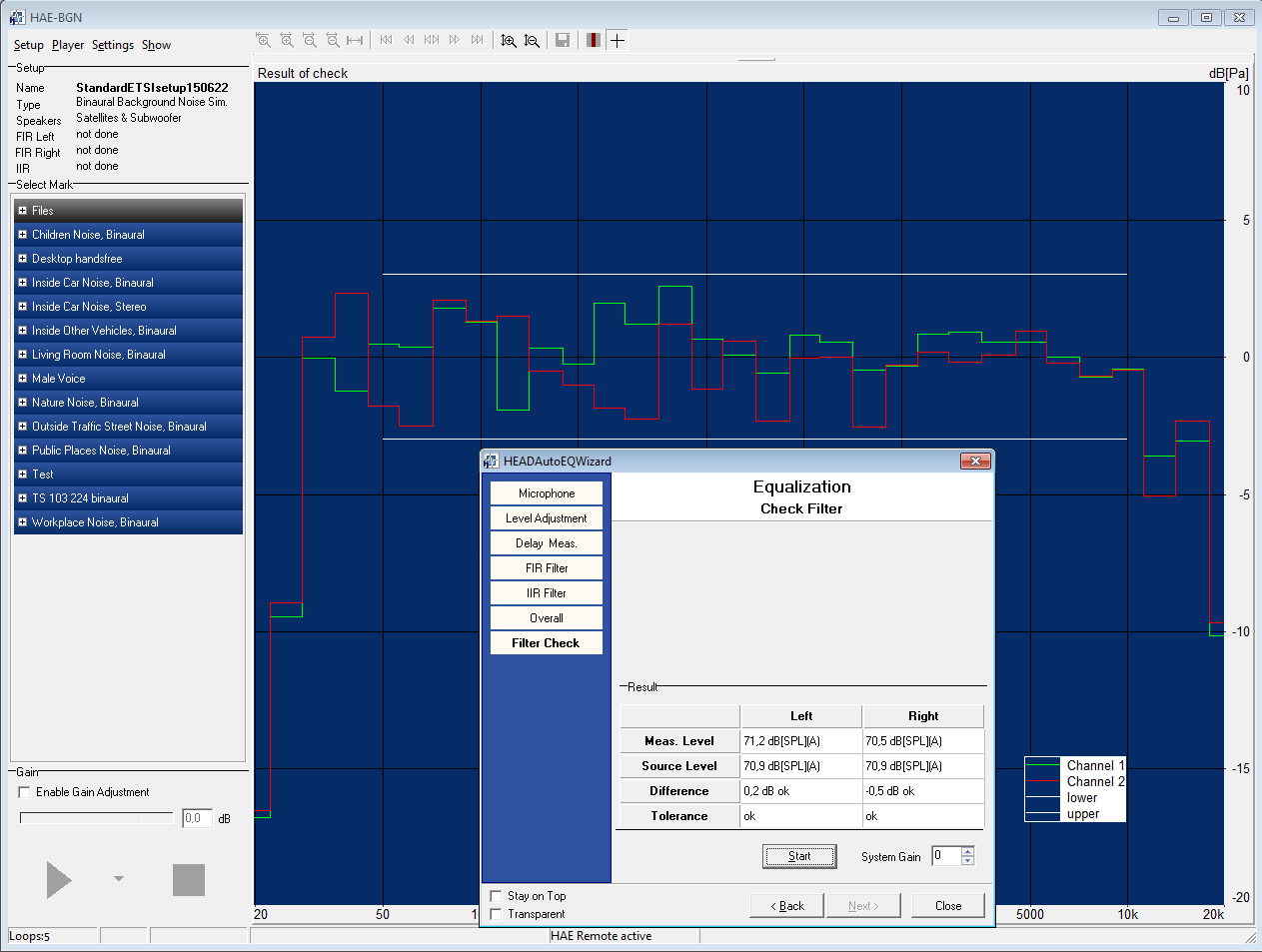
* 1. Head-mounted array equalization results (8 speaker method)

(For the additional recordings in the receiving direction.)

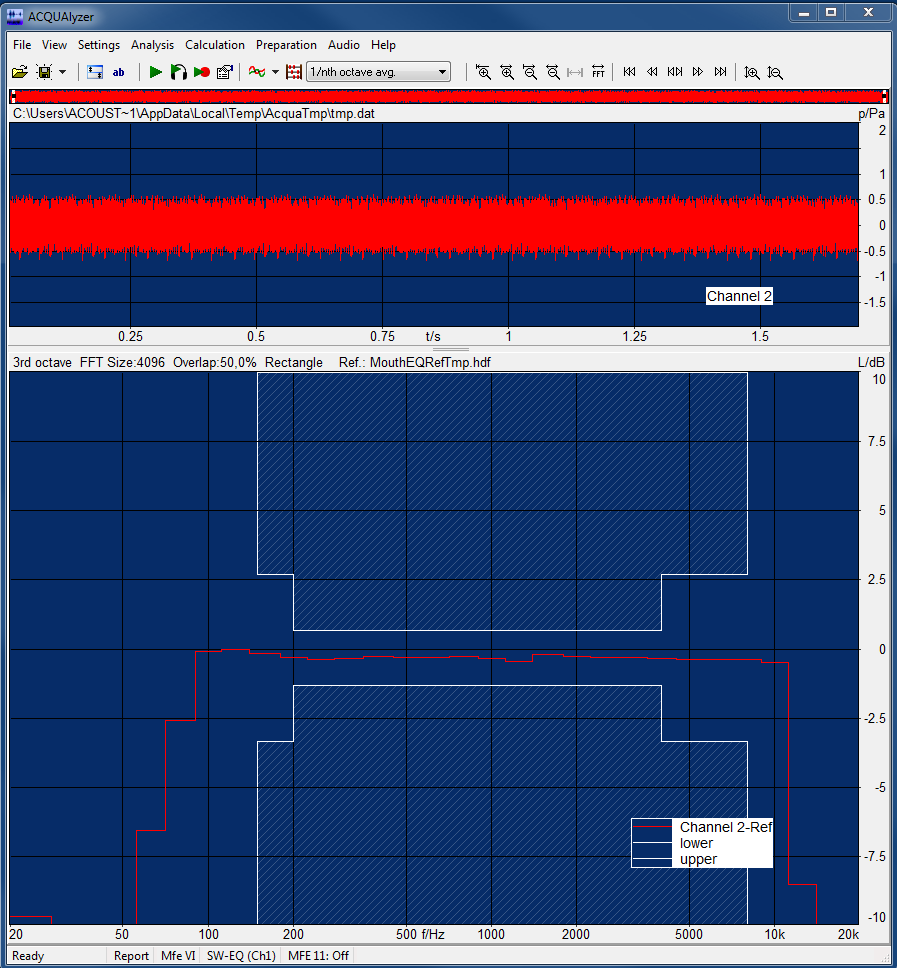
Diagrams of validation results

|  |  |
| --- | --- |
| Calibration Position | Fine tuning position |
| Frequency Response I | 50 Hz to 10000 Hz |
|  |  |
| Frequency Response II | 10000 Hz to 16000 Hz |
|  |  |
| Average Frequency Response | 50 Hz to 20000 Hz |
|  |  |
| Mag. of Complex Coherence | 100 Hz to 1000 Hz |
|  |  |
| Phase of Complex Coherence I | 100 Hz to 1000 Hz |
|  |  |
| Phase of Complex Coherence II | 1000 Hz to 1500 Hz |
|  |  |

* 1. Equalization results for the 4.1 speaker method (ES 202 396-1)



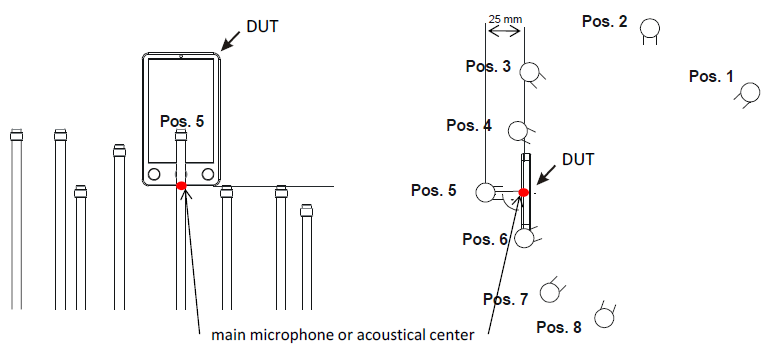
* 1. Mouth equalization results



1. Additional positioning info for the ES 103 224 method

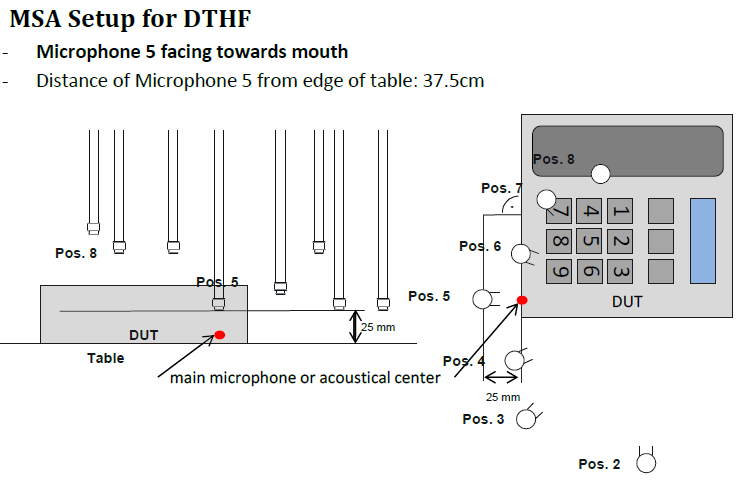
The positioning was made to replicate what was done at the HEAD acoustics lab.

* 1. Hand-held handsfree



Distance from Pos.5 to HATS lips: 27.5 cm horizontal, 5 cm vertical

* 1. Desk-top handsfree



Distance from Pos.5 to HATS lips: 37.5 cm horizontal, 30 cm vertical (from table top)

1. Test results

The detailed results will be available all labs have completed their measurements. We can however analyze the differences between the background noise methods using the same noise scenarios for the HHHF mode, in our lab.

Almost no offset in results were found between the methods, when using the same noise scenarios.

* For all DUT:s and background noises averaged, the mean difference between methods was between -0.01 MOS to 0.01 MOS, for S-, N- or G-MOS
* For each DUT, the mean difference over noise types was between -0.1 MOS to 0.1 MOS, for S-, N- or G-MOS
* However, one case of N-MOS difference of 0.6 MOS was observed, for one DUT in one noise type.



Figure 6 Comparison of test methods for HHHF for narrowband (MOS values for TS 103 224 minus MOS values for ES 202 396-1 with noise scenarios from TS 102 224). For each noise scenario, the DUTS are plotted in sequence DUT1, DUT2…DUT7. Absolute values ranged from 2.0 to 4.1 (S-MOS), 1.2 to 3.4 (N-MOS) and 1.3 to 3.5 (G-MOS) for the TS 103 224 method.



Figure 7 Comparison of test methods for HHHF for wideband (MOS values for TS 103 224 minus MOS values for ES 202 396-1 with noise scenarios from TS 102 224). For each noise scenario, the DUTS are plotted in sequence DUT1, DUT2…DUT7. Absolute values ranged from 2.3 to 4.2 (S-MOS), 1.7 to 3.5 (N-MOS) and 1.7 to 3.4 (G-MOS) for the TS 103 224 method.

1. Conclusions from the tests

The detailed conclusions about lab-to-lab variability will be drawn when all labs have completed their measurements.

At this point, we can conclude that:

* For desktop handsfree, the equalization was not successful at high frequencies (frequency response at the fine-tuning position), despite several attempts. The influence of the large reflective table surface makes the setup sensitive. (During preparations, successful equalization had been obtained at a slightly different location.)
* Both background noise methods produced almost identical hand-held handsfree results (averaged over noise types).
* A lot of preparations and efforts were required to conduct the tests, it is quite resource intensive, especially if accuracy is a priority.
* Both methods use multiple speakers with highly correlated signals and are thus sensitive to positioning (either small errors in practical lab work or systematic offsets between calibration point and DUT microphone point).
* MOS results averaged over noise scenarios were very similar between the two methods

1. References

[1] Tdoc SQ-AHQ099 Proposed test plan for a Round Robin Test for comparison of background noise simulations – Rev. 1

[2] ETSI ES 202 396-1 Speech quality performance in the presence of background noise; Part 1: Background noise simulation technique and background noise database

[3] ETSI TS 103 224 A sound field reproduction method for terminal testing including a background noise database

1. The test plan [1] prescribes 8 noise types but it was decided offline to reduce to 6 noise types to increase chances of completing the round-robin on time [↑](#footnote-ref-2)