|  |  |
| --- | --- |
| 3GPP TR 26.801 V18.0.0 (2023-12) | |
| Technical Specification | |
| 3rd Generation Partnership Project;  Technical Specification Group Services and System Aspects;  User Equipment (UE) supporting handset mode with  non-traditional earpieces  (Release 18) | |
|  | |
|  |  |
|  | |
| The present document has been developed within the 3rd Generation Partnership Project (3GPP TM) and may be further elaborated for the purposes of 3GPP. The present document has not been subject to any approval process by the 3GPPOrganizational Partners and shall not be implemented. This Specification is provided for future development work within 3GPPonly. The Organizational Partners accept no liability for any use of this Specification. Specifications and Reports for implementation of the 3GPP TM system should be obtained via the 3GPP Organizational Partners' Publications Offices. | |

|  |
| --- |
|  |
| ***3GPP***  Postal address  3GPP support office address  650 Route des Lucioles - Sophia Antipolis  Valbonne - FRANCE  Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16  Internet  https://www.3gpp.org |
| ***Copyright Notification***  No part may be reproduced except as authorized by written permission. The copyright and the foregoing restriction extend to reproduction in all media.  © 2023, 3GPP Organizational Partners (ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, TTC).  All rights reserved.  UMTS™ is a Trade Mark of ETSI registered for the benefit of its members  3GPP™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners LTE™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners  GSM® and the GSM logo are registered and owned by the GSM Association |

Contents

Foreword 5

Introduction 6

1 Scope 7

2 References 7

3 Definitions of terms, symbols and abbreviations 7

3.1 Terms 7

3.2 Symbols 7

3.3 Abbreviations 8

4 Report on studies with UEs featuring non-traditional earpieces 8

4.1 Challenges on conducting objective tests according to 3GPP existing specifications for UEs featuring non-traditional earpieces 8

4.1.1 Handset positioning on head and torso simulator 8

4.1.1.1 Background 8

4.1.1.2 Definition - within the FS\_HaNTE study - of handset position for UEs featuring non-traditional earpieces 8

4.1.1.3 Template structure - within the FS\_HaNTE study - of handset position for UEs featuring non-traditional earpieces 9

4.2 Reports on user studies with UEs featuring Non-traditional earpieces and suitability of existing test equipment 10

4.2.1 On the suitability of HATS for HaNTE measurements 10

4.2.1.1 Introduction 10

4.2.1.2 Description of experimental apparatus 10

4.2.1.2.1 Traditional and Non-Traditional Earpiece prototypes 10

4.2.1.2.2 Playback System 11

4.2.1.3 Description of experimental procedure 11

4.2.1.3.1 Wideband objective measurements on HATS 11

4.2.1.3.2 Band-passed objective measurements on HATS 13

4.2.1.3.3 Subjective loudness matching 13

4.2.1.4 Loudness Matching Results 14

4.2.1.5 Discussion 15

4.2.2 On the impact of handset mounting on HaNTE measurements 16

4.2.2.1 Introduction 16

4.2.2.2 Devices 16

4.2.2.2.1 Traditional and Non-Traditional Earpiece prototypes 16

4.2.2.2.2 Commercial devices 16

4.2.2.3 Measurement system 16

4.2.2.3.1 Device mounting 16

4.2.2.3.2 Playback system 17

4.2.2.3.3 Recording system 17

4.2.2.4 Measurement results 17

4.2.2.5 Discussion 18

4.3 Acoustic measurements according to 3GPP TS 26.131 and TS 26.132 and additional test cases to complement TS 26.131 and TS 26.132 18

4.3.1 Test setup 18

4.3.2 Test results in receiving 19

4.3.2.1 WB RLR 19

4.3.2.2 WB RFR 20

4.3.3 Additional results on user studies and objective measurements 21

4.3.3.1 General 21

4.3.3.2 Additional test cases to complement TS 26.131 and TS 26.132 21

4.3.3.2.1 Out-of-band energy level measurement from SWB RFR measurement 21

4.3.3.2.2 P.863 measurement 24

4.3.3.2.3 WB RFR at maximum volume control 24

4.3.4 Discussion 25

4.4 Acoustic measurements according to 3GPP TS 26.131 and TS 26.132 with various mounting conditions and test equipment 26

4.4.1 Introduction 26

4.4.2 Test Setup 26

4.4.3 Measurement Results 27

4.4.3.1 Device 1 on HATS 1 Measurements 27

4.4.3.1.1 Receiver Loudness Rating Results 27

4.4.3.1.2 Receiver Sensitivity/Frequency Response Results 28

4.4.3.2 Device 1 on HATS 2 Measurements 30

4.4.3.3 Device 2 on HATS 1 Measurements 30

4.4.3.3.1 Receiver Loudness Rating Results 30

4.4.3.3.2 Receiver Sensitivity/Frequency Response Results 31

4.4.3.4 Device 2 on HATS 2 Measurements 32

4.4.4 Discussion 32

4.5 HaNTE round robin test in different labs 33

4.5.1 Introduction 33

4.5.2 Devices 33

4.5.3 Test plan 33

4.5.3.1 Overview 33

4.5.3.2 Test 1: Speech Quality 34

4.5.3.3 Test 2: Privacy 35

4.5.3.4 Test 3a: Robustness (variation of ECRP) 35

4.5.5.5 Test 3b: Robustness (variation of fork position) 36

4.5.4 Results 37

4.5.4.1 Preparations 37

4.5.4.2 Test 1: Speech Quality 42

4.5.4.3 Test 2: Privacy 43

4.5.4.4 Test 3a: Robustness (variation of ECRP) 46

4.5.4.5 Test 3b: Robustness (variation of fork position) 51

4.5.5 Discussion 54

5 Conclusion 55

Annex A: Tables for the study on suitability of HATS for HaNTE measurements 57

Annex B: Change history 60

# Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 or greater indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the document.

In the present document, certain modal verbs have the following meanings:

**shall** indicates a mandatory requirement to do something

**shall not** indicates an interdiction (prohibition) to do something

NOTE 1: The constructions "shall" and "shall not" are confined to the context of normative provisions, and do not appear in Technical Reports.

NOTE 2: The constructions "must" and "must not" are not used as substitutes for "shall" and "shall not". Their use is avoided insofar as possible, and they are not used in a normative context except in a direct citation from an external, referenced, non-3GPP document, or so as to maintain continuity of style when extending or modifying the provisions of such a referenced document.

**should** indicates a recommendation to do something

**should not** indicates a recommendation not to do something

**may** indicates permission to do something

**need not** indicates permission not to do something

NOTE 3: The construction "may not" is ambiguous and is not used in normative elements. The unambiguous constructions "might not" or "shall not" are used instead, depending upon the meaning intended.

**can** indicates that something is possible

**cannot** indicates that something is impossible

NOTE 4: The constructions "can" and "cannot" shall not to be used as substitutes for "may" and "need not".

**will** indicates that something is certain or expected to happen as a result of action taken by an agency the behaviour of which is outside the scope of the present document

**will not** indicates that something is certain or expected not to happen as a result of action taken by an agency the behaviour of which is outside the scope of the present document

**might** indicates a likelihood that something will happen as a result of action taken by some agency the behaviour of which is outside the scope of the present document

**might not** indicates a likelihood that something will not happen as a result of action taken by some agency the behaviour of which is outside the scope of the present document

In addition:

**is** (or any other verb in the indicative mood) indicates a statement of fact

**is not** (or any other negative verb in the indicative mood) indicates a statement of fact

NOTE 5: The constructions "is" and "is not" do not indicate requirements.

# Introduction

The present document investigates on the suitability of 3GPP existing specifications for testing UEs featuring non-traditional earpieces.

UEs featuring non-traditional earpieces may pose challenges for handset mode acoustic testing. For example, a UE may exclusively use a vibrating display to produce sound when operating in handset mode, offering no clearly identifiable centre of an earpiece to position the headset for testing. Additionally, such UE could have its acoustic response affected by the choice of handset positioner mechanism.

The present document documents studies conducted and concludes with a summary of challenges identified and gap analysis of existing 3GPP specifications.

# 1 Scope

The present document reports on investigations of testing UEs featuring non-traditional earpieces, and identifies related gaps to existing 3GPP specifications and recommended test equipment.

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] ITU-T Recommendation P.57 (12/2011): "Artificial ears".

[3] ITU-T Recommendation P.58 (05/2013): "Head and torso simulator for telephonometry".

[4] ITU-T Recommendation P.64 (06/2019): "Determination of sensitivity/frequency characteristics of local telephone Systems".

[5] 3GPP TS 26.131: "Terminal acoustic characteristics for telephony; Requirements".

[6] 3GPP TS 26.132: "Speech and video telephony terminal acoustic test specification".

[7] ITU-T Recommendation P.79 (06/2007): "Determination of sensitivity/frequency characteristics of local telephone Systems".

[8] ITU-T Recommendation P.863 (09/2014): "Perceptual objective listening quality assessment".

[9] ITU-T Recommendation P.501 (05/2020): "Test signals for use in telephony and other speech-based applications".

[10] ITU-T Recommendation P.56 (12/2011): "Objective measurement of active speech level".

[11] ISO 532-1:2017: "Acoustics - Methods for calculating loudness - Part 1: Zwicker method".

[12] ITU-T Recommendation P.700 (06/2021): "Calculation of loudness for speech communication".

# 3 Definitions of terms, symbols and abbreviations

## 3.1 Terms

For the purposes of the present document, the terms given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1].

## 3.2 Symbols

Void.

## 3.3 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

ASL Active Speech Level

ECRP Ear Cap Reference Point

HaNTE Handsets featuring Non-Traditional Handsets

MAX Maximum volume setting

MECRP Manufacturer-defined Ear Cap Reference Point

NOM Nominal volume setting

RFR Receive Frequency Response

RLR Receive Loudness Rating

# 4 Report on studies with UEs featuring non-traditional earpieces

## 4.1 Challenges on conducting objective tests according to 3GPP existing specifications for UEs featuring non-traditional earpieces

### 4.1.1 Handset positioning on head and torso simulator

#### 4.1.1.1 Background

TS 26.132 [6] references ITU-T Recommendation P.64 [4] regarding handset position on artificial ears. However, UEs featuring non-traditional earpieces may not feature a centre of an acoustic port raising the question of how to properly position a handset for testing. Therefore, an update to P.64 would be required for SA4 to reference this Recommendation also for handsets with non-traditional earpieces, if the present study is followed by normative work in 3GPP.

Such update to P.64 may describe how to specify a reference point for handsets where there are no acoustic outlets in the earpiece area of the handset.

This sub-clause suggests an alternative method for positioning the handset that is suitable for UEs featuring non-traditional earpieces.

#### 4.1.1.2 Definition - within the FS\_HaNTE study - of handset position for UEs featuring non-traditional earpieces

Within the context of FS\_HaNTE, the MECRP is defined in terms of distance from the upper edge of the handset (this is equivalent what is sometimes referred to as "end stop" for positioning fixtures) and the distance from a centre/symmetry line.



Figure 4-1: Definition of MECRP for handsets in FS\_HaNTE

*ye* is allowed to vary significantly to cover also the following case, or more unusual shapes:



Figure 4-2: Example of a handset shape where, due to a curvature,   
some extra allowance of ye is needed.

The acceptable range of offset of the MECRP should not exceed:

- No limit of ye value.

- ±10 mm along unit vector ze.

#### 4.1.1.3 Template structure - within the FS\_HaNTE study - of handset position for UEs featuring non-traditional earpieces

A template for providing handset position for UEs featuring non-traditional earpieces is provided in Tables 4-1 to 4-3.

Table 4-1: MECRP

|  |  |
| --- | --- |
| Axis | [mm] |
| ye |  |
| ze |  |

Table 4-2: Angle Settings

|  |  |
| --- | --- |
| Angle | Delta from standard angle [°] |
| A |  |
| B |  |
| C |  |

Table 4-3: Application Force

|  |  |
| --- | --- |
| Application force [N] |  |

See ITU-T P.64 for details on standard angle.

## 4.2 Reports on user studies with UEs featuring Non-traditional earpieces and suitability of existing test equipment

### 4.2.1 On the suitability of HATS for HaNTE measurements

#### 4.2.1.1 Introduction

3GPP UE's Receive Frequency Response (RFR) and Receive Loudness Rating (RLR) are measured with Head and Torso Simulators (HATS). With traditional earpiece designs, the perceived loudness comes from acoustic radiation into the ear canal. With non-traditional earpieces, such as vibrating displays, sound may be transmitted to the user by other mechanisms (e.g. tissue conduction). Because HATS primarily measure acoustic radiation through the ear canal, one question is whether HATS is suitable for measurement of handsets featuring non-traditional earpieces.

To compare objective and subjective assessments of loudness on devices featuring a traditional and non-traditional earpiece (vibrating display), an experimental methodology was developed by the source.

#### 4.2.1.2 Description of experimental apparatus

##### 4.2.1.2.1 Traditional and Non-Traditional Earpiece prototypes

Two handset prototypes were developed for this experiment. Prototype 1 incorporates a 27ohm 15x6 mm earpiece (traditional earpiece) radiating sound through an acoustic port on the display, and Prototype 2 incorporates an 8ohm 15x6mm actuator directly attached to the prototype display. Apart from this difference, both prototypes are identical. See Figure 4-3. Improved prototype acoustic sealing was achieved through an added layer of material at the bottom of the prototype screen and putty for screen fastening.



Figure 4-3: Prototypes produced (one with traditional earpiece, other with vibrating display)

Handsets were positioned on two HATS models for objective frequency response measurement. In both cases, the prototypes were positioned according to ITU-T Rec. P.64 (A = 21.2°, B = 12.9°, C = 2.3°) [4]. HATS 1 was a Bruel & Kjaer Type 5128 with Type 4606 Handset Positioner and Centering Fork UA-1537. HATS 2 was a Head Acoustics HMS II.3 artificial head with HHP III handset positions system. In both cases, the prototypes were mounted with End-Stop adjustment set to 17mm (MECRP defined as ye = 17mm, yz = 0mm). See Figure 4-4.



Figure 4-4: Handset positioned on HATS 1 and HATS 2

NOTE: The Bruel & Kjaer Type 5128 HATS does not comply with the current ITU-T Recommendations P.58 [3] and P.64 [4].

##### 4.2.1.2.2 Playback System

To drive Prototype 1 and Prototype 2, a personal computer (PC) was connected to an RME MADIface Pro USB digital audio interface. Prototype 1 was directly driven by the MADIface headphone output (Analog 3). Prototype 2 was driven by a Bruel & Kjaer Type 2716C amplifier with +12dB of gain which was in turn connected to the MADIface Analog 1 output.

For wideband objective measurements in clause 4.2.1.3.1 and band-passed level adjustment in clause 4.2.1.3.2, Adobe Audition was used for digital audio playback. For level adjusted measurements in clause 4.2.1.3.2 and the subjective loudness matching in clause 4.2.1.3.3, Max/MSP was used for digital audio playback. A Griffin Technology PowerMate USB Volume Knob was used to allow control of the playback level during subjective loudness matching. See Figure 4-5 for the block diagram.

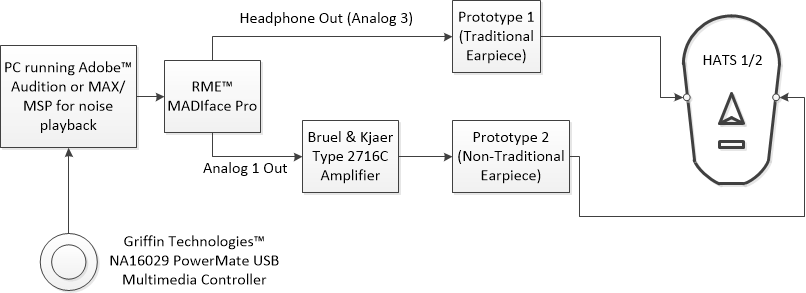


Figure 4-5: Block diagram of test apparatus for the experiment

#### 4.2.1.3 Description of experimental procedure

##### 4.2.1.3.1 Wideband objective measurements on HATS

The wideband pink noise responses of Prototypes 1 and 2 were objectively measured on HATS 1 and 2, using the handset positioner settings described in clause 4.2.1.2.1. The force was adjusted to 2N, 8N, 13N to span a range of mounting conditions. The measurements were accomplished by playing -18dBFS pink noise through the playback system described in clause 4.2.1.2.2 (using Adobe Audition for digital audio playback). The system playback gain was adjusted such that the acoustic sound pressure level was 79dBSPL at the Ear Reference Position (ERP) for Prototype 1 (traditional earpiece) mounted on HATS 1 (Bruel & Kjaer Type 5128) with a force of 8N. The same system gain was then used for Prototype 2 (non-traditional earpiece), HATS 2 (Head Acoustics HMS II.3), and all other mounting forces (2N and 13N). Figures 4-6a and 4-6b show the ERP level measurements in 3rd octave bands for all Prototype, HATS, and force conditions.

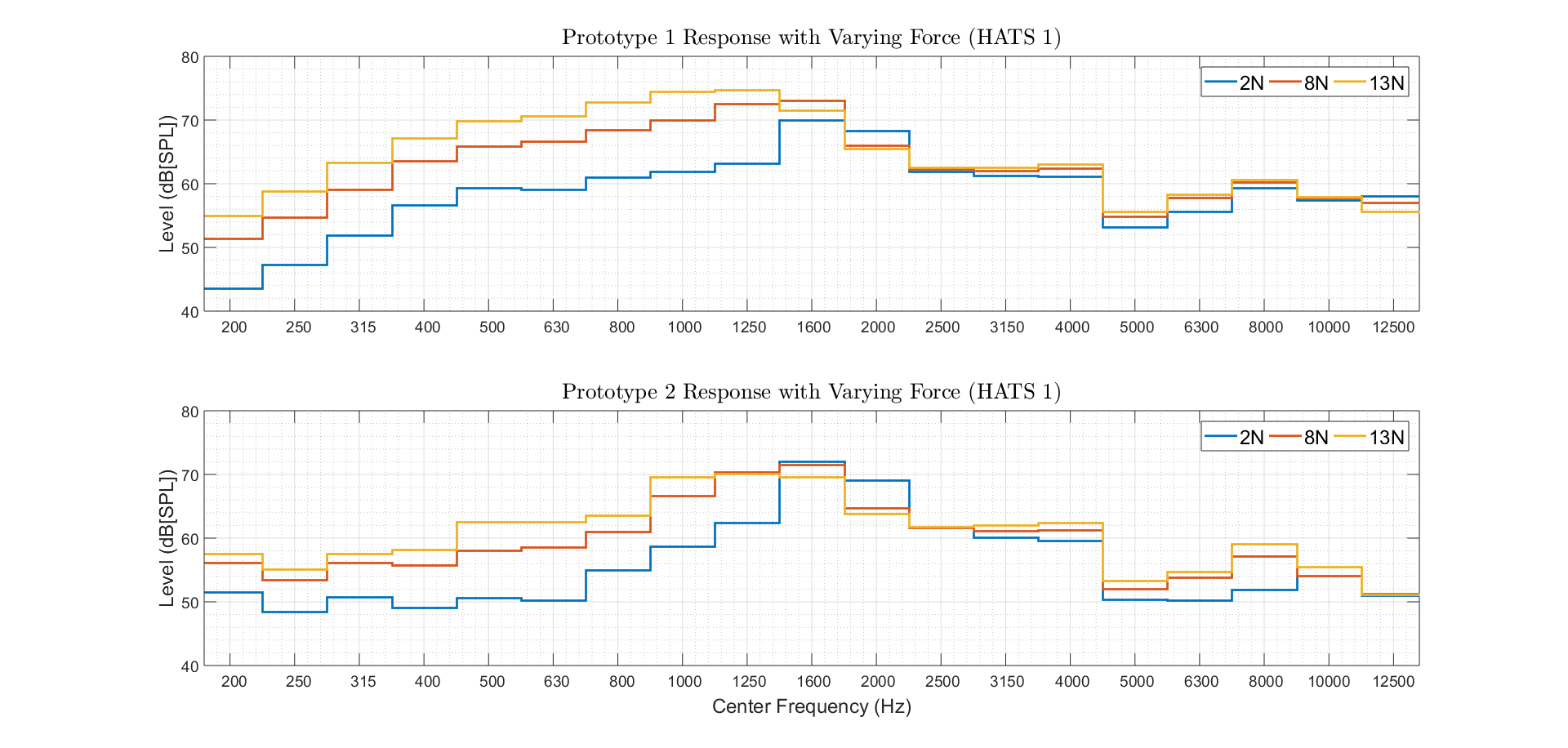


Figure 4-6a: Earpiece level (Prototype 1) and Actuator level (Prototype 2) level on HATS 1

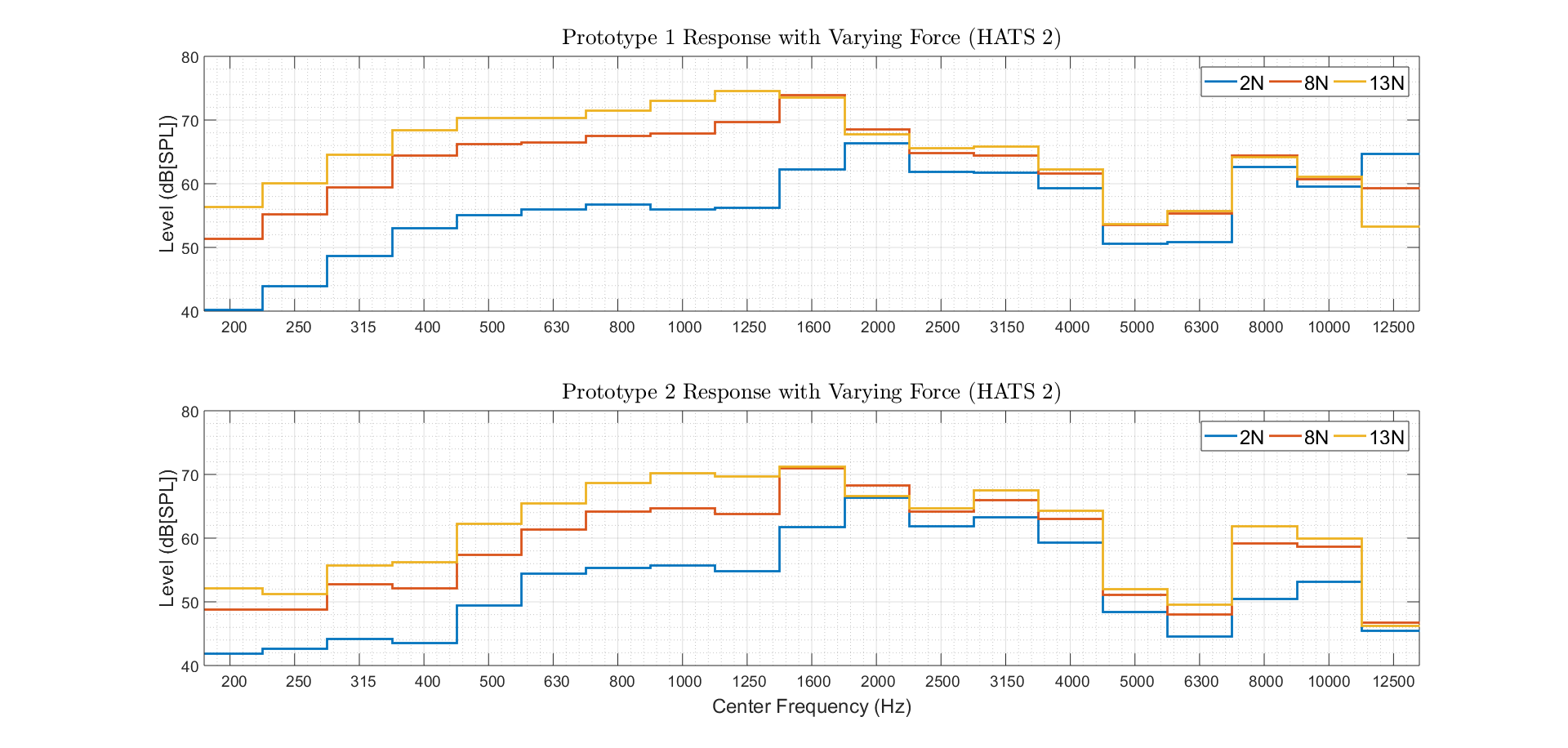


Figure 4-6b: Earpiece level (Prototype 1) and Actuator level (Prototype 2) level on HATS 2

Attenuation at low frequencies (below 1.6kHz for HATS 1 and below 2kHz for HATS 2) is seen for both HATS and both prototypes as the application force is decreased from 13N to 2N. Table 4-4 presents the mean attenuation (200Hz-1.6kHz) between 13N and 2N of application force for all conditions. For both HATS, Prototype 1 demonstrates more low frequency leakage than Prototype 2 (~3dB). Furthermore, HATS 2 demonstrates more low frequency loss than HATS 1 independent of prototype (~4dB).

Table 4-4: Mean attenuation (200Hz-1.6kHz) with varying application force

|  |  |  |
| --- | --- | --- |
| HATS | Prototype | Mean 13N to 2N Attenuation |
| HATS 1 | Earpiece (Prototype 1) | -10.57dB |
| HATS 1 | Actuator (Prototype 2) | -8.11dB |
| HATS 2 | Earpiece (Prototype 1) | -15.48dB |
| HATS 2 | Actuator (Prototype 2) | -11.74dB |

##### 4.2.1.3.2 Band-passed objective measurements on HATS

Band-passed pink noise level calibration was accomplished by playing 1/3rd octave band-passed pink noise through the playback system described in clause 4.2.1.2.2 with Adobe Audition. The level of each noise band was adjusted to match roughly 79dBSPL at the Ear Reference Position (ERP) when Prototype 1 (traditional earpiece) was mounted with a force of 8N on HATS 1.

After calibration, band-passed pink noise responses were measured through Max/MSP with both prototypes on both HATS using the level calibrated band-passed pink noise as stimuli. Figure 4-7 shows the ERP level measurements in each 1/3rd octave band for both prototypes on both HATS at 8N of force.

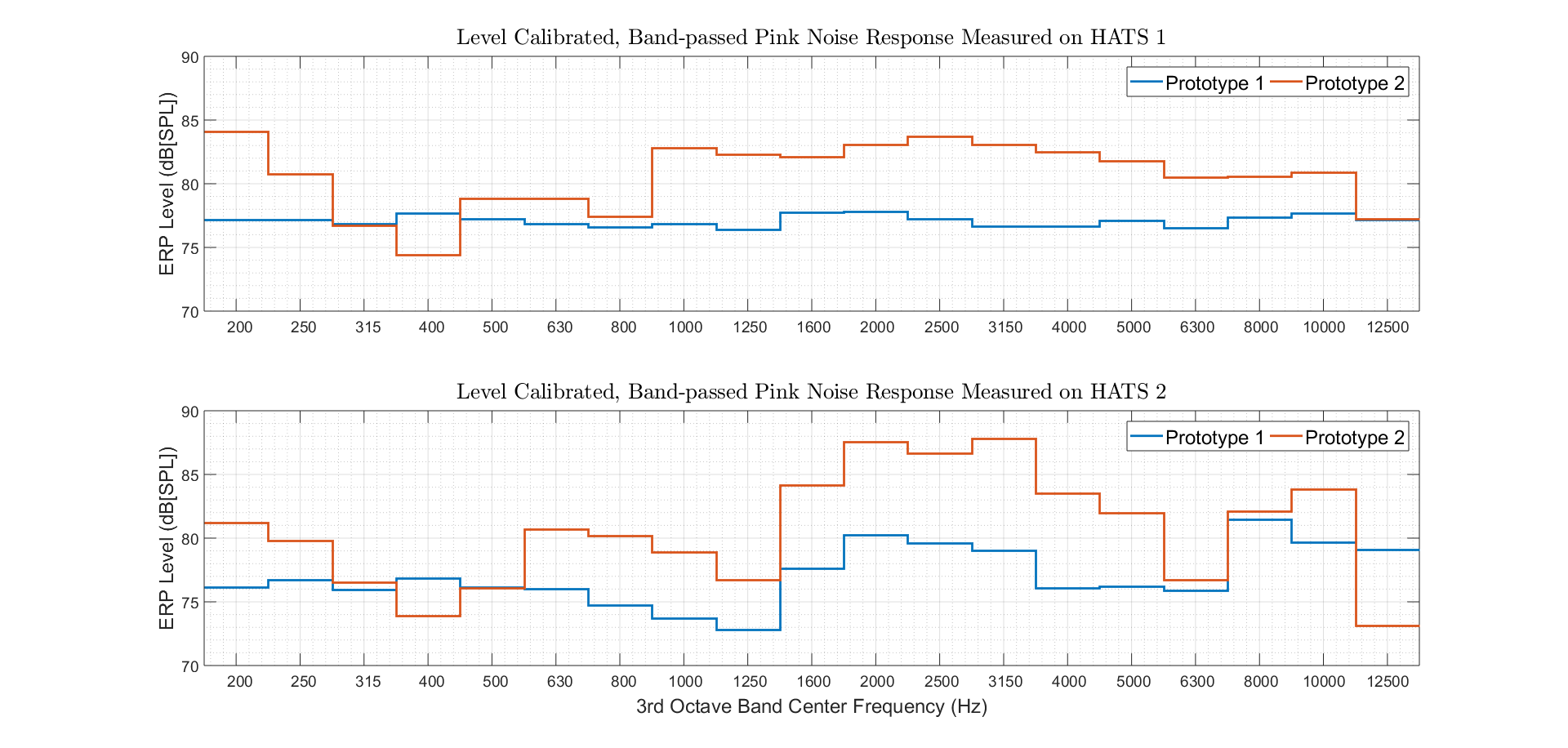


Figure 4-7: Band-passed Earpiece level (Prototype 1) and Actuator level (Prototype 2) on both HATS

Prototype 1 was successfully calibrated to +/-1dB around a nominal level of 77dBSPL for all bands from 200Hz to 12.5kHz on HATS 1. The -2dB difference between the target level of 79dBSPL and achieved nominal level of 77dBSPL may be due to a software switch between Audition and Max/MSP for digital audio playback.

The relative level between Prototype 1 and Prototype 2 for each 3rd octave band is used as the objective transfer function in the following section. Figure 4-10 includes the transfer functions () computed for both HATS, where and are the 3rd octave band responses of Prototype 1 and 2, respectively.



##### 4.2.1.3.3 Subjective loudness matching

For the subjective loudness matching, eight participants were tasked with comparing the perceived loudness of Prototypes 1 and 2 for each level adjusted, band-passed pink noise stimulus. Participants were asked to hold both handsets in a comfortable position. The test administrator ensured consistent positioning throughout the test. Figure 4-8 demonstrates a standard positioning for one subject.



Figure 4-8: Indicative handset positioning during loudness matching

Each participant completed two sessions. In each session, nineteen 3rd octave bands (200Hz-12.5kHz) were assessed. Between each session the side where Prototype 2 was held (left or right) was reversed. In each trial, participants were asked to adjust the level of Prototype 2 to match the perceived loudness of Prototype 1 using the Griffin Technology PowerMate USB Volume Knob. The volume knob allowed adjustment in steps of 0.5dB and a range of -24dB to +24dB. The starting gain of Prototype 2 was randomly adjusted between -5dB and +5dB. The order of the frequencies presented was also randomized.

An interface for the listening test was developed in Max/MSP and is shown in Figure 4-9. Playback switching between the two prototypes (left and right) was signalled by the subject (verbally or through gestures) and carried out by the test administrator. All tests were completed in an anechoic chamber.

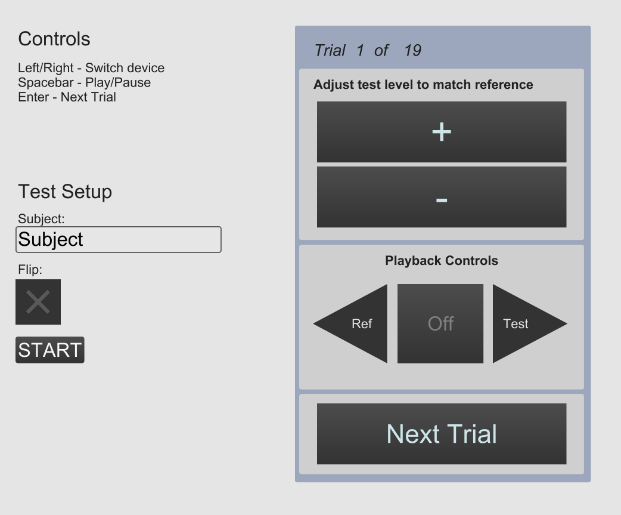


Figure 4-9: Graphical User Interface for listening test

#### 4.2.1.4 Loudness Matching Results

Figure 4-10 shows the results of the listening tests (i.e. the gain that will be applied to Prototype 2 to match Prototype 1) including mean and 95% confidence intervals, along with the equivalent objective measurement on both HATS at 8N. A full table of descriptive statistics can be found in Table A-1 of Annex A.

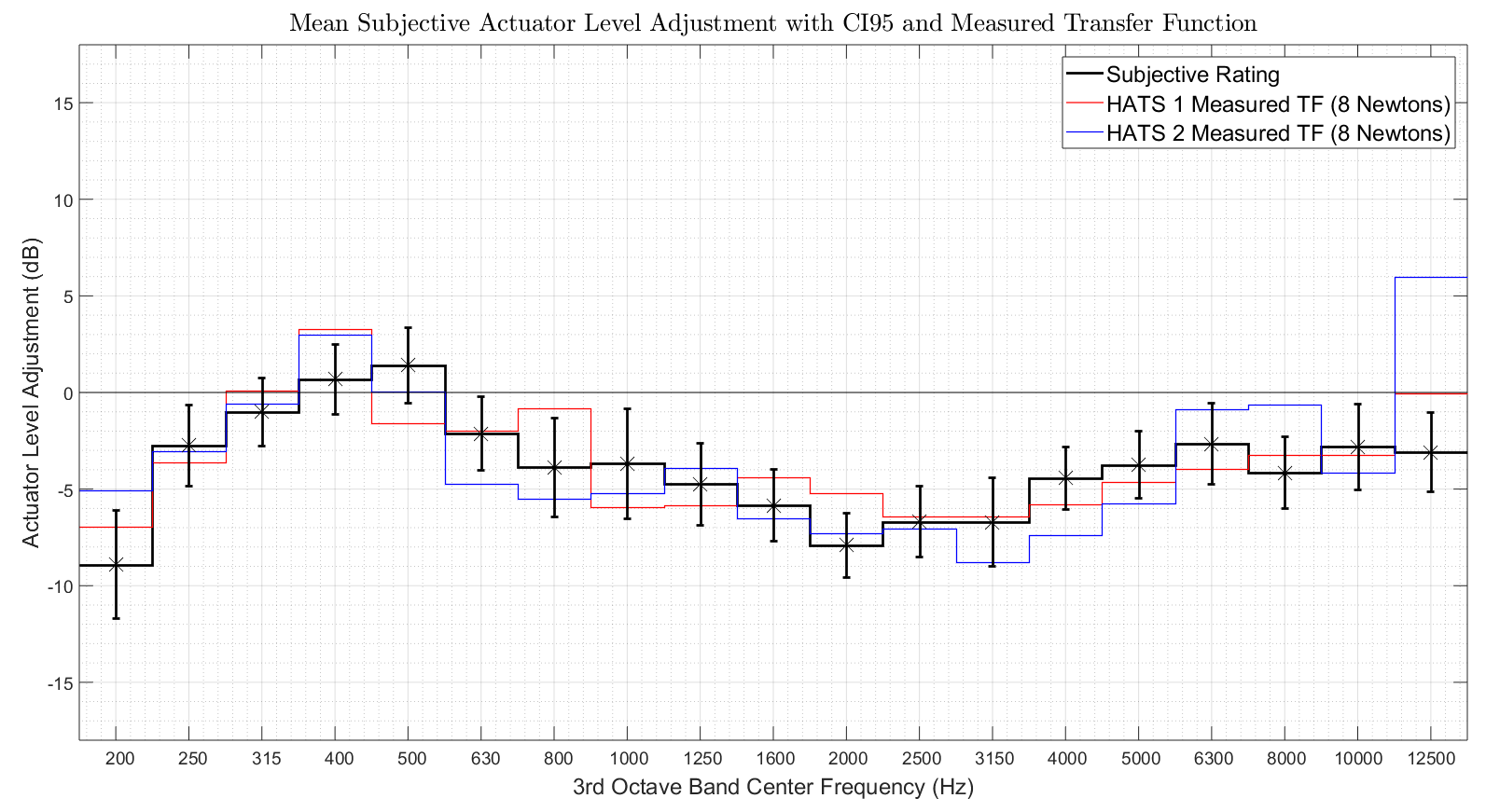


Figure 4-10: Results of subjective loudness matching and objective transfer functions from HATS

Generally strong agreement exists between the measured transfer function between earpiece and actuator prototypes on HATS and the mean subjective actuator level adjustment. There is a strong positive correlation between subjective level matching and the HATS 1 transfer function (ρ=0.775; p<0.000) and a strong (although slightly lower) positive correlation between subjective level matching and the HATS 2 transfer function (ρ=0.656; p<0.001). Table A-2 of Annex A provides more details on the correlation analysis.

Table A-3 of Annex A presents squared error (SE) per 3rd octave band and wideband root mean squared error (RMSE) calculations comparing the subjective level matching and the two HATS generated transfer functions. The RMSE between subjective results and HATS 1 measurements (RMSE=1.79dB) is ~1dB lower than the RMSE between subjective results and HATS 2 measurements (RMSE=2.83dB). It is interesting to note that the RMSE is nearly identical for the two HATS below 8kHz. However, at and above 8kHz the HATS 2 RMSE is nearly 4dB higher than that of HATS 1. Table 4-5 presents RMSE data for these two frequency ranges.

Table 4-5: RMSE by frequency range

|  |  |  |
| --- | --- | --- |
| Frequency Range | HATS 1 RMSE (dB) | HATS 2 RMSE (dB) |
| 200Hz – 6.3kHz | 1.783 | 1.866 |
| 8kHz – 12.5kHz | 1.851 | 5.677 |

Subjective loudness matching results for each 3rd octave band were compared to HATS data using a two-tailed, one sample t-test with 15 degrees of freedom (t[15] = 2.1314) and a 95% confidence interval. Full t-test results with HATS 1 as reference are presented in Table A-4 of Annex A. Of the nineteen tested frequency bands, subjective adjustment varies significantly from HATS 1 measurements in five (400Hz, 500Hz, 800Hz, 2kHz, and 12.5kHz). A maximum absolute error of 3.065dB occurs at 800Hz.

Table A-5 of Annex A presents the same t-test methodology with HATS 2 as the reference. Seven of the nineteen frequency bands demonstrate a statistically significant difference between subjective and objective level matching (200Hz, 400Hz, 630Hz, 4kHz, 5kHz, 8kHz, and 12.5kHz). A maximum absolute error of 9.074dB occurs at 12.5kHz.

#### 4.2.1.5 Discussion

The results support the performance of HATS as a measurement tool for the vibrating display prototype in the frequency range 200Hz to 12.5kHz. There exists a significant and strong positive correlation between subjective loudness matching of traditional and non-traditional earpieces and objectively measured transfer functions between the same two prototypes. The positive correlation exists for both HATS models investigated herein. For these specific prototypes, HATS 1 demonstrated less error than HATS 2 (particularly at high frequencies) between the measured transfer function and subjective level matching. Furthermore, the transfer function measured on HATS 1 has fewer frequency bands, which vary significantly from the corresponding subjective level adjustment, than the transfer function measured on HATS 2.

### 4.2.2 On the impact of handset mounting on HaNTE measurements

#### 4.2.2.1 Introduction

3GPP UE's Receive Frequency Response (RFR) and Receive Loudness Rating (RLR) are measured with Head and Torso Simulators (HATS) and handset positioners. Handset positioners exert force at multiple points on the Device Under Test (DUT), including stationary supports on the back of the DUT. These supports may or may not be indicative of a user's hand position while interacting with the DUT.

The receiver transducer of a traditional earpiece design is decoupled from the surrounding mechanical system and will not be greatly affected by the existence and placement of physical supports. However, with non-traditional earpiece designs (e.g. a vibrating display) the receiver transducer may be coupled with the entire mechanical structure of the device. Therefore, it is important to consider any physical supports and the force they exert when mounting non-traditional earpiece devices for RFR and RLR testing. This contribution explores mounting induced RFR variability on devices with both traditional and non-traditional earpieces.

#### 4.2.2.2 Devices

##### 4.2.2.2.1 Traditional and Non-Traditional Earpiece prototypes

See clause 4.2.1.2.1 for a detailed description of the prototypes used in this study.

##### 4.2.2.2.2 Commercial devices

Five commercial devices were tested along with the two prototypes mentioned in the previous section. Four of the five commercial devices employ non-standard, vibrating display earpieces, while one device has as traditional earpiece. Table 4-6 provides more information.

Table 4-6: Further information on commercial devices

|  |  |
| --- | --- |
| Device Number | Earpiece |
| 1 | Traditional |
| 2 | Vibrating Display (HaNTE) |
| 3 | Vibrating Display (HaNTE) |
| 4 | Vibrating Display (HaNTE) |
| 5 | Vibrating Display (HaNTE) |

#### 4.2.2.3 Measurement system

##### 4.2.2.3.1 Device mounting

For RFR measurement, handsets were positioned on a Bruel & Kjaer Type 5128 HATS as described in clause 4.2.1.2.1.

For a comparison of mounting conditions, all devices were measured with and without the handset positioner support pins in place. See Figure 4-11.

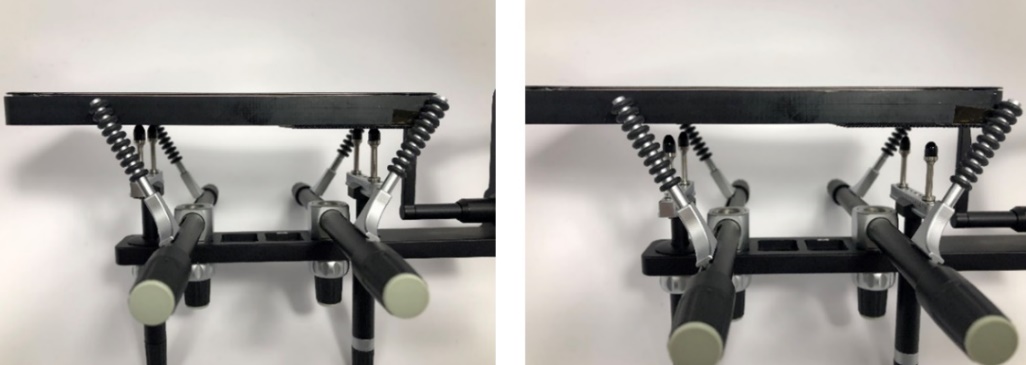


Figure 4-11: Handset positioned with and without support pins

##### 4.2.2.3.2 Playback system

See clause 4.2.1.2.2 for a description of the prototype playback system. Wideband pink noise was used for RFR measurement on both prototypes. Playback of the pink noise stimulus was accomplished using Adobe Audition.

The five commercial devices were driven by a Rohde & Schwarz CMW500 Wideband Radio Communication Tester. AMR-WB (20Hz–8kHz) at 12.65kbps was used for transmission speech coding. The measurement stimulus was an ITU-T Recommendation P.501 speech file with twelve utterances (six male and six female). ACQUA from Head Acoustics was used for stimulus playback. See Figure 4-12.

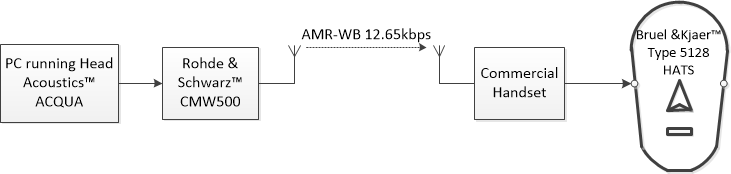


Figure 4-12: Block diagram of commercial device playback system

##### 4.2.2.3.3 Recording system

The RFR for each DUT was measured at the drum reference point (DRP) of Bruel & Kjaer Type 5128 HATS. Recordings were captured by a Head Acoustics MFE VI frontend and analysed with a Head Acoustics ACQUA system. All recordings were DRP-to-ERP (ear reference point) equalized prior to third octave band analysis.

#### 4.2.2.4 Measurement results

All devices were measured under two conditions, as described in clause 4.2.2.3.1. The transfer function between the measured RFR with supports and the RFR without supports was computed for each device. Figure 4-13 shows the results for all devices, divided between standard and non-standard earpieces. Table 4-7 presents the same data along with the root mean squared error (RMSE).

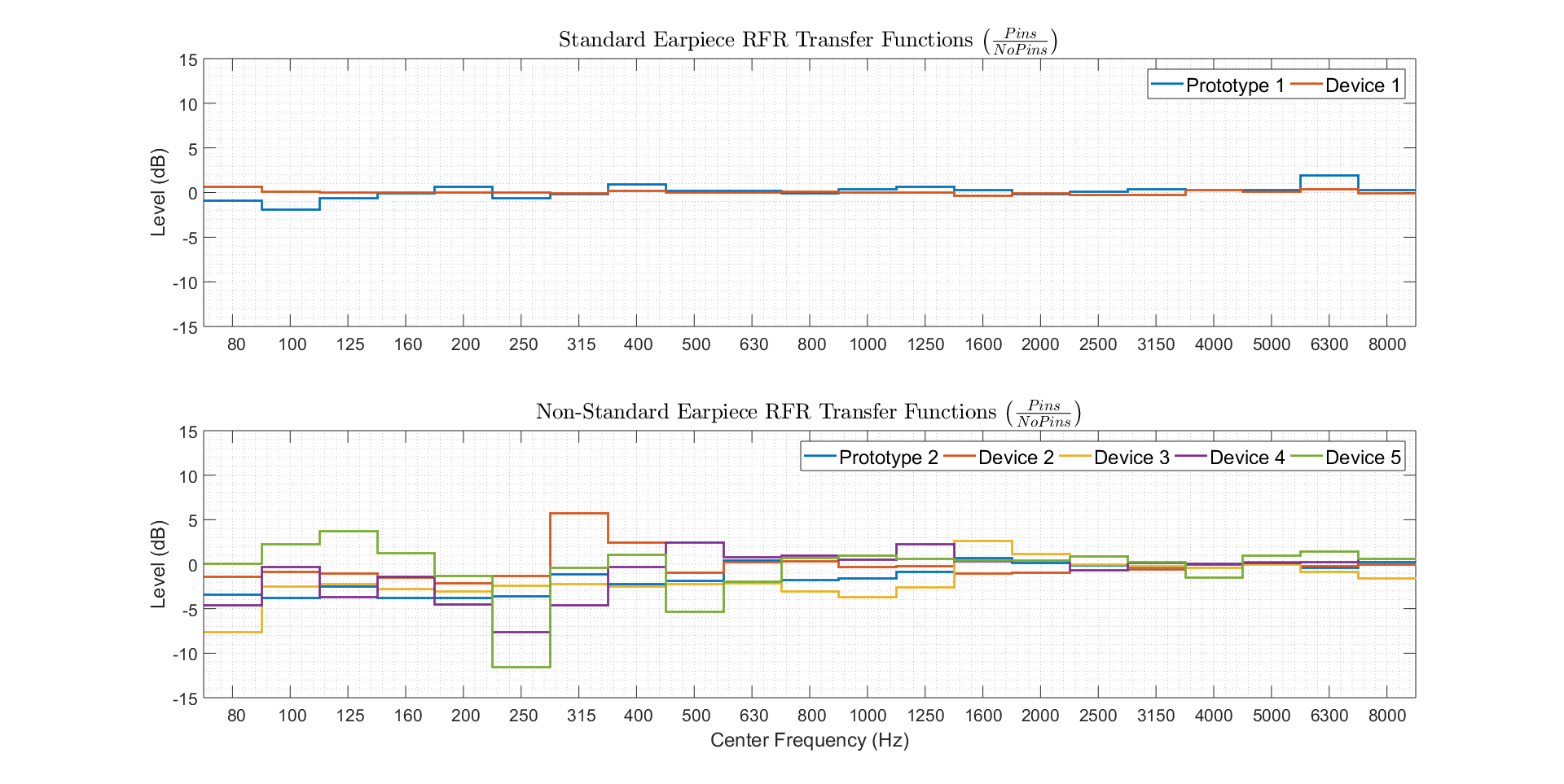


Figure 4-13: Comparing device RFR with and without supporting pins

Table 4-7: Comparing device RFR with and without supporting pins

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Center Frequency (Hz) | Prototype 1 | Device 1 | Prototype 2 | Device 2 | Device 3 | Device 4 | Device 5 |
| 80 | -0.84 | 0.69 | -3.35 | -1.38 | 0.05 | -4.54 | -7.56 |
| 100 | -1.89 | 0.13 | -3.79 | -0.86 | 2.27 | -0.31 | -2.46 |
| 125 | -0.60 | 0.05 | -2.49 | -1.00 | 3.72 | -3.64 | -2.23 |
| 160 | -0.07 | 0.02 | -3.78 | -1.47 | 1.27 | -1.40 | -2.77 |
| 200 | 0.72 | 0.03 | -3.71 | -2.12 | -1.25 | -4.47 | -3.02 |
| 250 | -0.60 | 0.03 | -3.53 | -1.26 | -11.48 | -7.61 | -2.37 |
| 315 | -0.18 | -0.01 | -1.11 | 5.80 | -0.34 | -4.55 | -2.24 |
| 400 | 1.00 | 0.26 | -2.16 | 2.51 | 1.10 | -0.28 | -2.46 |
| 500 | 0.27 | 0.09 | -1.86 | -0.89 | -5.29 | 2.51 | -2.17 |
| 630 | 0.22 | 0.01 | 0.43 | 0.26 | -1.93 | 0.83 | -2.08 |
| 800 | -0.01 | 0.14 | -1.71 | 0.37 | 0.69 | 0.97 | -2.98 |
| 1000 | 0.38 | 0.09 | -1.59 | -0.25 | 0.99 | 0.55 | -3.66 |
| 1250 | 0.68 | 0.02 | -0.79 | -0.20 | 0.64 | 2.28 | -2.58 |
| 1600 | 0.31 | -0.35 | 0.71 | -0.97 | 0.43 | 0.38 | 2.66 |
| 2000 | -0.13 | -0.06 | 0.16 | -0.88 | 0.44 | 0.46 | 1.21 |
| 2500 | 0.10 | -0.20 | -0.08 | -0.68 | 0.94 | -0.65 | -0.01 |
| 3150 | 0.43 | -0.20 | -0.25 | -0.53 | 0.31 | 0.21 | -0.17 |
| 4000 | 0.32 | 0.36 | -0.02 | 0.12 | -1.43 | 0.05 | -0.34 |
| 5000 | 0.36 | 0.18 | -0.04 | 0.24 | 1.04 | 0.22 | 0.01 |
| 6300 | 2.00 | 0.41 | -0.34 | -0.17 | 1.44 | 0.29 | -0.84 |
| 8000 | 0.32 | -0.08 | 0.29 | -0.03 | 0.62 | 0.63 | -1.54 |
| RMSE (80Hz-1kHz) | 0.75 | 0.22 | 2.70 | 2.10 | 3.96 | 3.45 | 3.33 |
| RMSE (80Hz-8kHz) | 0.75 | 0.23 | 2.06 | 1.63 | 3.05 | 2.67 | 2.68 |

It is shown in Figure 4-13 and Table 4-7 that standard earpiece devices (Prototype 1 and Device 1) demonstrate low levels of variability (< 1dB) as supporting pins are added or removed from the back of the device. Devices with a non-standard earpiece (e.g. Prototype 2 and Devices 2-5) tend to demonstrate higher levels of variability (> 2dB), particularly at frequencies below 1.25kHz. Furthermore, the non-standard earpiece variability is device and frequency dependent.

#### 4.2.2.5 Discussion

The results in clause 4.2.2.4 support the hypothesis that the receiver frequency response of a non-standard, vibrating display (HaNTE) is subject to mounting specific variability. Furthermore, the decoupled nature of a standard earpiece mitigates such variability.

## 4.3 Acoustic measurements according to 3GPP TS 26.131 and TS 26.132 and additional test cases to complement TS 26.131 and TS 26.132

### 4.3.1 Test setup

Three devices under test (DUTs), denoted A, B and C, were used for testing. They all have the form factor of a smartphone/phablet. As described in Table 4-8, DUTs A and B use a vibrating display to produce sounds in handset mode with no ear piece, while DUT C is a traditional handset UE with an ear piece.

Table 4-8: DUT description

|  |  |
| --- | --- |
| DUT | UE Type |
| A | Handset with vibrating display |
| B | Handset with vibrating display |
| C | Handset with traditional ear piece |

For each DUT, MECRP and mounting instructions were provided by the manufacturer – the positioning instructions can translated to P.64 Annex E [3]. Each DUT was mounted on a B&K 4128C HATS placed in an acoustically treated room complying with requirements in TS 26.132. The test simulator was based on Head Acoustics ACQUA 4.0.100, with Head Acoustics MFE VI.1 and MFE VIII.1, and Rohde & Schwarz CMW500 was used for LTE connections.

Measurements (in handset mode) were conducted according to TS 26.131 and 26.132 (Rel-13). For the sake of conciseness, only wideband (WB) conditions are reported for LTE connections, using the AMR-WB (12.65 kbit/s) codec; the NB case was also tested, and in the following some remarks are also provided to illustrate relevant DUT behavior in NB when necessary. The default application force of 8 N was applied. A subset of test results, mainly RLR and RFR results, are reported hereafter.

### 4.3.2 Test results in receiving

#### 4.3.2.1 WB RLR

The measured WB RLR at all volume control setting (according to TS 26.131, clause 6.2.2 and TS 26.132, clause 8.2.2.2) is shown in Figures 4-14a, 4-14b, and 4-14c for DUT A, B and C, respectively. The nominal values of RLR are constrained in the non-hashed area corresponding to the interval 2 ± 3 dB. The nominal position 'nom.' is chosen for each DUT such that the target RLR value of 2 dB is met as close as possible. Note that the figures also show the WB RLR constraints for the maximal position 'max.' (> -13 dB, < -3 dB) and for minimum position 'min' (< 18 dB).

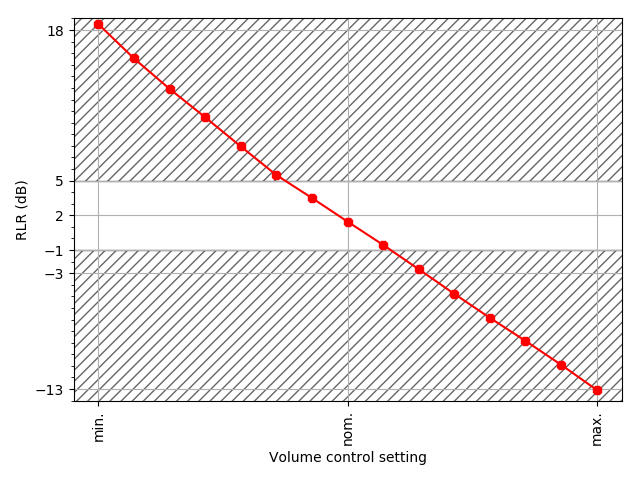


Figure 4-14a: WB RLR for DUT A.

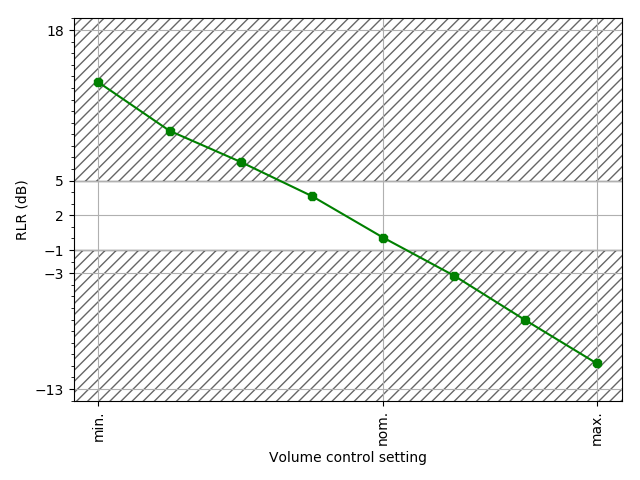


Figure 4-14b: WB RLR for DUT B.

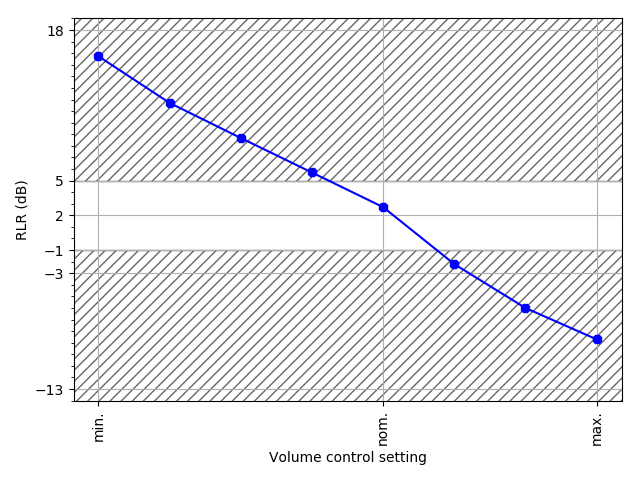


Figure 4-14c: WB RLR for DUT C.

#### 4.3.2.2 WB RFR

For each DUT, the WB frequency response was measured at the nominal volume control setting according to TS 26.131 clause 6.4.2 and TS 26.132 clause 8.4.2. The results for DUT A, B, and C, are provided in Figure 4-15. It may be observed that DUT A and DUT B do not meet WB RFR requirements, however DUT A is very close to fitting in the WB frequency mask.

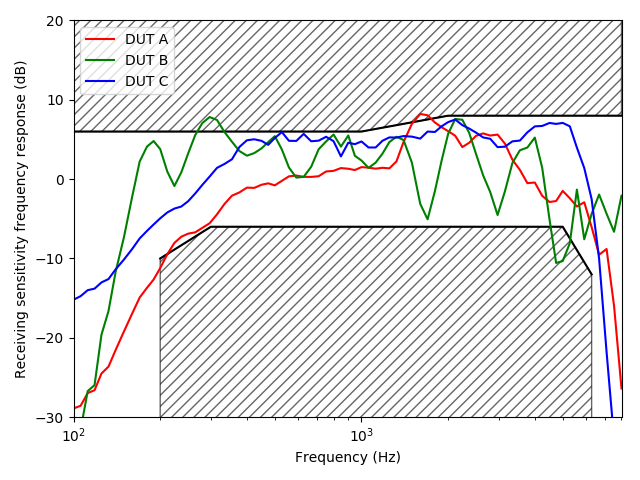


Figure 4-15: WB RFR for DUT A, B, and C at nominal volume control (1/12 octave band resolution)

### 4.3.3 Additional results on user studies and objective measurements

#### 4.3.3.1 General

In addition to acoustic measurements, two other types of evaluations were reported:

- Listening tests: Informal expert listening tests were conducted to evaluate the subjective quality of LTE WB voice calls with DUT A, B and C.

- User feedback from commercial deployment.

The following observations can be made based on these evaluations for DUT B:

- The vibration of the display was found to be very noticeable at maximum volume control, including some vibration felt in fingers and the hand, and some issues of privacy were also reported (i.e. the voice call may be heard by neighbouring people from the back of the phone, in contrast to traditional UEs with an ear piece in handset mode which do not exhibit such behaviour).

- Depending on far-end voice characteristics (e.g. male or female talker), it was noted that the voice timbre can be quite distorted when listening at maximum volume control.

- Voice quality was found to be distorted by out-of-band noise, especially at higher maximum volume settings. This noise is present in the WB case (frequencies> 7kHz) and it is even more noticeable in NB (frequencies > 4kHz).

To address these issues, the following test cases are proposed to be included in TS 26.131 and TS 26.132.

#### 4.3.3.2 Additional test cases to complement TS 26.131 and TS 26.132

##### 4.3.3.2.1 Out-of-band energy level measurement from SWB RFR measurement

Figures 4-16a and 4-16b show the SWB RFR data measured at maximum and nominal volume control, when using the SWB RFR test method specified in TS 26.131 clause 7.4.2.1 and TS 26.132 clause 9.4.2.1, for the WB phones under test (DUT A to C). This measurement provides a frequency analysis range extended to the at least [100, 16000] Hz.

NOTE 1: The SWB RFR test method in TS 26.132 specifies that a double resolution (1/12 octaves and 1/3 octaves) is used for measurement, however the 1/12 octave band is hard-limited to a maximum frequency of about 8 kHz in the setup used in this study. Therefore only 1/3 octave band results are reported in Figures 4-16a and 4-16b.

NOTE 2: The SWB frequency masks in Figures 4-16a and 4-16b do not apply in the WB case and they are only used for information.

NOTE 3: It was observed that DUT A has signal saturations (with clipped audio) in two specific time segments in the 23-25s time interval when volume control is set to maximum in receiving (recalling that the RFR test signal consisting of concatenated British English single sentences is about 35s long); such saturations were observed only for DUT A and only at maximum volume control and it should be considered when interpreting the estimated frequency response.

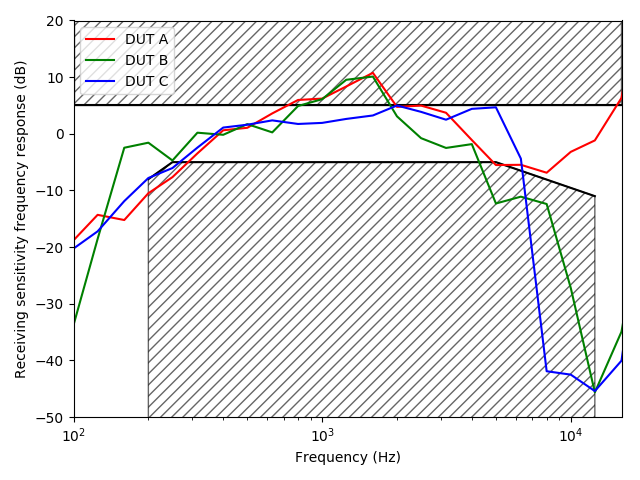


Figure 4-16a: RFR for DUT A, B, and C with extended frequency range and 1/3 octave band analysis (maximum volume control)

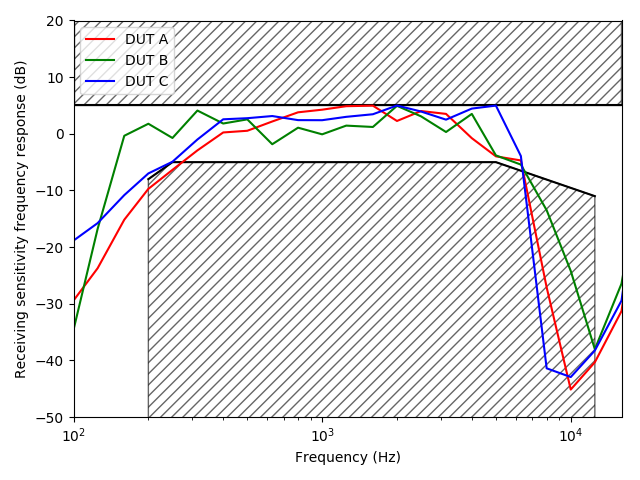


Figure 4-16b: RFR for DUT A, B, and C with extended frequency range and 1/3 octave band analysis (nominal volume control)

It can be noted that DUT B has some out-of-band energy in the 1/3 octave bands centered at 8 and 10 kHz. One caution is that the test signal is typically pre-processed with a pass-band filter which does not allow conducting a spectral analysis in the stop band region, therefore results above 12.5 kHz are not relevant here.

To complement the 1/3 octave band analysis, the test signals recorded when measuring SWB RFR at nominal volume control were also directly analysed. These recorded signals are sampled at 48 kHz, and a periodogram using an FFT size of 4096 samples, with Hanning window and 75 % overlap is reported in Figure 4-17.

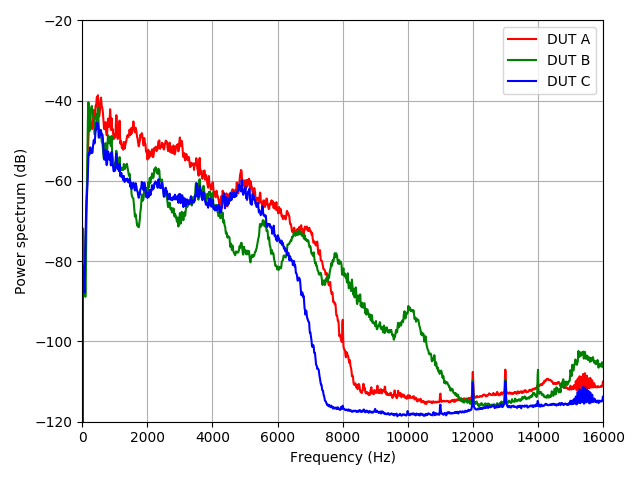


Figure 4-17: Periodogram of SWB RFR output test signal for DUT A, B, and C (nominal volume control)

The out-of-band energy level is shown to be significantly higher for DUT B.

Test results are provided here only for WB terminals; however, it was found that the same out-of-band energy issue is even more noticeable when the same phone (DUT B) operates in NB.

##### 4.3.3.2.2 P.863 measurement

The measured values (using POLQA v2.4) for DUT A, B, and C, at nominal and maximum volume control are reported in Table 4-9. It may be noted that the MOS-LQO value measured at nominal level ('Test Condition 0') for delay tests specified in TS 26.131, clause 6.11.1 and TS 26.132, clause 8.10.4 is already part of existing test cases. In principle, no extra measurement is therefore necessary for the nominal case; one may consider at least reporting this value for the nominal case.

The values reported in Table 4-9 may be used to quantify quality issues reported in user feedbacks.

Table 4-9: DUT description

|  |  |  |
| --- | --- | --- |
| DUT | MOS-LQOTEST at  nominal volume control | MOS-LQOTEST at  maximum volume control |
| A | 3.3 | 3.5 |
| B | 2.6 | 2.6 |
| C | 3.8 | 3.5 |

##### 4.3.3.2.3 WB RFR at maximum volume control

TS 26.132 clause 5.1 specifies that nominal volume control should be used for RFR measurement. The measured WB RFR at three volume control settings (minimum, nominal, and maximum) is shown in Figures 4-18a, 4-18b, and 4-18c for DUT A, B and C.

It can be observed that RFR at maximum volume control is spectrally less balanced than at nominal and minimum volume position for DUTs A and B, while for a traditional UE such as DUT C the RFR does not vary significantly across the volume range. This may explain the user feedback on timbre degradation reporting for the maximum volume settings.

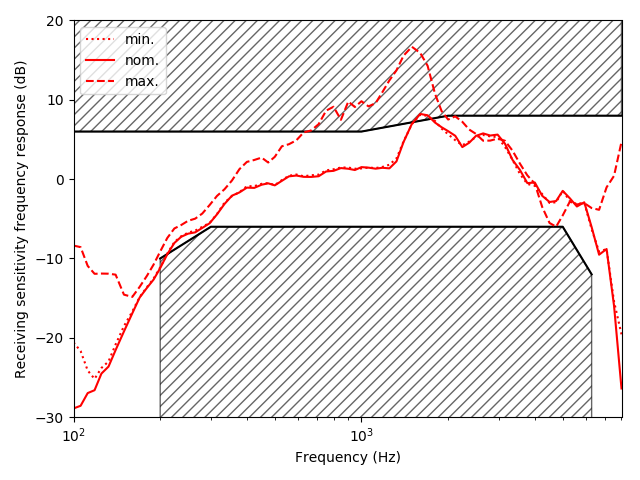


Figure 4-18a: WB RFR for DUT A at minimum, nominal and maximum volume control

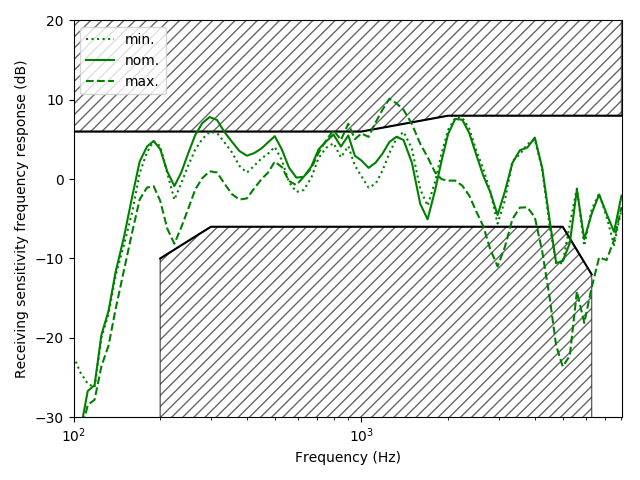


Figure 4-18b: WB RFR for DUT B at minimum, nominal and maximum volume control

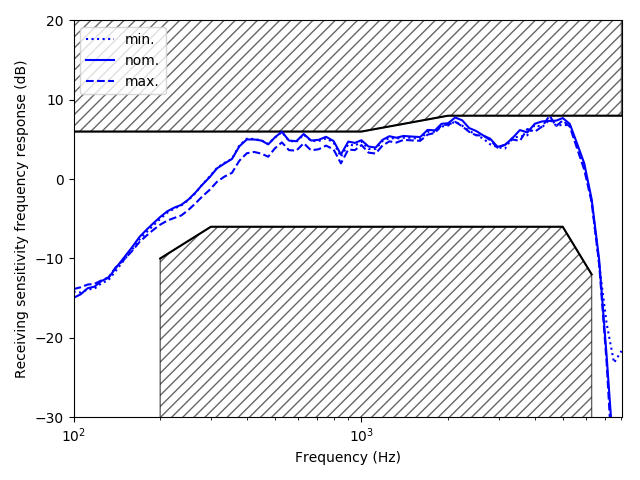


Figure 4-18c: WB RFR for DUT C at minimum, nominal and maximum volume control

NOTE: It was observed that DUT A has signal saturations (with clipped audio) in two specific time segments in the 23-25s time interval when volume control is set to maximum in receiving (recalling that the RFR test signal consisting of concatenated British English single sentences is about 35s long); such saturations were observed only for DUT A and only at maximum volume control and it should be taken into account when interpreting the estimated frequency response.

### 4.3.4 Discussion

The results presented in clause 4.3.2 show that acoustic measurements according to 3GPP TS 26.131 and 26.132 can be conducted on handset UEs featuring non-traditional earpieces, when MECRP and mounting position information are provided.

## 4.4 Acoustic measurements according to 3GPP TS 26.131 and TS 26.132 with various mounting conditions and test equipment

### 4.4.1 Introduction

3GPP handset terminal acoustic test methodologies [1] and requirements [2] were developed for UE designs supporting traditional earpieces. However, handset manufacturers are now selling UE's featuring non-traditional earpieces (HaNTE devices) such as vibrating display handsets. In the present study, standard acoustic terminal tests defined in [1] are performed on two commercially available HaNTE devices to determine the applicability of such methodologies to vibrating display handsets. The measurements are repeated with multiple mounting positions and at several measurement points along the handset's display. Variability and considerations specific to HaNTE acoustic measurements are discussed.

### 4.4.2 Test Setup

Two measurements are made for each device:

- Receiving Loudness Rating (RLR) defined in clause 8.2.2.2 of [1]

- Receiving Sensitivity/Frequency Characteristics (RFR) defined in clause 8.4.2 of [1]

All measurements are conducted in an acoustically isolated anechoic chamber. Two head and torso simulators (HATS) are used for the acoustic measurements, each with its own handset positioning system. Figure 4-19 shows the two HATS with their respective handset positioners.

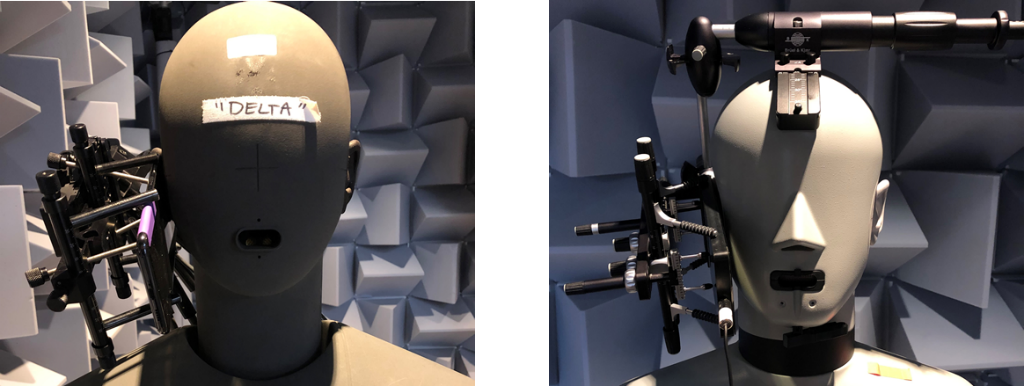


Figure 4-19: HATS 1 and 2 with handset positioners

Measurements are made with an MECRP of ye=17mm, ze=0mm for all devices. HATS 1 measurements are made on a grid of 9 additional MECRP shifts taken from clause 5 of [3]. Figure 4-20 demonstrates the 10 measurement points.

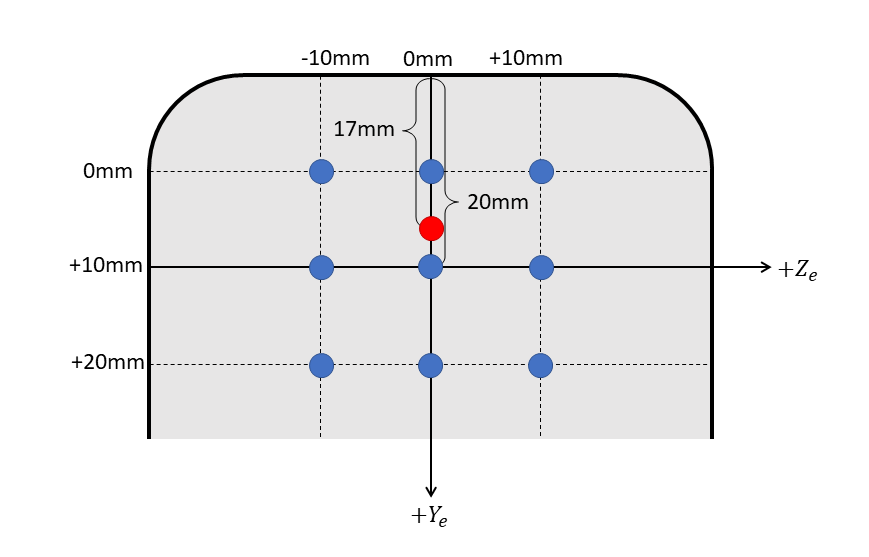


Figure 4-20: Arbitrary MECRP (red) and all other RLR/RFR measurement points for HATS 1

HATS 1 measurements are repeated under two handset support conditions shown in Figure 4-21. Mounting Position 1 (MP1) is a standard device mounting with support pins placed on the back of the handset. In Mounting Position 2 (MP2), the supporting pins are retracted, and the bottom fork is adjusted lower on the handset.



Figure 4-21: Variable handset support conditions used for HATS 1 measurements

On HATS 2, devices are mounted according to ITU-T Rec. P.64 (A = 21.2°, B = -12.9°, C = 2.3°). Due to contact with the HATS 1 cheek in the standard mounting position, a 5⁰ offset is added to the B angle for measurements on HATS 2.

Device volume is set as close as possible to the nominal RLR of 2 ± 3dB at the defined MECRP to comply with clause 6.2.2 of [2]. Test calls are established using the AMR-WB codec at 12.65 kbps. In what follows, RLR and RFR measurement results are provided for the two commercial HaNTE devices.

### 4.4.3 Measurement Results

#### 4.4.3.1 Device 1 on HATS 1 Measurements

##### 4.4.3.1.1 Receiver Loudness Rating Results

Device 1 was adjusted to a nominal RLR of 1.91dB at the MECRP with MP1 on HATS 1. Table 4-10 presents Device 1 RLR results for all MECRP shifts at MP1 and MP2 on HATS 1. Figure 4-22 provides a visualization of the positions which pass the requirements specified in [2].

Table 4-10: Device 1 RLR results (bold\* = within 2±3dB tolerance [2])

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Position | | RLR with MP 1 (dB) | RLR with MP 2 (dB) | Difference |
| ye | ze |
| 17 | 0 | **1.91\*** | **3.32** | -1.41 |
| 10 | -10 | 7.99 | 10.53 | -2.54 |
| 10 | 0 | 9.79 | 12.00 | -2.21 |
| 10 | 10 | 13.83 | 16.22 | -2.39 |
| 20 | -10 | **-0.29\*** | **-0.20** | -0.09 |
| 20 | 0 | **1.01\*** | **0.98** | 0.03 |
| 20 | 10 | **3.84\*** | **3.75** | 0.09 |
| 30 | -10 | **-0.86\*** | -3.35 | 2.49 |
| 30 | 0 | **0.74\*** | -1.48 | 2.22 |
| 30 | 10 | **3.42\*** | **1.01** | 2.41 |
| Mean | | 4.14\* | 4.28 | -0.14 |

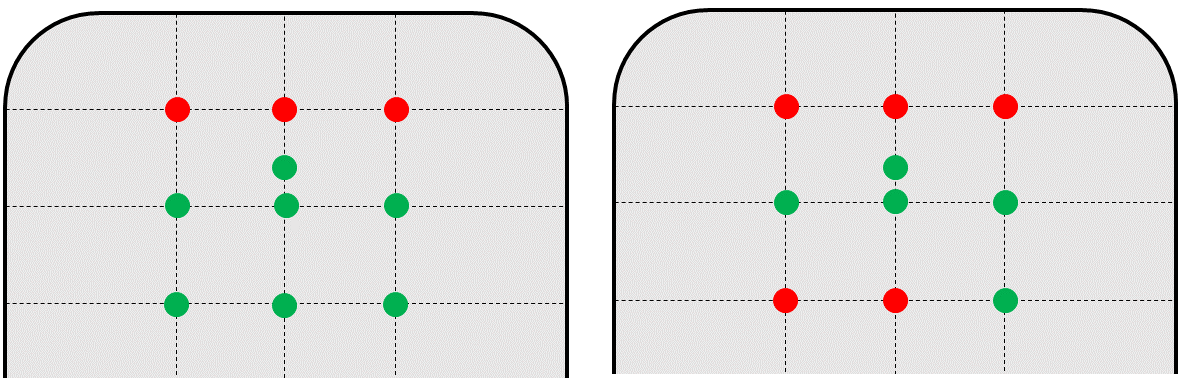


Figure 4-22: Device 1 RLR results with MP1 (left) and MP2 (right) (green = pass, red = fail)

Seven of ten tested points (including the MECRP) are within the RLR tolerance specified in [2] for MP1 and five of ten for MP2. The average RLRs for both MP1 (4.14dB) and MP2 (4.28dB) lie within the tolerance as well. MP1 RLR results are on average 0.14dB louder than MP2 results with a maximum variation of -2.54 dB at ye=10mm, yz=-10mm. For all ye = 10 positions, MP1 has a higher loudness rating than MP2. For positions with ye = 30, MP1 has a lower loudness rating. And for ye = 20, the RLR is consistent between MP1 and MP2.

##### 4.4.3.1.2 Receiver Sensitivity/Frequency Response Results

Using the previously defined nominal RLR of 1.68dB at the MECRP, Device 1 RFRs are measured on HATS 1 with two mounting positions at 10 points along the display. Figure 4-23 shows all RFR measurements (normalized to 0dB [Pa/V] maximum) along with the wideband tolerance from [2]. The arithmetic mean for each mounting position is presented as well. All individual measurements fail the wideband mask. Table 4-11 reports the maximum deviation from the tolerance mask for each measurement.

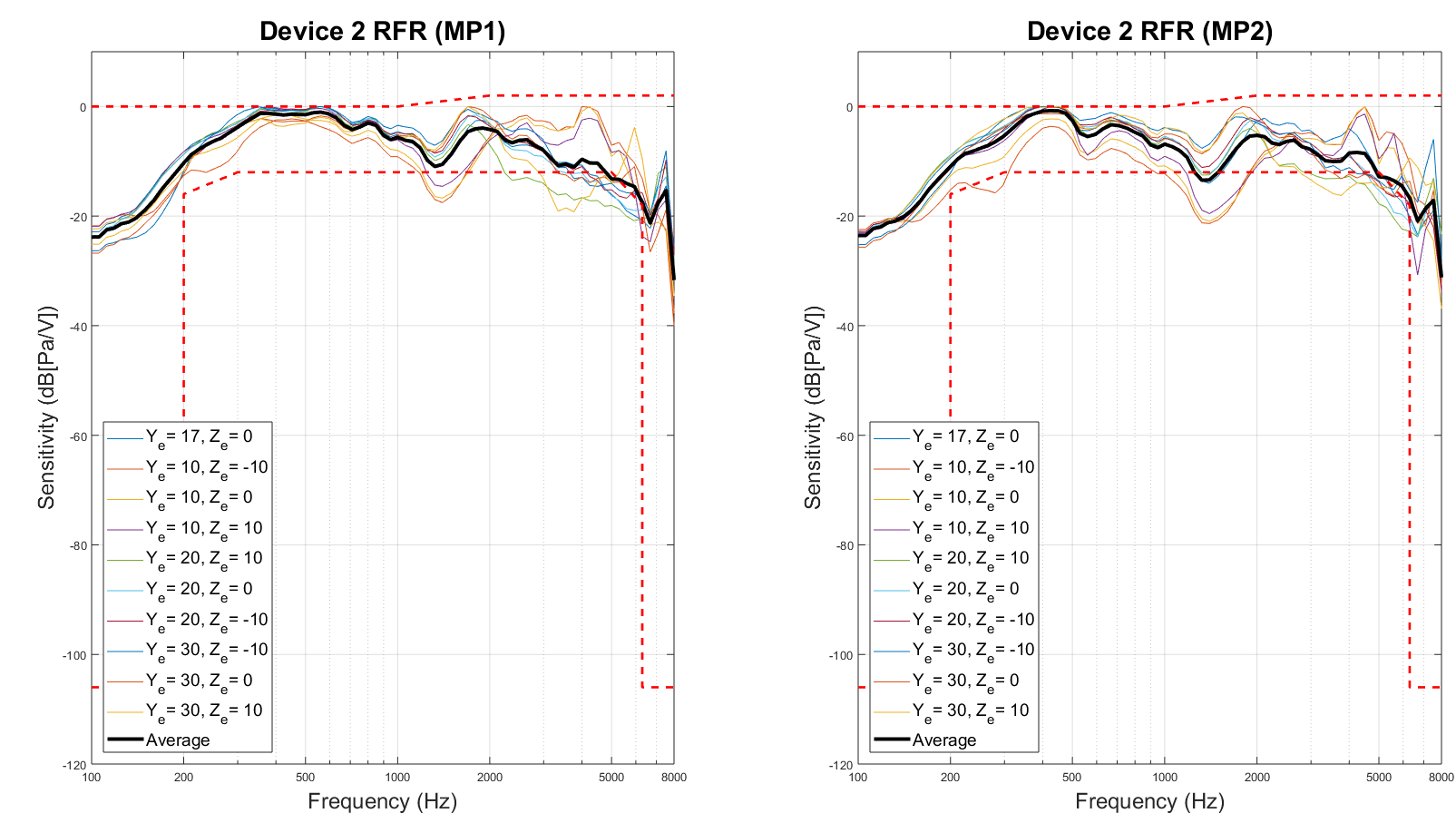


Figure 4-23: Device 1 RFR measured on HATS 1 at two mounting positions

Table 4-11: Device 1 RFR errors on HATS 1

|  |  |  |  |
| --- | --- | --- | --- |
| Position | | MP1 Error (dB) | MP2 Error (dB) |
| ye | ze |
| 17 | 0 | -4.12 | -2.07 |
| 10 | -10 | -3.54 | -7.34 |
| 10 | 0 | -3.11 | -6.96 |
| 10 | 10 | -5.68 | -7.55 |
| 20 | -10 | -3.23 | -2.37 |
| 20 | 0 | -4.27 | -4.42 |
| 20 | 10 | -6.20 | -5.87 |
| 30 | -10 | -2.57 | -1.58 |
| 30 | 0 | -0.39 | -1.34 |
| 30 | 10 | -5.69 | -4.23 |
| Mean Error | | -3.88 | -4.37 |

Variability of RFR results between MP1 and MP2 is presented in Figure 4-24. A positive value indicates higher sensitivity with MP1. For Device 1, there is a peak in variability (~4dB) at 550Hz and 1.8kHz for all measurement points. Variability never exceeds 6dB.

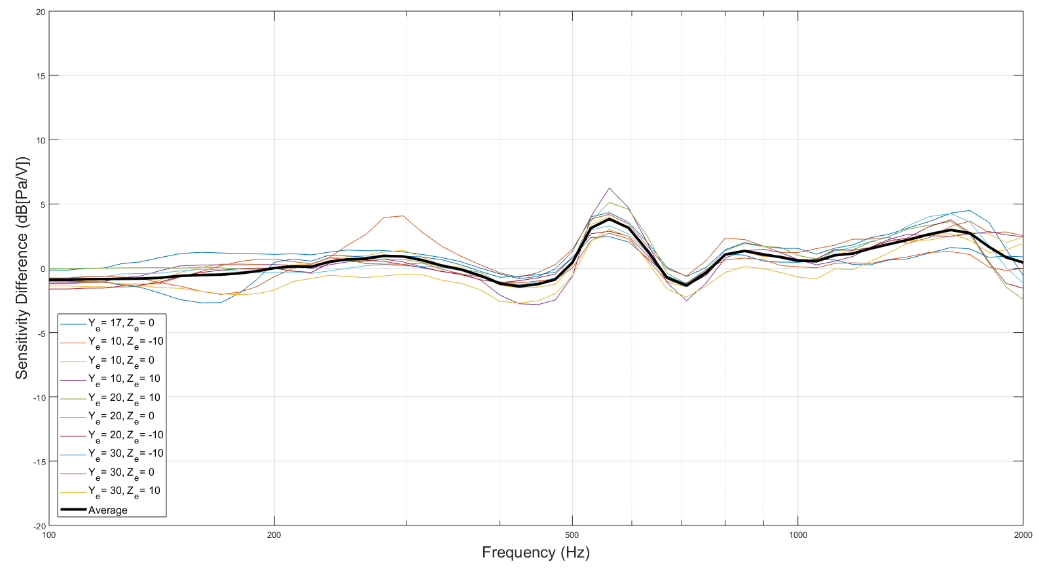


Figure 4-24: Device 1 RFR mounting variability

#### 4.4.3.2 Device 1 on HATS 2 Measurements

Device 1 RFR was also measured on HATS 2 with the previously defined MECRP. Device playback was calibrated to a nominal RLR of 0.61dB. Figure 4-25 shows the HATS 2 results along with HATS 1 MECRP and average RFRs using MP1. Device 1 passes the RFR mask from [2] when measured on HATS 2.

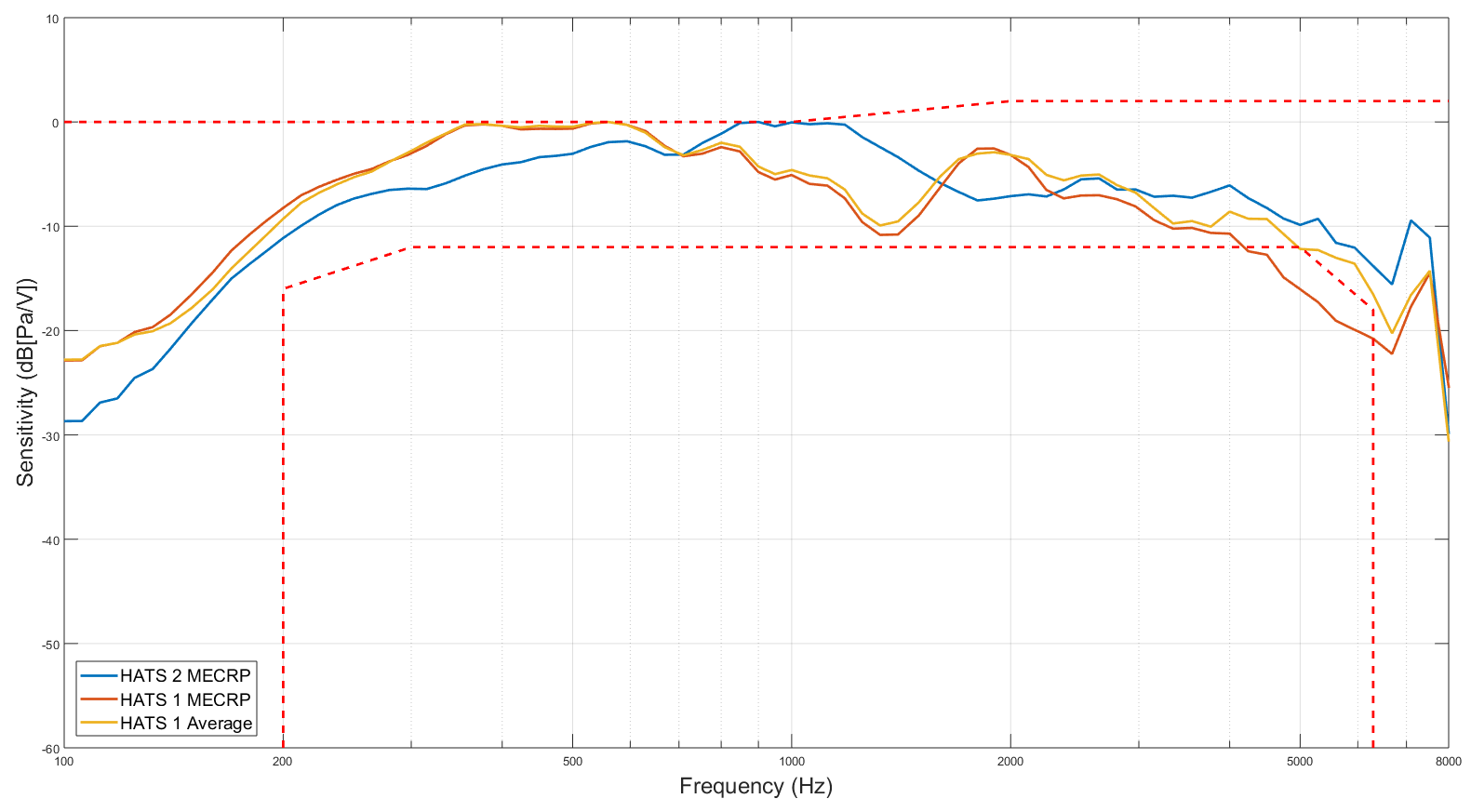


Figure 4-25: Device 1 RFR measured on HATS 2

#### 4.4.3.3 Device 2 on HATS 1 Measurements

##### 4.4.3.3.1 Receiver Loudness Rating Results

Device 2 was adjusted to a nominal RLR of -0.11dB at the MECRP with MP1. Table 4-12 presents Device 1 RLR results for all MECRP shifts at MP1 and MP2. Figure 4-26 provides a visualization of which positions pass the requirements specified in [2].

Table 4-12: Device 2 RLR results (bold\* = within 2±3dB tolerance [2])

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Position | | RLR with MP1 (dB) | RLR with MP2 (dB) | Difference |
| ye | ze |
| 17 | 0 | **-0.11\*** | **0.91\*** | -1.02 |
| 10 | -10 | 8.17 | 9.59 | -1.42 |
| 10 | 0 | 7.39 | 8.59 | -1.2 |
| 10 | 10 | 7.04 | 7.78 | -0.74 |
| 20 | -10 | **0.34\*** | **-0.12\*** | 0.46 |
| 20 | 0 | -1.09 | -1.09 | 0.00 |
| 20 | 10 | -1.79 | -2.02 | 0.23 |
| 30 | -10 | **2.18\*** | **0.04\*** | 2.14 |
| 30 | 0 | **0.67\*** | **-0.95\*** | 1.62 |
| 30 | 10 | **0.15\*** | -1.72 | 1.87 |
| Mean | | 2.30\* | 2.10\* | 0.19 |

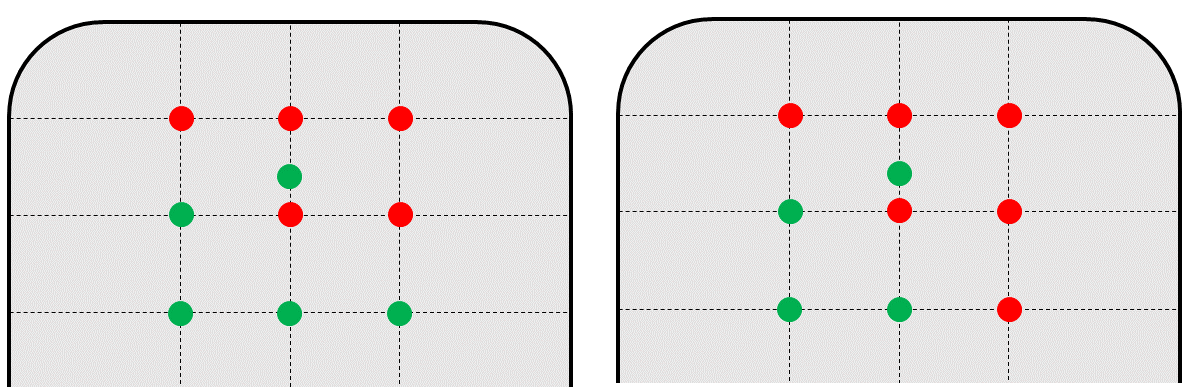


Figure 4-26: Device 2 RLR results with MP1 (left) and MP2 (right) (green = pass, red = fail)

Of the ten points tested, five (including the MECRP) are within the RLR tolerance specified in [2] for MP1 and four (including the MECRP) for MP2. The average RLRs for both MP1 (2.30) and MP2 (2.10) lie within the tolerance. MP2 RLR results are on average 0.19dB louder than MP1 with a maximum variation of 2.14dB at ye=30mm, yz=-10mm.

##### 4.4.3.3.2 Receiver Sensitivity/Frequency Response Results

Using the previously defined nominal RLR of -0.11dB at the MECRP, Device 2 RFRs are measured on HATS 1 with two mounting positions at 10 points along the display. Figure 4-27 shows all RFR measurements (normalized to 0dB [Pa/V] maximum) along with the wide-band tolerance from [2]. Furthermore, the arithmetic mean for each mounting position is presented. All individual measurements as well as the mean RFR fail the wide-band mask. Table 4-13 reports the maximum deviation from the tolerance mask for each measurement

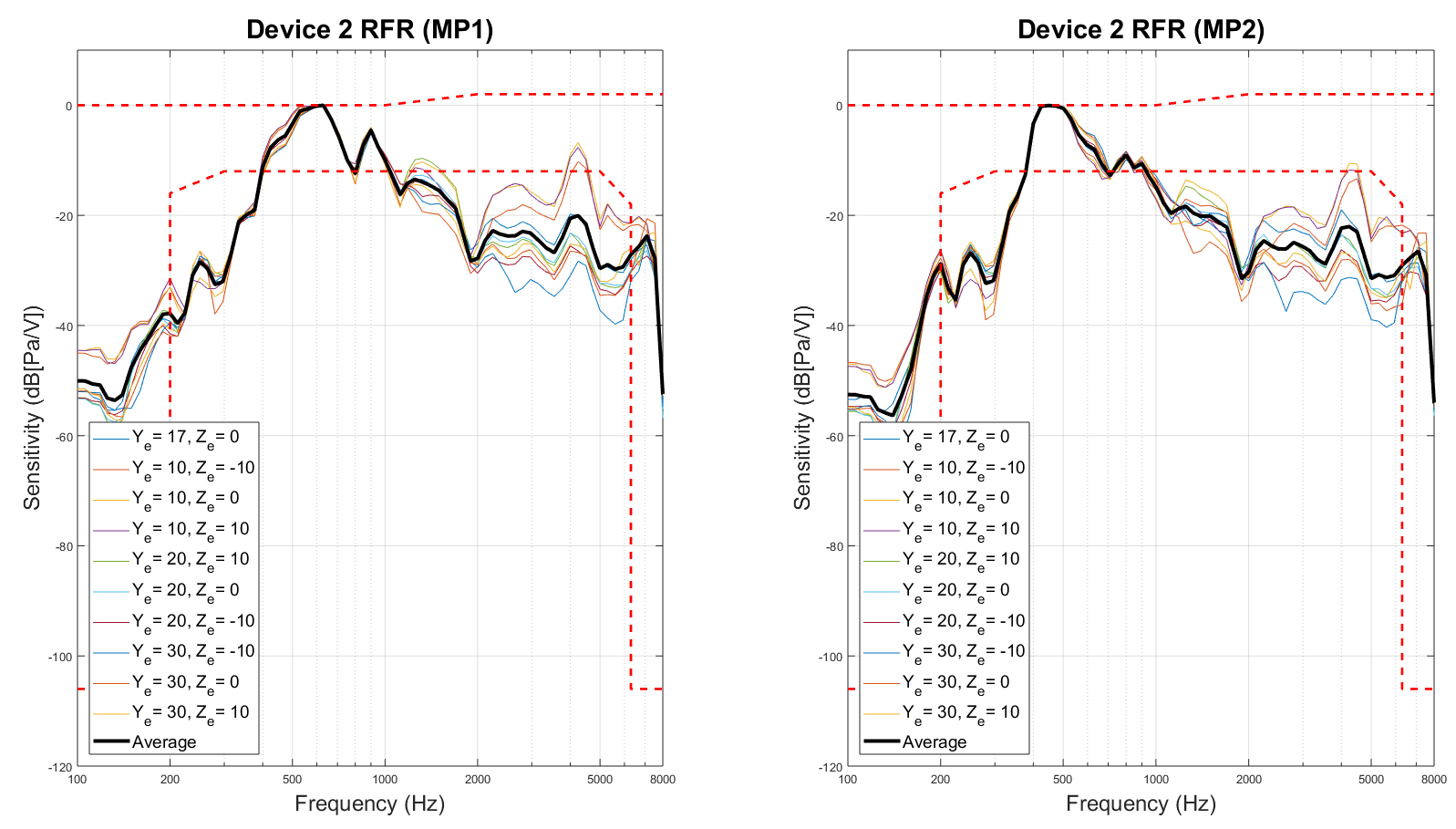


Figure 4-27: Device 2 RFR measured on HATS 1 at two mounting positions

Table 4-13: Device 2 RFR errors on HATS 1

|  |  |  |  |
| --- | --- | --- | --- |
| Position | | MP1 Error (dB) | MP2 Error (dB) |
| ye | ze |
| 17 | 0 | -25.43 | -20.58 |
| 10 | -10 | -25.23 | -26.30 |
| 10 | 0 | -23.78 | -24.61 |
| 10 | 10 | -22.12 | -22.49 |
| 20 | -10 | -26.54 | -23.54 |
| 20 | 0 | -25.95 | -21.33 |
| 20 | 10 | -25.63 | -21.25 |
| 30 | -10 | -25.25 | -26.96 |
| 30 | 0 | -26.49 | -24.07 |
| 30 | 10 | -25.74 | -22.01 |
| Mean | | -25.22 | -23.31 |

Variability of RFR results between MP1 and MP2 is presented in Figure 4-28. A positive value indicates higher sensitivity with MP1. For Device 2, all individual measurements demonstrate consistent and substantial (at some points > 10dB) variability between 100Hz and 2kHz.

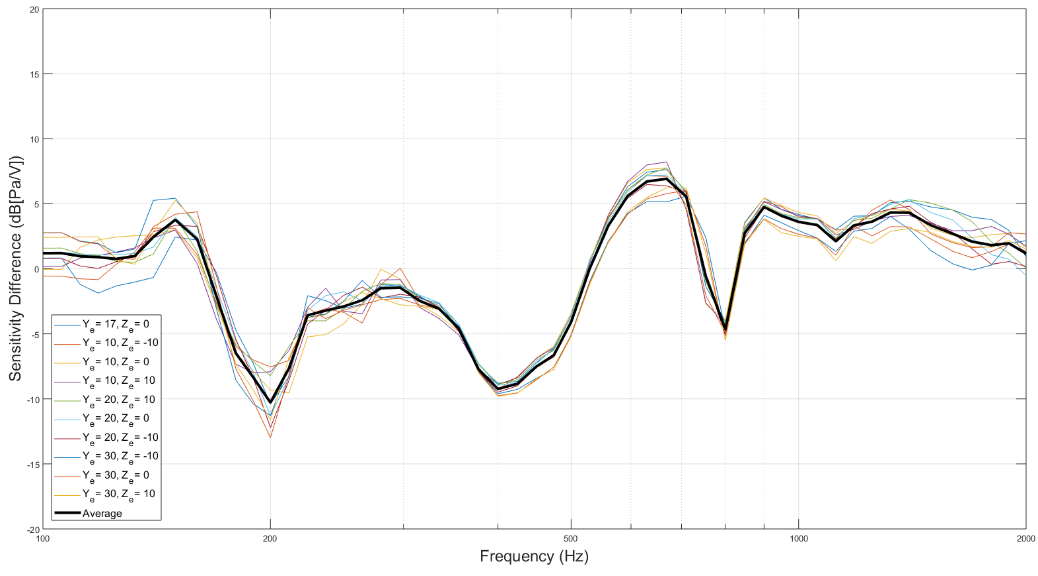


Figure 4-28: Device 2 RFR mounting variability

#### 4.4.3.4 Device 2 on HATS 2 Measurements

Device 2 RFR was also measured on HATS 2 with the previously defined MECRP. Device playback was calibrated to a nominal RLR of 2.0dB. Figure 4-29 shows the HATS 2 results along with HATS 1 MECRP and average RFRs using MP1. In-band sensitivity differences between the two HATS exceed 10dB at 250Hz and reach 8dB at 2.5kHz.

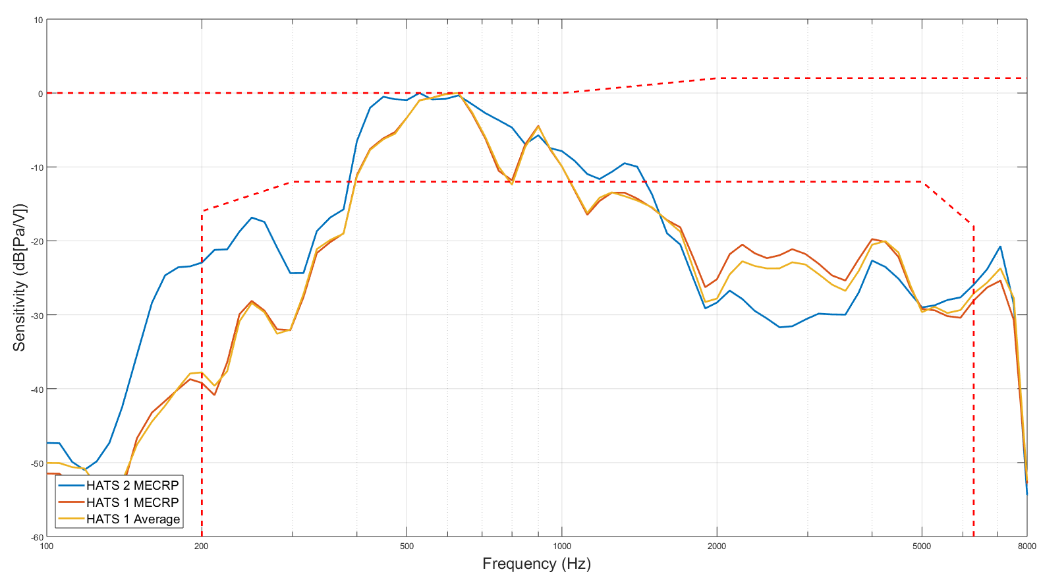


Figure 4-29: Device 2 RFR measured on HATS 2

### 4.4.4 Discussion

In the previous clauses, acoustic measurement results for two commercial HaNTE devices were presented. For each device the MECRP was defined as ye=17mm, ze=0mm. Each device was measured on two HATS at the MECRP. On HATS 1, the devices were measured at an additional 9 points in an equally spaced grid proposed in [3]. Furthermore, HATS 1 measurements were made with a standard mounting position and with an alternative mounting position. RLR and RFR results for all measurements were assessed according to the requirements in [2] and variability between HATS, measurement points, and mounting positions was reported.

Device 1 exemplifies a relatively high performing HaNTE device. RLRs stay within an acceptable range over a larger area of the device display when mounted in a standard position (MP1). Although no RFR measurements on HATS 1 pass the mask defined in [2], all positions are within 8dB of passing no matter the mounting position. Furthermore, Device 1 passes the RFR sensitivity mask when measured on HATS 2. Finally, mounting specific variability is limited to < 6dB in the frequency range 100Hz – 2kHz.

Device 2, however, demonstrates poor performance for a HaNTE device. The RLR is more susceptible to measurement point variability and fewer points fulfil the requirements of [2]. Furthermore, Device 2 RFR is far from passing the frequency mask no matter the measurement point (>20dB absolute error) or HATS. Lastly, the Device 2 RFR is prone to high levels of mounting specific variability at frequencies below 2kHz.

## 4.5 HaNTE round robin test in different labs

### 4.5.1 Introduction

This contribution reports on the Handsets featuring Non-Traditional Earpieces (HaNTE) round robin results at laboratories #1 (Qualcomm), #2 (HEAD acoustics GmbH), #3 (Orange) and #4 (Huawei Finland).

Tests were conducted according to an agreed test plan. However, due to some observed issues, some measurements and analyses were extended during the round robin test.

### 4.5.2 Devices

Seven commercially available mobile phones were evaluated as devices under test (DUT), which all provide a non-traditional earpiece. For reference, an additional device with a traditional earpiece was evaluated as an eighth device. Since the form factor of all devices is similar (smart phones), the *alternative handset position* (also called *flat handset position*) according to ITU-T P.64 [4], where B-axis is rotated by ‑5° (∆B=5°).

For the HaNTE-devices (DUT1 to DUT7), a common ECRP of Ye=-21mm was used for testing. ECRP of DUT8 (non-HaNTE device) was determined via the visible acoustic outlet.

### 4.5.3 Test plan

#### 4.5.3.1 Overview

Two different types of head and torso simulators (HATS) and corresponding handset positionier as listed in Table 4-14 were used for testing. Both HATS comply with ITU-T P.58 [3] and are equipped with ear simulators of Type 3.3 [2]. The handset positioners both comply with ITU‑T P.64 [4].

Table 4-14: HATS and Positioner systems used in the round robin test

|  |  |  |  |
| --- | --- | --- | --- |
| ID | Manufacturer | HATS | Handset positioner |
| H | HEAD acoustics | HMS II.3 | HHP IV (automated), HHP III (manual) |
| B | Brüel & Kjær | Type 4128C | Type 4606 (manual) |
| NOTE: The letters in the ID column of Table 4-14 are used in the following to distinguish between these two systems. | | | |

Participating labs in the round robin test could also test with more than one HATS/positioner system. An overview of the measurement systems vs labs is shown in Table 4-15.

Table 4-15: HATS and Positioner systems per lab

|  |  |  |  |
| --- | --- | --- | --- |
| Lab | Company | HATS/Positioner | Comment |
| #1 | Qualcomm, Inc. | B |  |
| H |  |
| #2 | HEAD acoustics GmbH | B |  |
| H |  |
| #3 | Orange | B | Some re-tests werde made after Lab #3 to correct some measurements with wrong volume settings |
| H | Due to lack of time, only frequency response and speech quality tests are available. Dataset still contains some measurements with wrong volume settings and were not re-tested |
| #4 | Huawei | B | Due to lack of time, only frequency response and speech quality tests are available |
| H |  |

For testing of receive loudness rating and frequency response, 3GPP TS 26.132 [6] specifies the usage of the British English single talk sequence according to clause 7 of ITU-T P.501 [9]. In order to reduce testing time, in several test cases also a shorter measurement signal according to Annex D of ITU-T P.501 [9] was utilized, too. Thus, results obtained with the longer default sequence are denoted as "Long" in the following, while results obtained with the short sequence are denoted as "Short".

In order to further minimize testing time and to re-use recordings as often as possible, all HATS measurements were conducted without DRP-ERP or DF-correction. For the analysis of receive frequency responses (RFR) and speech quality according to ITU-T P.863 [8], DF-correction according to ITU-T P.58 [3] was applied. For the calculation of receive loudness ratings, DRP-ERP-correction according to ITU-T P.57 [2] was applied.

For the automated handset positioner HHP IV, the test suite used for the evaluation provided an optional automated volume control check, which sweeps all possible volume settings of a device (via Bluetooth remote control). Volume control settings for maximum (MAX) and nominal (NOM) loudness ratings were determined with the short British English test sequence according to Annex D of ITU-T P.501 [9]. In addition, the resulting maximum volume steps were double-checked in lab #2 by manually counting the possible volume steps (from minimum to maximum) during a call.

Since RLR methodology was not yet investigated with (possibly) non-linear sound radition of HaNTE devices, in some cases the existing recordings were analyzed with the loudness measures ISO 532-1  [11] and ITU-T P.700 [12]:

- The loudness level (in phon) according to ISO 532-1 utilizes the time-varying loudness, which was calculated over the whole signal. The overall result is obtained by temporally aggregating with the 95% percentile versus time ("N5 percentile").

- The loudness level (in phon) according to ITU-T P.700 is also based on the time-varying loudness as per ISO 532-1, but the temporal aggregation is adapted for speech signals:

- Speech pauses larger than 400 ms are excluded.

- Instead of a percentile analysis, the average of the loudness versus time is used.

All DUTs were connected via 3G/UMTS to the test equipment, calls were setup with AMR-WB codec at 12.65 kbit/s. An application force of 8 N was used for all tests.

During the round robin test, DUT6 was damaged, and the display showed multiple compound cracks. For the re-testing in Lab1 at the end of the round robin, it was decided not to use this device again and was left out for the measurements.

#### 4.5.3.2 Test 1: Speech Quality

For speech quality testing at nominal and maximum volume, eight sentences of British English speech (two female, two male talkers) from ITU-T P.501 [9] were concatenated to an overall sequence of 32.0s. The source signal for the receive direction was pre-filtered to wideband and then calibrated to an active speech level (ASL) according to ITU-T P.56 [10] of -16 dBm0.

The recording is then analyzed with the speech quality prediction method according to ITU-T P.863 [8] in super-wideband mode (version 2.4), analyzing sentences pairs. The resulting four MOS values are averaged to an overall result value.

#### 4.5.3.3 Test 2: Privacy

As a measure for acoustic sound radiation of the devices, the differences in RLR between handset mode and far-field (measured at a radius of 42 cm in front of the HATS EEP) is evaluated as shown in Figure 4-30. Different angles between symmetry plane of the HATS and the measurement microphone are used for testing, i.e., points A, B, C, D and E as defined in Table 4-16. The test is run with maximum volume setting at the DUT. Only the short sequences are available for the analysis.

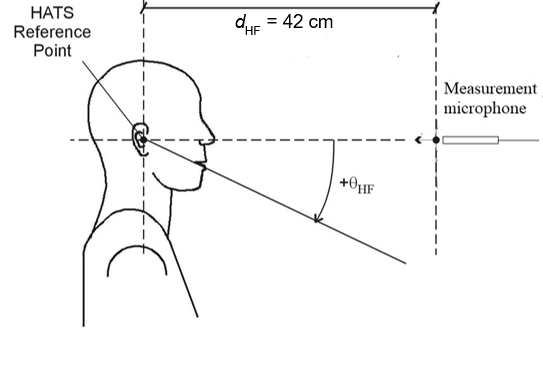
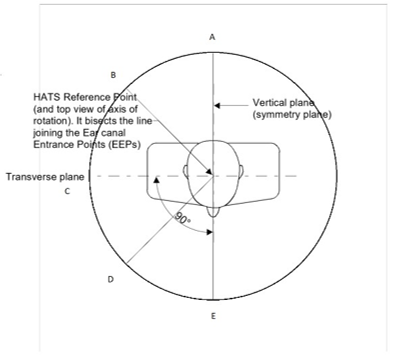


Figure 4-30: Test setup for privacy

Table 4-16: Angles for testing privacy

|  |  |
| --- | --- |
| Measuring position | Measuring angle (starting from E) |
| A | -180° |
| B | -135° |
| C | -90° |
| D | -45° |
| E | 0° |

#### 4.5.3.4 Test 3a: Robustness (variation of ECRP)

To investigate the robustness of the (chosen) ECRP, several shifts of 1 cm in Ze and Ye direction are evaluated. Figure 4-31 shows the resulting grid located around the default ECRP (marked in green). Four mandatory (marked in red) and four optional (marked in blue) points are evaluated. Table 4-17 provides unique labels (S0-S8) for the shifts to be applied. Table 4-18 shows a rearranged view of these labels according to the geometry of the grid in Figure 4-31.

RLR and RFR are evaluated with the long and short speech sequence as defined in clause . The test is run with nominal volume setting.

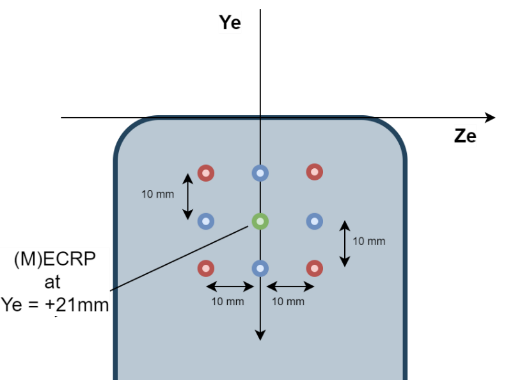


Figure 4-31: Variation of ECRP

Table 4-17: Shifts around ECRP (=S0)

|  |  |  |  |
| --- | --- | --- | --- |
| Shift | Offset Ze [mm] | Offset Ye [mm] | Type |
| S0 | 0 | 0 | Mandatory |
| S1 | -10 | -10 | Mandatory |
| S2 | +10 | -10 | Mandatory |
| S3 | +10 | +10 | Mandatory |
| S4 | -10 | +10 | Mandatory |
| S5 | 0 | -10 | Optional |
| S6 | +10 | 0 | Optional |
| S7 | 0 | +10 | Optional |
| S8 | -10 | 0 | Optional |

Table 4-18: Gemeotric representation of shifts

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | Offset Ze [mm] | | |
| -10 | 0 | +10 |
| Offset Ye [mm] | -10 | S1 | S5 | S2 |
| 0 | S8 | S0 | S6 |
| +10 | S4 | S7 | S3 |

#### 4.5.5.5 Test 3b: Robustness (variation of fork position)

To investigate the impact of the clamping fork positions, the round robin test evaluates three different positions. Initial results from lab #1 defined these three positions as offsets in Ye-axis. The general positioning strategy for the forks shown in Table 4-19 and is not bound to a vendor-specific handset positioner.

Table 4-19: Fork positions

|  |  |  |  |
| --- | --- | --- | --- |
|  | Bottom | Middle | Top |
| Fork position #1 | ✓ | ✓ |  |
| Fork position #2 | ✓ |  | ✓ |
| Fork position #3 |  | ✓ | ✓ |

All measurements except Test 3b are conducted with position #1. Only Test 3b evaluates positions #2 and #3. Exact values for bottom, middle and top fork position were initially provided by lab #1. However, it was found to be extremely impractical to measure the fork positions at each mounting.

Thus, the following general positioning strategy was chosen for lab #2 (and was then adopted by lab #3 and #4):

- Bottom position: the fork is moved and tightened as close as possible at the lower edge of the device. This fork positioning did not conflict with any button at the sides of the devices.

- Mid position: the fork is moved and tightened as close as possible at the center of the device (regarding Ye axis). In case of collisions with buttons at the sides of the device, the fork is moved to the closest collision-free position (typically, this is towards the lower edge of the device since most buttons are in the upper half of the device).

- Top position: the fork is moved and tightened as close as possible at the most upper edge of the device. In case of collisions with buttons at the sides of the device, the fork is moved downwards to the closest collision-free position.

In addition to the Mid and Top positions, care was taken that the clamps of the forks did not produce an overhang, as shown in Figure 4-32. In this case, the head of the clamp (red color) might push against the ear/cheek of the HATS and the screen of the device (orange color) is not mounted correctly.



Figure 4-32: Possible overhang of fork positions

RLR and RFR are evaluated with the long and short speech sequence as defined in clause 4.5.3.1. The test is run with nominal volume setting.

### 4.5.4 Results

#### 4.5.4.1 Preparations

As described in clause 4.5.3.1, the volume steps for nominal and maximum loudness ratings were initially determined and reported by lab #2. Even though it was suggested to use the same determined volume settings in the following labs, in some cases different volume settings were used. All used volume settings per lab and per HATS/positioner are provided in Table 4-20.

Table 4-20: Volume settings for MAX and NOM

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| DUT | MAX | NOM | | | | | |
| Lab #1 (B/H) | Lab #2 (B/H) | Lab #3 (B) | Lab #3 (H) | Lab #4 (B) | Lab #4 (H) |
| DUT1 | 7 | 4 | 4 | 4 | 4 | 4 | 3 |
| DUT2 | 7 | 6 | 6 | 6 | 6 | 6 | 6 |
| DUT3 | 14 | 6 | 6 | 6 | 6 | 8 | 7 |
| DUT4 | 14 | 7 | 7 | 7 | 6 | 7 | 6 |
| DUT5 | 6 | 2 | 2 | 2 | 2 | 2 | 1 |
| DUT6 | 10 | 6 | 6 | 6 | - | 7 | 7 |
| DUT7 | 5 | 1 | 1 | 1 | 1 | 2 | 1 |
| DUT8 | 15 | 6 | 6 | 6 | 6 | 6 | 6 |

Some DUTs showed strange behaviour here:

1) DUT2 does not provide a valid nominal volume setting. During incremental testing, RLR "jumps" from ~8.5 dB at step 5/7 to ~-2.5 dB at step 6/7.

2) DUT4 seems to have an incorrect implementation of volume control. At a certain volume increase step, the recorded signal level decreased. This was confirmed by measurements as well as by expert listening.

3) DUT5 provides a “boost mode” in the volume control. This mode was disabled and not used for testing.

The resulting RLR values for all labs, DUTs and long test sequence are shown in Figure 4-33 (nominal volume) and Figure 4-34 (maximum volume). Performance requirements according to clause 6.2 of 3GPP TS 26.131 [5] are indicated as tolerances in the figures as well. If available, results for the long and short sequence (see clause 4.5.3.1) are provided. Loudness levels in phon according to ITU-T P.700 are provided for reference in Figure 4-35 (nominal volume) and Figure 4-36 (maximum volume). Limits indicated in these figures are derived from the RLR requirements.



Figure 4-33: RLR for NOM volume

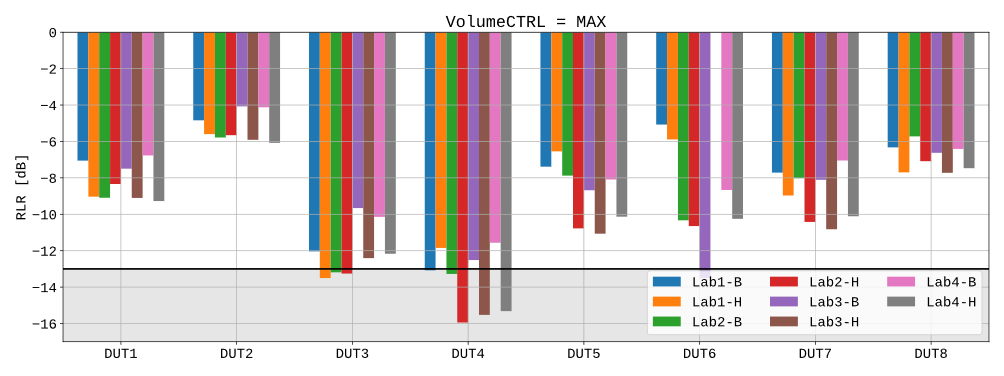


Figure 4-34: RLR for MAX volume

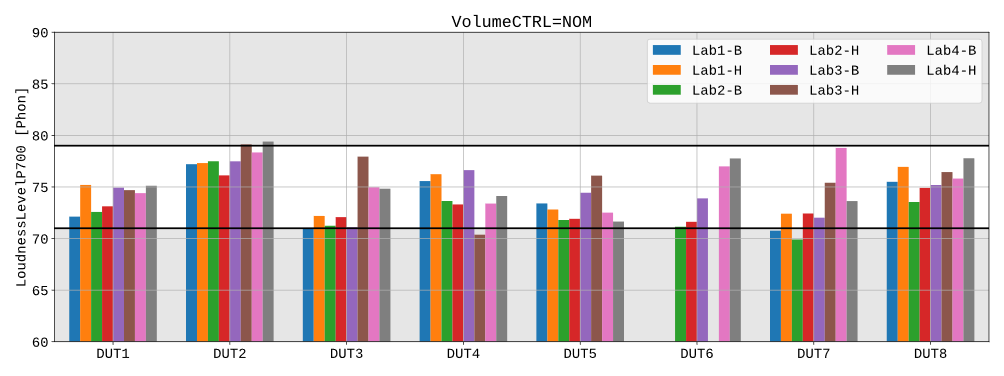


Figure 4-35: Loudness levels according to ITU-T P.700 for NOM volume

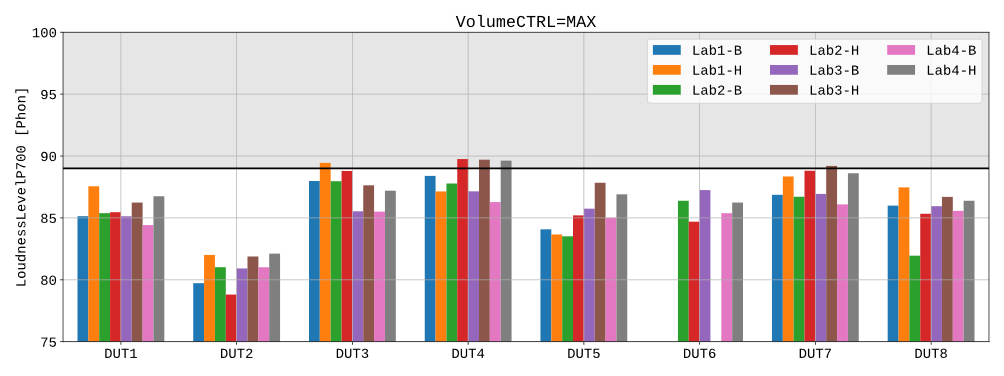


Figure 4-36: Loudness levels according to ITU-T P.700 for MAX volume

The results of the RFR measurements according to clause 8.4.2 of 3GPP TS 26.132 [6] (using the long single talk sequence from ITU-T P.501 [9]) are shown for NOM and MAX volumes and for all devices in Figure 4-37 to Figure 4-44.

NOTE 1: In all figures below, the tolerances according to clause 6.4.2 of 3GPP TS 26.131 [5] are provided only for reference and better visual orientation. The graphs are intended to illustrate the variance across labs, HATS and/or positioners. Any evaluation of the performance requirement is out of scope for this analysis. Furthermore, 3GPP TS 26.131 [5] also does not specify performance requirements for RFR and MAX volume setting.

NOTE 2: For the illustration of this variance, any level difference between curves in one graph should remain. For this purpose, the frequency responses in each graph are adjusted commonly with one single offset. This offset is calculated by adjusting the average RFR (not shown) to the upper tolerance.

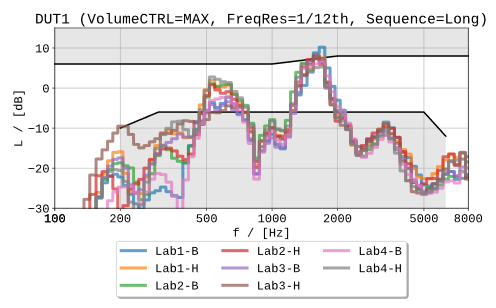
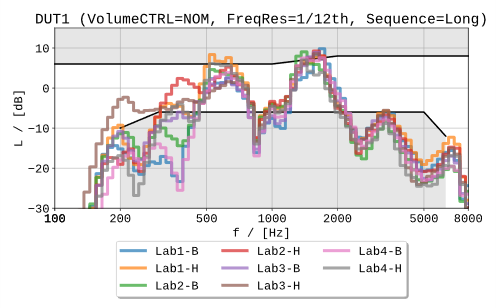


Figure 4-37: RFR results of DUT1 for NOM (left) and MAX (right)

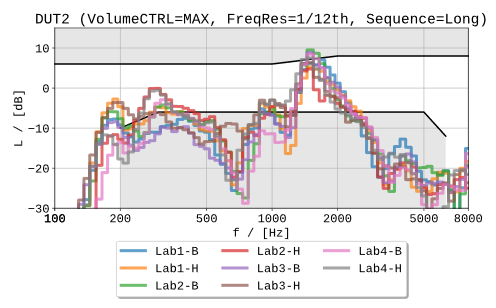
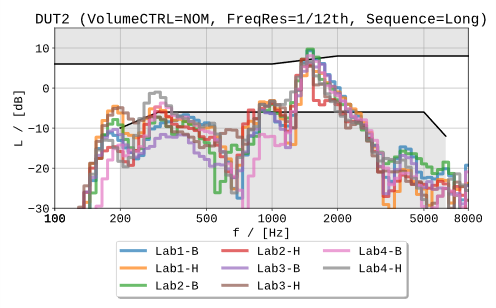


Figure 4-38: RFR results of DUT2 for NOM (left) and MAX (right)

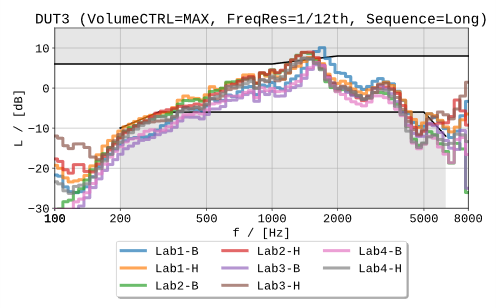
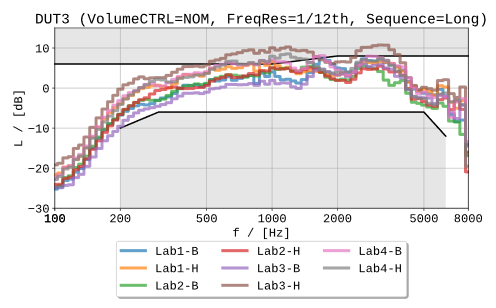


Figure 4-39: RFR results of DUT3 for NOM (left) and MAX (right)

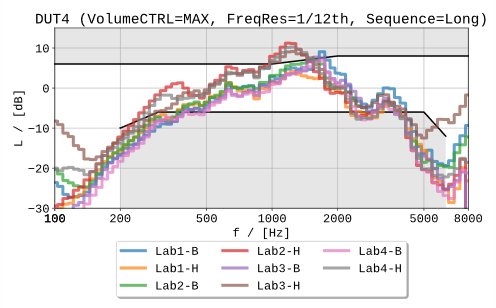
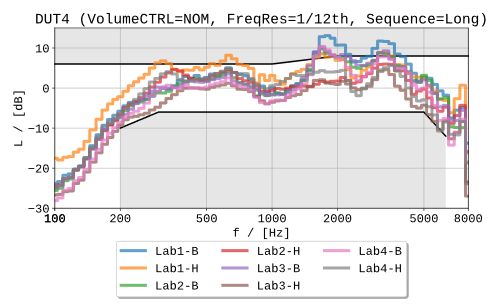


Figure 4-40: RFR results of DUT4 for NOM (left) and MAX (right)

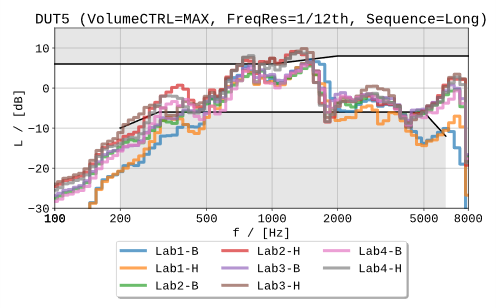
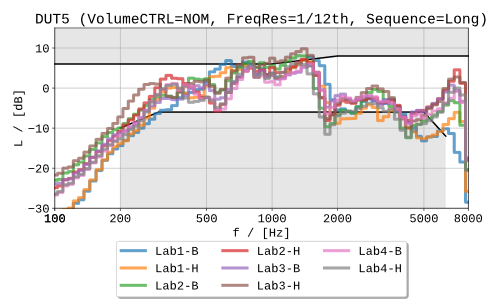


Figure 4-41: RFR results of DUT5 for NOM (left) and MAX (right)

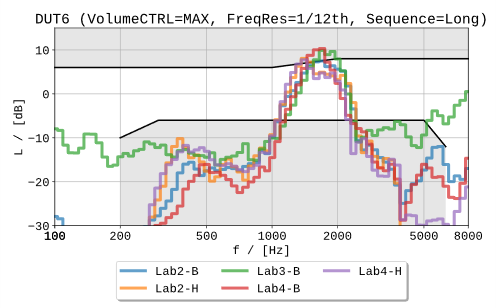
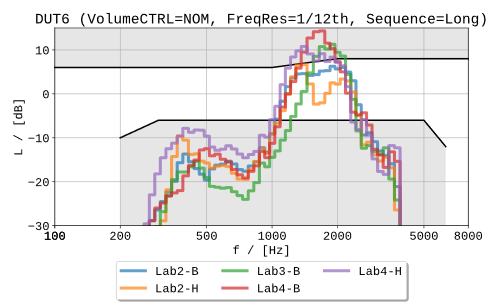


Figure 4-42: RFR results of DUT6 for NOM (left) and MAX (right)

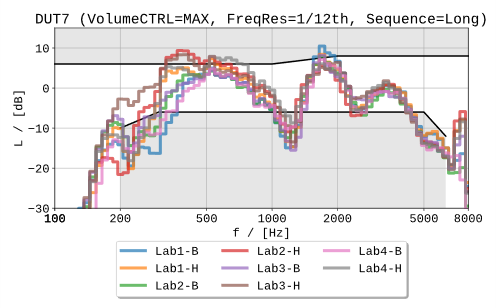
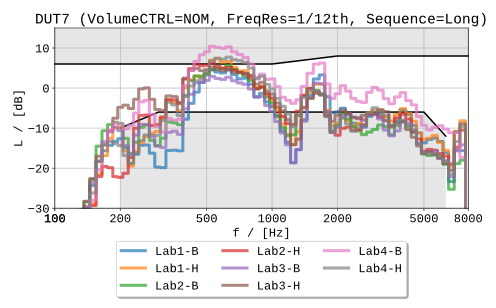


Figure 4-43: RFR results of DUT7 for NOM (left) and MAX (right)

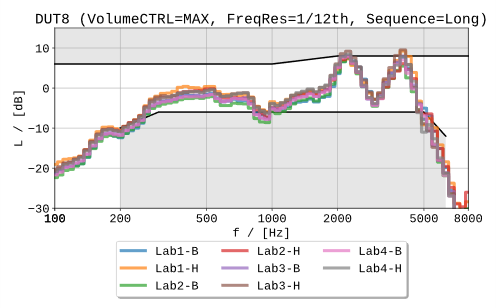
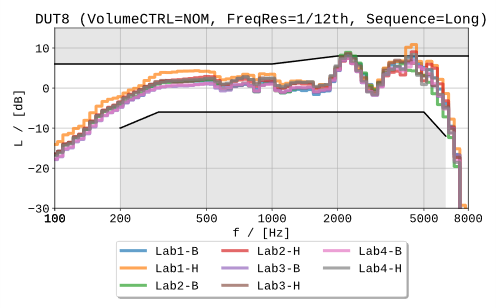


Figure 4-44: RFR results of DUT8 for NOM (left) and MAX (right)

#### 4.5.4.2 Test 1: Speech Quality

Figure 4-45 and Figure 4-46 provide the results of the speech quality testing according to ITU-T P.863 for NOM and MAX volume settings for all labs. In several cases, the decrease in quality between these two volume settings is more than 0.5 MOS. Lower performance of some devices may be explained by the frequency response results as shown in clause 4.1.

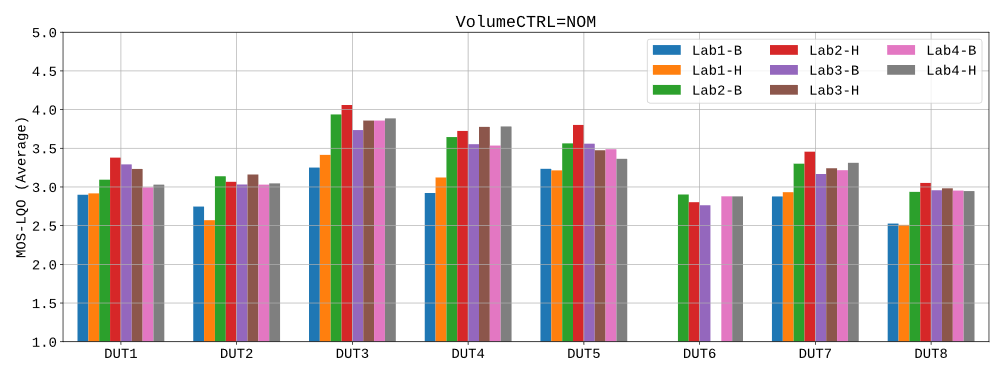


Figure 4-45: P.863 results for NOM volume setting vs labs

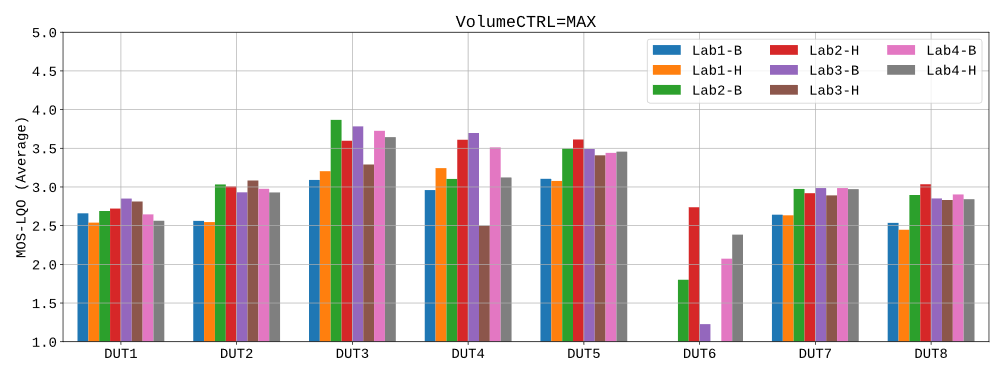


Figure 4-46: P.863 results for MAX volume setting vs labs

NOTE: The performance of DUT6 at MAX volume (see Figure 4-46) seems to be unreasonably low and/or inconsistent for lab #2 and #3. After investigation and expert listening, it was noticed that in some cases, DUT6 emits an extremely high peak level (> 30 dBPa!), which was not taken into account by the dynamic range settings in both labs and results in a clipped amplitude. This artifact leads to a clear degradation in listening quality prediction of P.863. This behaviour occurred for two different HATS and in labs (using different measurement equipment).

#### 4.5.4.3 Test 2: Privacy

For testing of privacy, the test plan for the round robin test specified RLR measurements at MAX volume setting. At the same time, RLR was measured at the ear (DUT) and in the vicinity (microphone) and the difference between both should reveal performance regarding sound radiation. After an initial analysis of the results, the method of (Delta-)RLR calculation seems not to be suitable.

Due to the low overall level of the recorded signals at the microphone position, the analysis of RLR or Delta-RLR seems not to provide reasonable results. Loudness ratings according to ITU-T P.79 [7] are based on a linearized loudness model, which only works for levels in a suitable "operational range" and the recorded signal at the microphone is not considered to be a valid "signal under test". In general, the level at the microphone should be as low as possible.

For this reason, a post-analysis of the RLR measurements was conducted. The unweighted and A-weighted levels are calculated for the microphone signals (short speech sequence). Leading/trailing silence segments are excluded (0.5 to 6.0 s) and frequency range from 100 to 8000 Hz was considered. Results for all angles and labs are shown in Figure 4-47 and Figure 4-48.

NOTE: Due to the low level (and thus, lower SNR), the usage of active speech level according to ITU-T P.56 [10] is not possible here in all cases.

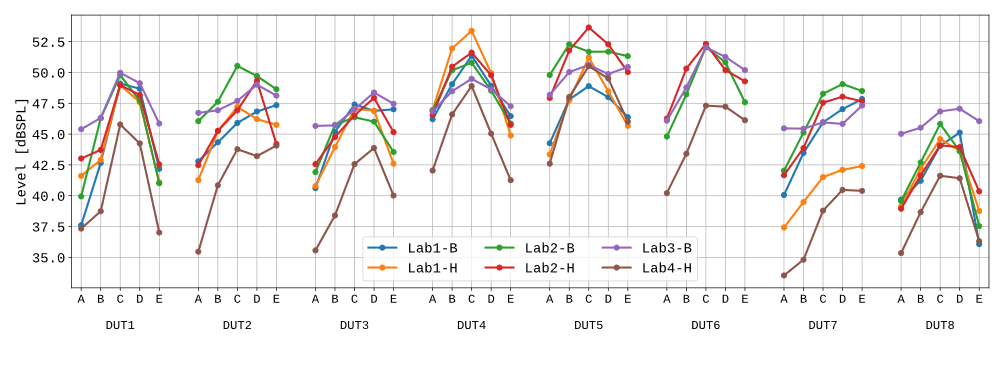


Figure 4‑47: Level (unweigthed) for privacy tests vs labs

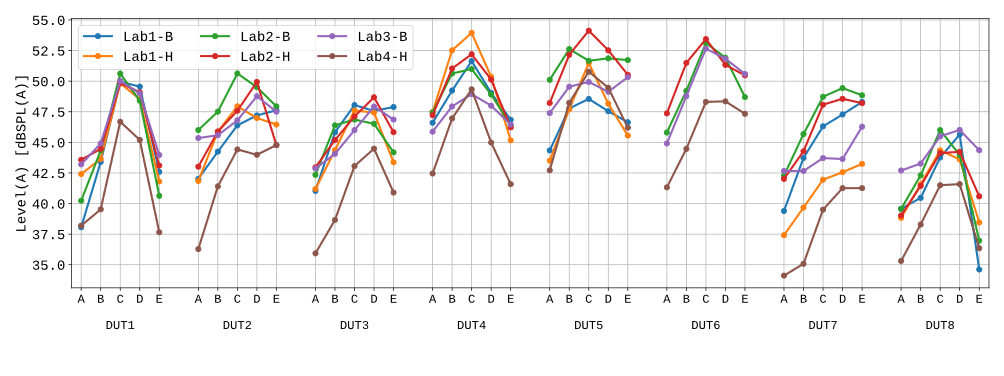


Figure 4-48: Level (A-weigthed) for privacy tests vs labs

As expected, for all devices the highest sound radiation is obtained at position C (microphone pointing directly to back side of the device). For some devices, the level of the test signal at the microphone position is rather high. One reason for this is the corresponding low RLR at the DUT (e.g., DUT3, DUT4). On the other hand, some devices with a low RLR at the DUT seem to radiate less than others (e.g., DUT3: similar as DUT4 / RLRmax of ~-13dB - but rather low level at the microphone). Vice versa, devices with moderate RLRmax values obtain higher levels at the microphone than devices with lower RLRmax values (e.g., DUT5, DUT6: RLRmax =~ -9dB, but highest overall levels).

The non-HaNTE device DUT8 obtains the lowest overall level at almost all angles/positions. However, the small level differences observed here do not indicate a sound radition issue for HaNTE devices: DUT8 also has one of the lowest RLRmax values, so a lower level is expected.

At first sight, it seems like those measurements from lab #4 are in general lower in level, even though the trend across angles and devices is similar. This can be explained by the rather low idle noise level, which is differently pronounced in the recordings. Table 4-21 provides the idle noise levels in dBSPL(A) for each lab, measured in a speech pause of the microphone recordings for DUT1.

Table 4-21: Idle Noise

|  |  |  |
| --- | --- | --- |
| Lab | Level [dBSPL (A)] | Comment |
| #1 | ~26 |  |
| #2 | ~21 |  |
| #3 | ~42 | Mainly quantization noise due to 16-bit usage and too high A/D conversion settings |
| #4 | ~15 |  |

To further investigate the psychoacoustic relevance of absolute speech levels and noise floor, the existing recordings were analyzed again with the loudness measures ISO 532-1  [11] and ITU-T P.700 [12]. Results of these two analyses are shown in Figure 4-49 (ISO 532-1) and Figure 4-50 (ITU-T P.700).

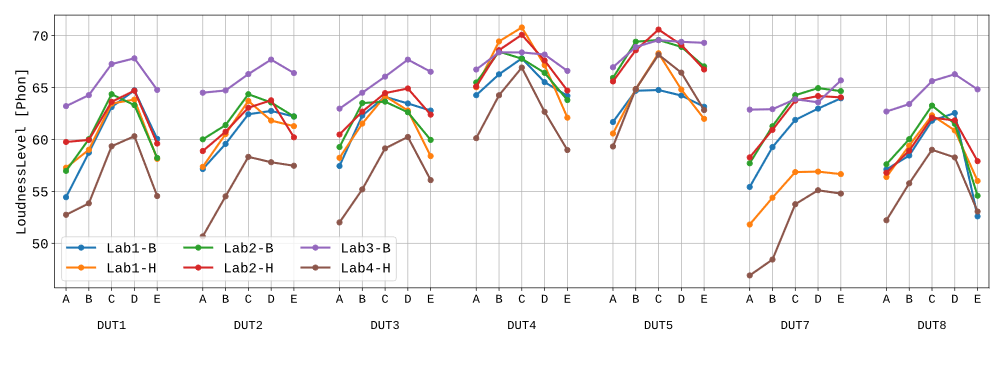


Figure 4-49: Loudness Level according to ISO 532-1 for privacy tests vs labs

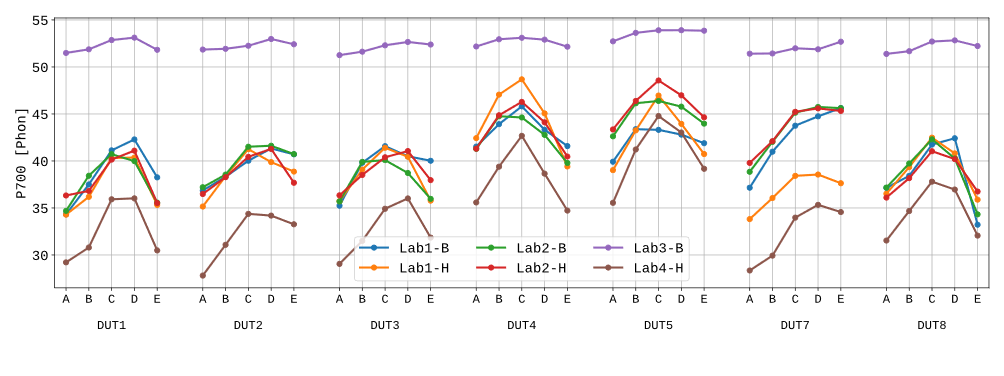


Figure 4-50: Loudness Level according to ITU-T P.700 for privacy tests vs labs

#### 4.5.4.4 Test 3a: Robustness (variation of ECRP)

Results of the RLR measurements for the robustness against shifts regarding (M)ECRP across labs are shown for each device in Figure 4-51 to Figure 4-58. Each figure only provides results for the long speech sequence. For sake of clarity, RLR values are clipped to the range -6 dB to 20 dB. Requirements of clause 6.2.2 of 3GPP TS 26.131 [5] (-1 to +5 dB for NOM volume setting), are indicated in the figures as well.

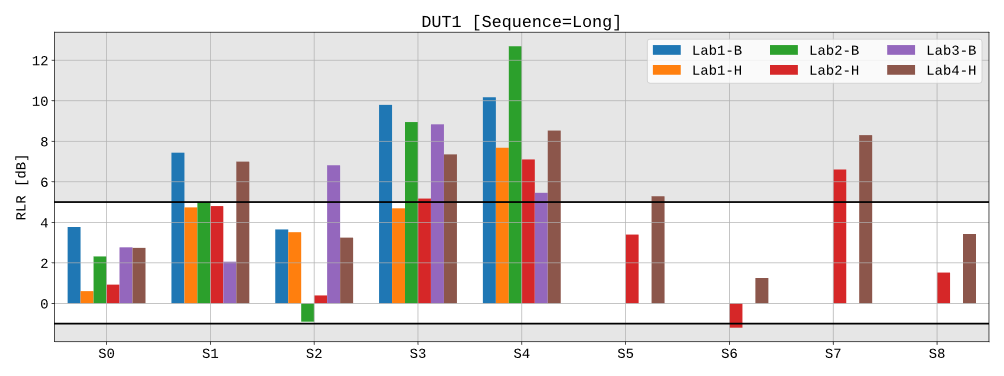


Figure 4-51: RLR Results for DUT1 vs shifts, sequences and labs

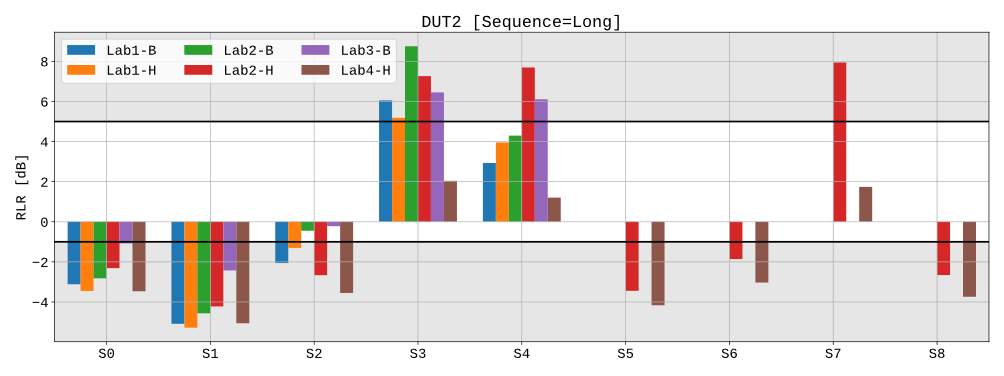


Figure 4-52: RLR Results for DUT2 vs shifts, sequences and labs

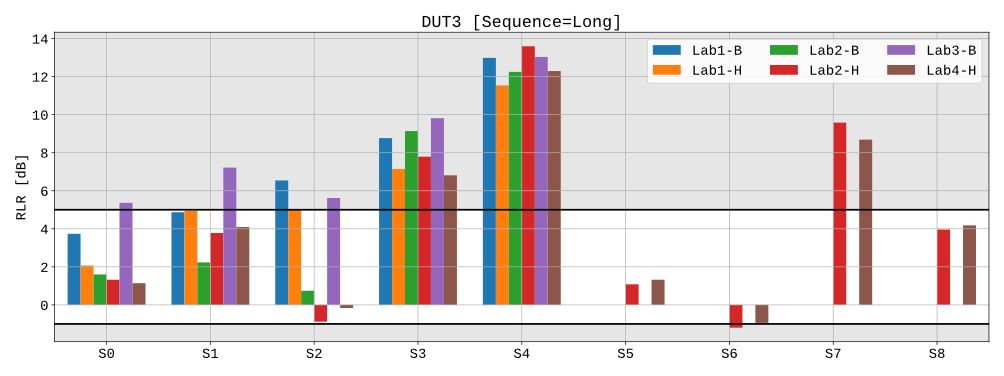


Figure 4-53: RLR Results for DUT3 vs shifts, sequences and labs

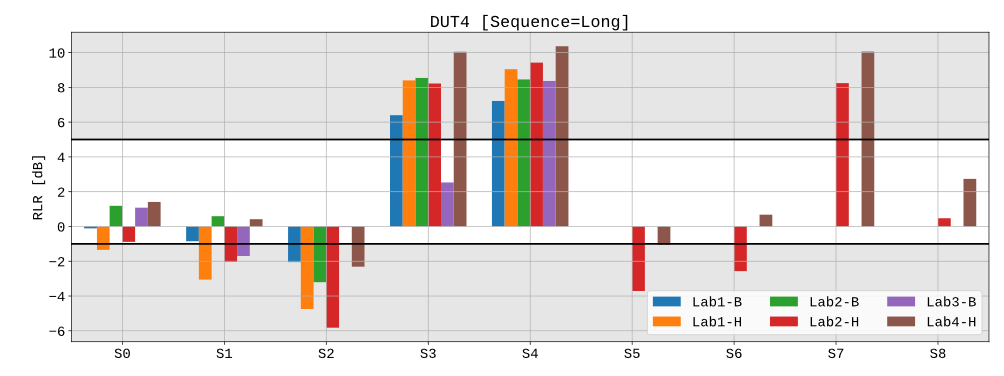


Figure 4-54: RLR Results for DUT4 vs shifts, sequences and labs

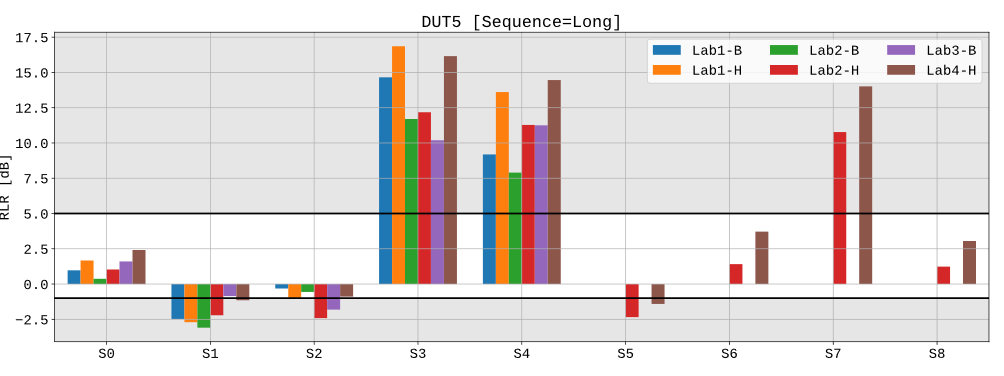


Figure 4-55: RLR Results for DUT5 vs shifts, sequences and labs

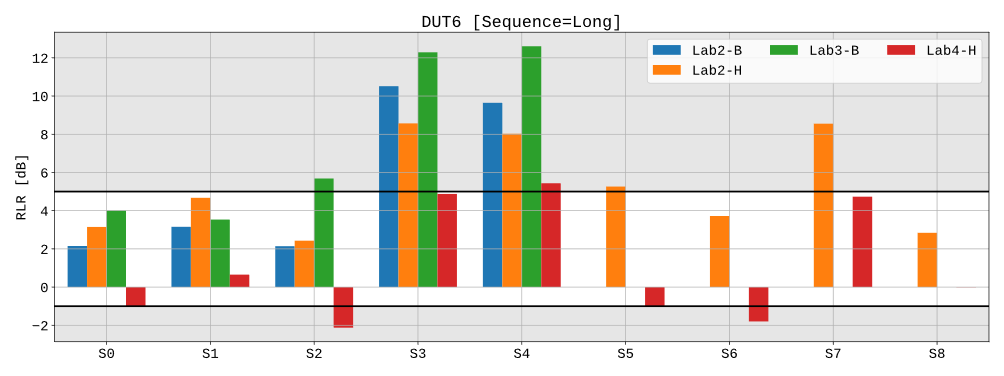


Figure 4-56: RLR Results for DUT6 vs shifts, sequences and labs

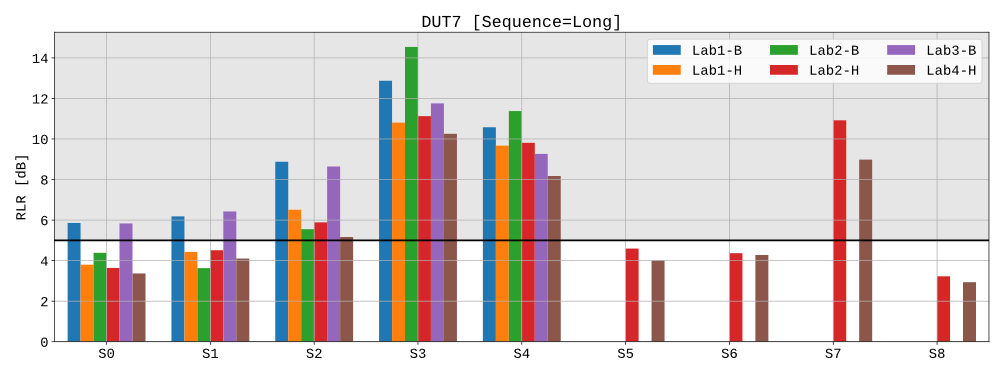


Figure 4-57: RLR Results for DUT7 vs shifts, sequences and labs

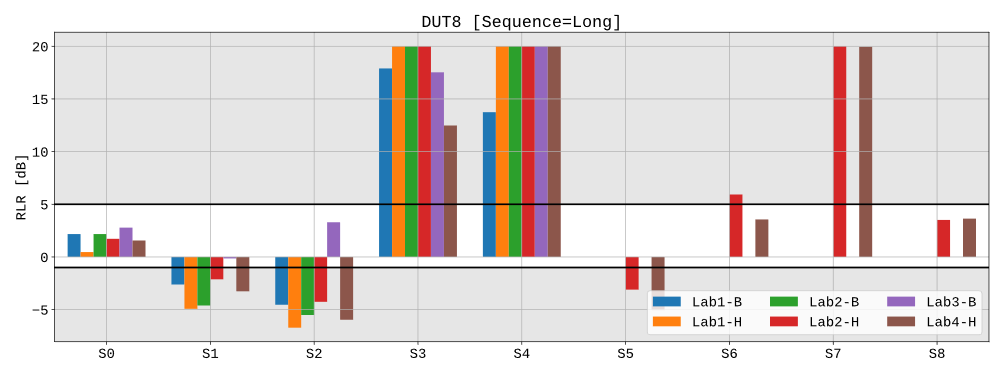


Figure 4-58: RLR Results for DUT8 vs shifts, sequences and labs

The left sides of Figure 4-59 to Figure 4-66 illustrate the RFR results for each device and each shift. Since many curves are provided per figure, for sake of clarity RFR is shown in 1/3rd octave bands. In both sub-figures, the tolerances according to clause 6.4.2 of 3GPP TS 26.131 [5] are provided for reference. Since only few data is available for the optional shifts, only S0 to S5 are considered in the following analyses.

While the left graphs are showing the robustness regarding positioning, the right sides of Figure 4-59 to Figure 4-66 provide a different analysis. Here the RFRs are first averaged across all shifts and are then reported for each lab. This can be interpreted as a kind of "robustness vs labs".

All frequency responses in each figure are adjusted with the same offset for better (overall) comparison and to remain level differences, which may be caused by the shifts. The single offset per figure is determined by the average curve to the upper tolerance, which is provided in both subplots as well (black dashed line).

|  |  |
| --- | --- |
|  |  |

Figure 4-59: RFR Results for DUT1, averaged vs labs (left) and vs shifts (right)

|  |  |
| --- | --- |
|  |  |

Figure 4-60: RFR Results for DUT2, averaged vs labs (left) and vs shifts (right)

|  |  |
| --- | --- |
|  |  |

Figure 4-61: RFR Results for DUT3, averaged vs labs (left) and vs shifts (right)

|  |  |
| --- | --- |
|  |  |

Figure 4-62: RFR Results for DUT4, averaged vs labs (left) and vs shifts (right)

|  |  |
| --- | --- |
|  |  |

Figure 4-63: RFR Results for DUT5, averaged vs labs (left) and vs shifts (right)

|  |  |
| --- | --- |
|  |  |

Figure 4-64: RFR Results for DUT6, averaged vs labs (left) and vs shifts (right)

|  |  |
| --- | --- |
|  |  |

Figure 4-65: RFR Results for DUT7, averaged vs labs (left) and vs shifts (right)

|  |  |
| --- | --- |
|  |  |

Figure 4-66: RFR Results for DUT8, averaged vs labs (left) and vs shifts (right)

Since the results for default RFR and RLR seem to include quite some variance across labs and also equipment (see clause 4.5.3.1, Figure 4-37 to Figure 4-44), this average measure may help to reduce such scatter. In most cases, RFR results here are a) similar to the ones of clause 4.5.3.1 and b) seem to be less variant. Considering all eight shift positions would possibly decrease the variance even more.

#### 4.5.4.5 Test 3b: Robustness (variation of fork position)

Results of the RLR measurements for the robustness against different fork positions are shown in Figure 4-67 for the long test sequence. Requirements of clause 6.2.2 of 3GPP TS 26.131 [5] (-1 to +5 dB for NOM volume setting) are indicated in the figures as well.

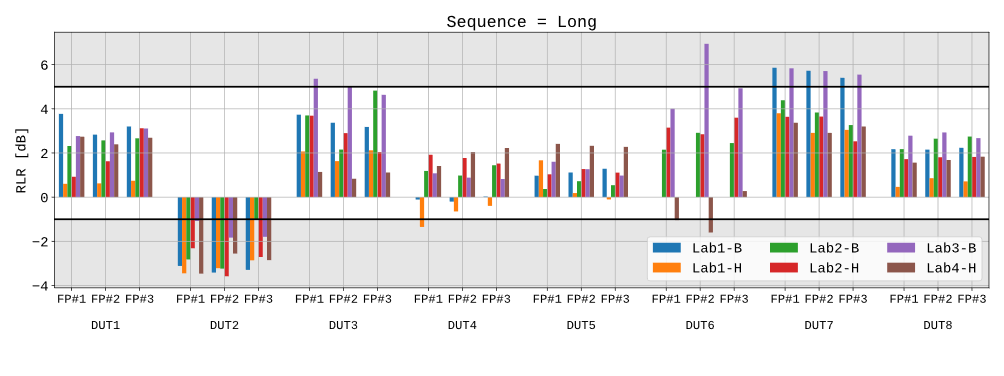


Figure 4-67: RLR Results for all DUTs vs fork positions (long sequence)

The results of the RFR measurements according to clause 8.4.2 of 3GPP TS 26.132 [6] (using only the long single talk sequence from ITU-T P.501 [9]) versus fork positions and labs are shown in Figure 4-68 to Figure 4-75 for all devices. To evaluate differences between labs and fork positions, an RFR averaged across fork positions and labs is additionally provided. For sake of clarity, curves are shown in 1/3rd octave bands.

In both sub-figures, the tolerances according to clause 6.4.2 of 3GPP TS 26.131 [5] are provided for reference. All frequency responses in each figure are adjusted with the same offset for better (overall) comparison and to remain level differences, which may be caused by the different fork positions. The single offset per figure is determined by adjusting the average curve to the upper tolerance, which is provided in both subplots as well (black dashed line).

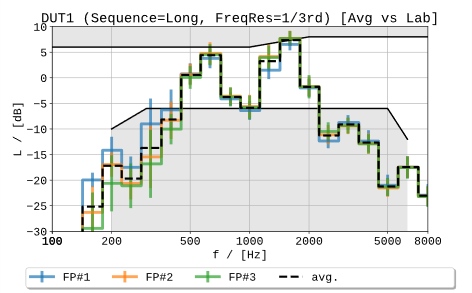
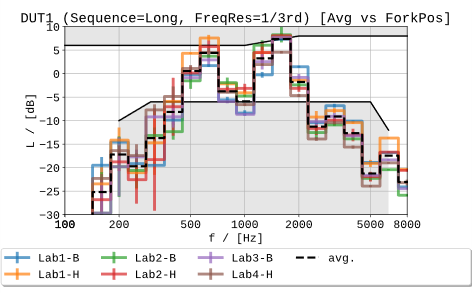


Figure 4-68: RFR Results for DUT1 average vs fork positions (left) and vs labs (right)

|  |  |
| --- | --- |
|  |  |

Figure 4-69: RFR Results for DUT2 average vs fork positions (left) and vs labs (right)

|  |  |
| --- | --- |
|  |  |

Figure 4-70: RFR Results for DUT3 average vs fork positions (left) and vs labs (right)

|  |  |
| --- | --- |
|  |  |

Figure 4-71: RFR Results for DUT4 average vs fork positions (left) and vs labs (right)

|  |  |
| --- | --- |
|  |  |

Figure 4-72: RFR Results for DUT5 average vs fork positions (left) and vs labs (right)

|  |  |
| --- | --- |
|  |  |

Figure 4-73: RFR Results for DUT6 average vs fork positions (left) and vs labs (right)

|  |  |
| --- | --- |
|  |  |

Figure 4-74: RFR Results for DUT7 average vs fork positions (left) and vs labs (right)

|  |  |
| --- | --- |
|  |  |

Figure 4-75: RFR Results for DUT8 average vs fork positions (left) and vs labs (right)

### 4.5.5 Discussion

Round robin test results from several combinations of lab/HATS were presented from four labs (with up two HATS/positioner types), which were obtained according to the agreed test plan (see clause 4.5.3). Several pecularities of the devices were identified and described to reduce effort for the subsequent labs in the round robin test. The following observations per sub-test were made:

Test 1 (Speech Quality)

Due to the low performance in frequency response for NOM and/or MAX, some devices cannot provide constant and acceptable quality across volume settings. However, since also good performance was achieved by other devices (e.g., DUT3 meets RFR requirement and MOS > 3.5), it seems in general possible to provide acceptable quality for HaNTE-devices.

Test 2 (Privacy)

After reviewing (Delta-)RLR results, an absolute level measure was found to be more suitable for the evaluation of direction-dependent sound radiation. Levels measured at the microphone between ~34 dBSPL / dBSPL(A) and 53-54 dBSPL / dBSPL(A) were obtained. The rather high variance in the results per DUT can be explained by the quite different idle noise floor, which depends highly on the measurement equipment/configuration, measurement room and/or microphone used.

However, there are noticeable differences between the devices and some reach about 50 dBSPL/dBSPL(A), which is rather soft, but already audible. To investigate the degree of sound radiation with regard to psychoacoustics/intelligibility, also loudness levels according to ISO 532-1 and ITU-T P.700 were provided.

Test 3a (Robustness vs. shifts)

Large differences in performance regarding RLR and RFR could be observed, depending on the shifts and devices. Only few devices can be regarded as robust against shifts.

The average measure across shift positions seem to reduce the scatter versus labs and provids better consistency between labs. This approach should be investigated more in detail (for example, if the same result can be achieved with the short sequences).

Test 3b (Robustness vs. fork positions)

Most devices perform quite similar regarding the different fork positions. However, some devices show high deviations here, which refer back to issues in the corresponding frequency responses as well.

# 5 Conclusion

The present document reports on investigations of testing UEs featuring non-traditional earpieces ("HaNTE device") and identifies related gaps to existing 3GPP specifications and recommended test equipment. The feasibility study resulted in the following main findings:

- UEs featuring non-traditional earpieces may not feature a center of an acoustic port, raising the question of how to properly position a handset for testing. Therefore, an update to ITU-T Recommendation P.64 was found required to reference this Recommendation also for handsets with non-traditional earpieces.

- When measured with manufacturer provided MECRP and mounting instructions, at least one HaNTE device was found compliant to current 3GPP terminal acoustic specifications.

- However, even when measured with manufacturer provided MECRP and mounting instructions, certain HaNTE devices were found not compliant to current 3GPP terminal acoustic specifications.

- In contrast to devices featuring a traditional earpiece, the manner of holding the device with the handset positioner was found of significance for the receive frequency response and receive loudness rating of certain HaNTE devices. The mechanical contact of the handset positioner to the HaNTE device structure may alter the device's acoustic radiating properties.

- Because a significant portion of the display vibrates and radiates sound, the listening "sweet spot" may be larger for certain HaNTE devices when compared to handsets featuring a traditional earpiece.

- Privacy was found of concern for certain HaNTE devices studied, with audible sound radiating to areas in proximity to the user.

- High distortion was identified as a concern with certain HaNTE devices studied.

- Out of band noise was identified as a concern with certain HaNTE devices studied.

A round robin test in four different labs was conducted, which targeted at introducing possible new test methods for 3GPP TS 26.132 and performance requirements for 3GPP TS 26.131. The round robin resulted in the following main findings:

- Similar as in previous studies, it was noted that certain HaNTE devices are in general able to achieve a sufficient level of quality (ITU-T P.863 prediction of ~3.5 MOS). New test methods for 3GPP TS 26.132 (and in consequence, also requirements for 3GPP TS 26.131) were not considered to be necessary.

- The sound radiation of HaNTE devices was identified to be higher than for traditional handsets. However, since it could not be determined if this leads to an issue regarding privacy, i.e., if the receive signal is intelligible at the distant position, new test methods for 3GPP TS 26.132 (and in consequence, also requirements for 3GPP TS 26.131) were not considered to be necessary.

- The results of the robustness experiments (variation of ECRP) were found to be relevant for positioning HaNTE devices in cases where no MECRP is provided. An iterative procedure in 3GPP TS 26.132 was specified to identify the most suitable testing position for such handsets.

- The results of the robustness experiments (variation of fork position) indicated that only a few devices seem to have issues with different fork positions, which are also reflected in corresponding RFR measurements. Thus, new test methods for 3GPP TS 26.132 (and in consequence, also requirements for 3GPP TS 26.131) were not considered to be necessary.

Annex A:  
Tables for the study on suitability of HATS for HaNTE measurements

Table A-1: Subjective level matching descriptive statistics

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Center Frequency (Hz) | N | Mean (dB) | Std. Deviation | Std. Error |
| 200 | 16 | -8.91 | 5.25 | 1.31 |
| 250 | 16 | -2.75 | 3.92 | 0.98 |
| 315 | 16 | -1.00 | 3.30 | 0.82 |
| 400 | 16 | 0.69 | 3.39 | 0.85 |
| 500 | 16 | 1.41 | 3.67 | 0.92 |
| 630 | 16 | -2.13 | 3.61 | 0.90 |
| 800 | 16 | -3.88 | 4.82 | 1.21 |
| 1000 | 16 | -3.69 | 5.33 | 1.33 |
| 1250 | 16 | -4.75 | 3.95 | 0.99 |
| 1600 | 16 | -5.84 | 3.50 | 0.88 |
| 2000 | 16 | -7.91 | 3.11 | 0.78 |
| 2500 | 16 | -6.69 | 3.47 | 0.87 |
| 3150 | 16 | -6.72 | 4.32 | 1.08 |
| 4000 | 16 | -4.44 | 3.05 | 0.76 |
| 5000 | 16 | -3.75 | 3.29 | 0.82 |
| 6300 | 16 | -2.66 | 3.91 | 0.98 |
| 8000 | 16 | -4.16 | 3.50 | 0.87 |
| 10000 | 16 | -2.81 | 4.17 | 1.04 |
| 12500 | 16 | -3.09 | 3.84 | 0.96 |

Table A-2: Objective-Subjective correlation analysis

|  |  |  |  |
| --- | --- | --- | --- |
| Center Frequency (Hz) | HATS 1 (dB) | HATS 2 (dB) | Subjective Mean (dB) |
| 200 | -6.93 | -5.06 | -8.91 |
| 250 | -3.63 | -3.04 | -2.75 |
| 315 | 0.09 | -0.58 | -1.00 |
| 400 | 3.28 | 2.97 | 0.69 |
| 500 | -1.58 | 0.04 | 1.41 |
| 630 | -1.99 | -4.71 | -2.13 |
| 800 | -0.81 | -5.48 | -3.88 |
| 1000 | -5.93 | -5.22 | -3.69 |
| 1250 | -5.85 | -3.92 | -4.75 |
| 1600 | -4.39 | -6.51 | -5.84 |
| 2000 | -5.23 | -7.30 | -7.91 |
| 2500 | -6.44 | -7.05 | -6.69 |
| 3150 | -6.43 | -8.80 | -6.72 |
| 4000 | -5.80 | -7.40 | -4.44 |
| 5000 | -4.64 | -5.76 | -3.75 |
| 6300 | -3.94 | -0.87 | -2.66 |
| 8000 | -3.23 | -0.61 | -4.16 |
| 10000 | -3.21 | -4.14 | -2.81 |
| 12500 | -0.05 | 5.98 | -3.09 |
| rho | 0.775 | 0.656 | - |
| tstat | 5.350 | 3.791 | - |
| p | 0.000 | 0.001 | - |

Table A-3: SE and RMSE calculated between subjective results and HATS data

|  |  |  |
| --- | --- | --- |
| Center Frequency (Hz) | HATS 1 (dB) | HATS 2 (dB) |
| 200 | -6.93 | -5.06 |
| 250 | -3.63 | -3.04 |
| 315 | 0.09 | -0.58 |
| 400 | 3.28 | 2.97 |
| 500 | -1.58 | 0.04 |
| 630 | -1.99 | -4.71 |
| 800 | -0.81 | -5.48 |
| 1000 | -5.93 | -5.22 |
| 1250 | -5.85 | -3.92 |
| 1600 | -4.39 | -6.51 |
| 2000 | -5.23 | -7.30 |
| 2500 | -6.44 | -7.05 |
| 3150 | -6.43 | -8.80 |
| 4000 | -5.80 | -7.40 |
| 5000 | -4.64 | -5.76 |
| 6300 | -3.94 | -0.87 |
| 8000 | -3.23 | -0.61 |
| 10000 | -3.21 | -4.14 |
| 12500 | -0.05 | 5.98 |
| RMSE | 1.794 | 2.832 |

Table A-4: One-sample, two-tailed t-test with HATS 1 as reference

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Center Frequency (Hz) | Test Value | t | df | Sig. (2-tailed) | Mean Difference | 95% Confidence Interval | |
| Lower | Upper |
| 200 | -6.93 | -1.504 | 15 | 0.153 | -1.976 | -11.706 | -6.106 |
| 250 | -3.63 | 0.898 | 15 | 0.383 | 0.880 | -4.839 | -0.661 |
| 315 | 0.09 | -1.323 | 15 | 0.206 | -1.090 | -2.757 | 0.757 |
| 400 | 3.28 | -3.059 | 15 | 0.008\* | -2.593 | -1.119 | 2.494 |
| 500 | -1.58 | 3.258 | 15 | 0.005\* | 2.986 | -0.547 | 3.360 |
| 630 | -1.99 | -0.150 | 15 | 0.883 | -0.135 | -4.047 | -0.203 |
| 800 | -0.81 | -2.543 | 15 | 0.023\* | -3.065 | -6.444 | -1.306 |
| 1000 | -5.93 | 1.683 | 15 | 0.113 | 2.243 | -6.527 | -0.848 |
| 1250 | -5.85 | 1.114 | 15 | 0.283 | 1.100 | -6.855 | -2.645 |
| 1600 | -4.39 | -1.661 | 15 | 0.117 | -1.454 | -7.709 | -3.978 |
| 2000 | -5.23 | -3.442 | 15 | 0.004\* | -2.676 | -9.564 | -6.249 |
| 2500 | -6.44 | -0.285 | 15 | 0.779 | -0.248 | -8.536 | -4.839 |
| 3150 | -6.43 | -0.268 | 15 | 0.793 | -0.289 | -9.019 | -4.419 |
| 4000 | -5.80 | 1.788 | 15 | 0.094 | 1.363 | -6.062 | -2.813 |
| 5000 | -4.64 | 1.083 | 15 | 0.296 | 0.890 | -5.501 | -1.999 |
| 6300 | -3.94 | 1.315 | 15 | 0.208 | 1.284 | -4.738 | -0.575 |
| 8000 | -3.23 | -1.060 | 15 | 0.306 | -0.926 | -6.019 | -2.293 |
| 10000 | -3.21 | 0.382 | 15 | 0.708 | 0.398 | -5.033 | -0.592 |
| 12500 | -0.05 | -3.175 | 15 | 0.006\* | -3.044 | -5.137 | -1.050 |

Table A-5: One-sample, two-tailed t-test with HATS 2 as reference

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Center Frequency (Hz) | Test Value | t | df | Sig. (2-tailed) | Mean Difference | 95% Confidence Interval | |
| Lower | Upper |
| 200 | -5.06 | -2.928 | 15 | 0.010\* | -3.846 | -11.706 | -6.106 |
| 250 | -3.04 | 0.296 | 15 | 0.771 | 0.290 | -4.839 | -0.661 |
| 315 | -0.58 | -0.510 | 15 | 0.618 | -0.420 | -2.757 | 0.757 |
| 400 | 2.97 | -2.693 | 15 | 0.017\* | -2.283 | -1.119 | 2.494 |
| 500 | 0.04 | 1.491 | 15 | 0.157 | 1.366 | -0.547 | 3.360 |
| 630 | -4.71 | 2.866 | 15 | 0.012\* | 2.585 | -4.047 | -0.203 |
| 800 | -5.48 | 1.331 | 15 | 0.203 | 1.605 | -6.444 | -1.306 |
| 1000 | -5.22 | 1.150 | 15 | 0.268 | 1.533 | -6.527 | -0.848 |
| 1250 | -3.92 | -0.841 | 15 | 0.414 | -0.830 | -6.855 | -2.645 |
| 1600 | -6.51 | 0.761 | 15 | 0.458 | 0.666 | -7.709 | -3.978 |
| 2000 | -7.30 | -0.780 | 15 | 0.448 | -0.606 | -9.564 | -6.249 |
| 2500 | -7.05 | 0.418 | 15 | 0.682 | 0.363 | -8.536 | -4.839 |
| 3150 | -8.80 | 1.929 | 15 | 0.073 | 2.081 | -9.019 | -4.419 |
| 4000 | -7.40 | 3.887 | 15 | 0.001\* | 2.963 | -6.062 | -2.813 |
| 5000 | -5.76 | 2.446 | 15 | 0.027\* | 2.010 | -5.501 | -1.999 |
| 6300 | -0.87 | -1.829 | 15 | 0.087 | -1.786 | -4.738 | -0.575 |
| 8000 | -0.61 | -4.057 | 15 | 0.001\* | -3.546 | -6.019 | -2.293 |
| 10000 | -4.14 | 1.274 | 15 | 0.222 | 1.328 | -5.033 | -0.592 |
| 12500 | 5.98 | -9.464 | 15 | 0.000\* | -9.074 | -5.137 | -1.050 |

Annex B:  
Change history

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Change history** | | | | | | | |
| **Date** | **Meeting** | **TDoc** | **CR** | **Rev** | **Cat** | **Subject/Comment** | **New version** |
| 2019-09 | SA#85 | SP-190645 |  |  |  | Presented to TSG SA#85 for approval | 1.0.0 |
| 2019-09 | SA#85 |  |  |  |  | Approved at TSG SA#85 | 16.0.0 |
| 2023-12 | SA#102 |  |  |  |  | Update to Rel-17 version (MCC) | 17.0.0 |
| 2023-12 | SA#102 | SP-231370 | 0001 | 2 | B | CR on HaNTE round robin test results in TR 26.801 | 18.0.0 |