ETSI TS 103 220 V1.1.1 (2014-03)

Technical Specification

**Speech and multimedia Transmission Quality (STQ);
Transmission requirements for Superwideband handheld
(handset and handsfree) terminals from a QoS perspective
as perceived by the user**

Reference
DTS/STQ-218

Keywords
QoS, speech, terminal

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://ipr.etsi.org).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Specification (TS) has been produced by ETSI Technical Committee Speech and multimedia Transmission Quality (STQ).

Introduction

Speech terminals are currently implementing narrowband and wideband. Nowadays, terminal equipment may offer superwideband, thanks to features already available in these terminals. Such equipment may implement conversational features that may be to the benefit of the electro acoustic equipment already available in the terminal and may provide wider quality for the end users.

The present document is intended to provide initial requirements and test methods for such equipment. Futures releases will take benefit of new requirements and test methods provided by TS 102 924 [17] and TS 102 925 [24].

1 Scope

The present document provides speech transmission performance requirements and measurement methods for handset and handsfree functions of superwideband handheld terminals, and requirements in order to optimize the end to end quality perceived by users.

Standardized superwideband coders are now available, some being also compatible with wideband coders.

The present document considers only conversational services (that may be mixed with other services) and does not cover the streaming-only services. Such applications include:

- Speech communication for handset and handsfree functions. Special care is taken to ensure that the quality offered both by handset and handsfree functions is equivalent and that the quality in send direction is similar to superwideband terminals fulfilling TS 102 924 [17] or TS 102 925 [24].
- Bandwidth extension which may allow usage for some mixed content applications. The frame of these applications is in the context of the mix found in ES 202 396-1 [10].

NOTE: Requirements and measurement methods for the headset function associated to handheld terminals can be found in TS 102 924 [17].

2 References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the reference document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

2.1 Normative references

The following referenced documents are necessary for the application of the present document.

[1] Recommendation ITU-T P.501 Amendment 1: "Test signals for use in telephonometry".

[2] Recommendation ITU-T P.10/G.100: "Vocabulary for performance and quality of service".

[3] Recommendation ITU-T P.58: "Head and torso simulator for telephonometry".

[4] Recommendation ITU-T P.581: "Use of head and torso simulator (HATS) for hands-free and handset terminal testing".

[5] Recommendation ITU-T P.79: "Calculation of loudness ratings for telephone sets".

[6] Recommendation ITU-T P.340: "Transmission characteristics and speech quality parameters of hands-free terminals".

[7] Recommendation ITU-T G.722.1 (Annex C): "Low-complexity coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss".

[8] Recommendation ITU-T G.729.1 (Annex E): "G.729-based embedded variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable with G.729".

[9] Recommendation ITU-T G.718 (Annex B): "Frame error robust narrow-band and wideband embedded variable bit-rate coding of speech and audio from 8-32 kbit/s".

[10] ETSI ES 202 396-1: "Speech and multimedia Transmission Quality (STQ); Speech quality performance in the presence of background noise; Part 1: Background noise simulation technique and background noise database".

[11] ETSI ES 202 740: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for wideband VoIP loudspeaking and handsfree terminals from a QoS perspective as perceived by the user".

[12] ETSI TS 103 740: "Speech and multimedia Transmission Quality (STQ);Transmission requirements for wideband wireless terminals (handsfree) from a QoS perspective as perceived by the user".

[13] ETSI ES 202 739: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for wideband VoIP terminals (handset and headset) from a QoS perspective as perceived by the user".

[14] Recommendation ITU-T P.863: "Perceptual objective listening quality assessment".

[15] Recommendation ITU-T G.711.1: "Wideband embedded extension for G.711 pulse code modulation".

[16] ETSI TS 103 739: "Speech and multimedia Transmission Quality (STQ);Transmission requirements for wideband wireless terminals (handset and headset) from a QoS perspective as perceived by the user".

[17] ETSI TS 102 924: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for Superwideband/Fullband headset terminals from a QoS perspective as perceived by the user".

[18] Recommendation ITU-T P.800: "Methods for subjective determination of transmission quality".

[19] Recommendation ITU-T P.830: "Subjective performance assessment of telephone-band and wideband digital codecs".

[20] Recommendation ITU-T G.722: "7 kHz audio-coding within 64 kbit/s".

[21] Recommendation ITU-T P.56: "Objective measurement of active speech level".

[22] ISO 3 (1973): "Preferred numbers -- Series of preferred numbers".

[23] ISO 3745: "Acoustics -- Determination of sound power levels and sound energy levels of noise sources using sound pressure -- Precision methods for anechoic rooms and hemi-anechoic rooms".

[24] ETSI TS 102 925: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for Superwideband/Fullband handsfree and conferencing terminals from a QoS perspective as perceived by the user".

[25] Recommendation ITU-T P.64: "Determination of sensitivity/frequency characteristics of local telephone systems".

[26] Recommendation ITU-T P.57: "Artificial ears".

2.2 Informative references

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

Not applicable.

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

handheld terminal: terminal intended to be held in hand, which includes either the handset function or the handsfree function or both

handset: terminal coupled to the ear by hand

handsfree: terminal in handsfree mode while terminal is hand held

superwideband telephony: transmission of speech with a nominal pass-band wider than 100 Hz to 7 000 Hz, usually understood to be 50 Hz to 14 000 Hz (definition from Recommendation ITU-T P.10/G.100 [2])

NOTE: Superwideband covers at least mono and stereo capabilities.

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

ACR	Absolute Category Rating
CSS	Composite Source Signal
EVS	Enhanced Voice Services
GAT	Group Audio Terminal
HATS	Head and Torso Simulator
HhHFRP	Handheld HandsFree Reference Point
HFRP	HandsFree Reference Point
MCU	Multiplexing Control Unit
MRP	Mouth Reference Point
PDA	Personal Digital Assistant
POI	Point of Interconnection
RLR	Receive Loudness Rating
SLR	Send Loudness Rating
SWB	Superwideband

4 Applications and Coder considerations

4.1 Applications

The following applications are within the scope of the present document:

- Bandwidth extension which may allow usage for some mixed content applications where wider bandwidth could bring a significant added value for the customer (support of 14 kHz bandwidth).

The send path can be characterized in two ways:

- The signal picked up by microphone(s) may combine speech, music and every type of environmental signal.

NOTE: For some applications (e.g. journalist reporting) the user should have the possibility to cancel the noise environment or to transmit it without degradation.

- Direct insertion of any type of signal.

For receive path, the signal may combine the two following types:

- Communication signal such as described for send path.

- Signal coming from distributed applications (e.g. advertisement, music on hold, etc.).

4.2 Coder considerations

As indicated in the scope only coders supporting conversational SWB services are applicable to the present document.

Table 4.2

Coder Reference	Speech	Other signals	Stereo	Remark
Recommendation ITU-T G.722.1 [7] Annex C	X	X Music		For low frame loss
Recommendation ITU-T G.729.1 [8] Annex E (extension SWB)	X	X background noise (X) music		
Recommendation ITU-T G.718 [9] Annex B	X	X Music		
Recommendation ITU-T G.711.1 [15] Annexes D and F	X	X	X (Annex F)	
Recommendation ITU-T G.722 [20] Annexes B and D	X	X	X (Annex D)	
NOTE: Recommendation ITU-T G.722.1 [7] is intended to be used for hand-free application.				

When X is in brackets, it means that the coder is not optimized for this application.

The following coders are recommended for superwideband:

- Recommendation ITU-T G.722.1 [7] Low-complexity coding at 24 kbit/s and 32 kbit/s for handsfree operation in systems with low frame loss. Annex C 14 kHz mode at 24 kbit/s, 32 kbit/s and 48 kbit/s.
  - The algorithm is recommended for use in handsfree applications such as conferencing where there is a low probability of frame loss. It may be used with speech or music inputs. The bit rate may be changed at any 20 ms frame boundary. New Annex C contains the description of a low-complexity extension mode to G.722.1 [7], which doubles the algorithm to permit 14 kHz audio bandwidth using a 32 kHz audio sample rate, at 24 kbit/s, 32 kbit/s and 48 kbit/s.
  - Annex C. This annex provides a description of the 14 kHz mode at 24 kbit/s, 32 kbit/s and 48 kbit/s for this Recommendation.
- Recommendation ITU-T G.729.1 [8], Annex E (extension SWB for G.729.1).
  - This annex provides the high-level description of the higher bit-rate extension of G.729 designed to accommodate a wide range of input signals, such as speech, with background noise and even music.
- Recommendation ITU-T G.718 [9], Annex B Superwideband scalable extension for Recommendation ITU-T G.718 [9]) *"This annex describes a scalable superwideband (SWB, 50-14000 Hz) speech and audio coding algorithm operating from 36 to 48 kbit/s and interoperable with Recommendation ITU-T G.718 [9]."*
- Recommendation ITU-T G.711.1 [15], Annex D defines the superwideband extension.
  - Annex F defines the Stereo embedded extension for Recommendation ITU-T G.711.1 [15].

- *"The Annex F is intended as a stereo extension to the G.711.1 wideband coding algorithm and its superwideband Annex D. Compared to discrete two-channel (dual-mono) audio transmission, this stereo extension G.711.1, Annex F saves valuable bandwidth for stereo transmission. It is specified to offer the stereo capability while providing backward compatibility with the monaural core in an embedded scalable way. The Annex provides very good quality for stereo speech contents (clean speech and noisy speech with various stereo sound pickup systems: binaural, MS, etc.), and for most of the conditions it provides significantly higher quality than low bitrate dual-mono. For some music contents, e.g. highly reverberated and/or with diffuse sound, the algorithm may have some performance limitations and may not perform as good as dual-mono codecs, however it achieves the quality of state-of-the-art parametric stereo codecs."*

- Recommendation ITU-T G.722 [20], Annex B defines the superwideband extension and Annex D defines the Stereo embedded extension for Recommendation ITU-T G.722 [20].

  - *"Annex B describes a scalable superwideband (SWB, 50-14 000 Hz) speech and audio coding algorithm operating at 64, 80 and 96 kbit/s. The Recommendation ITU-T G.722 [20] superwideband extension codec is interoperable with Recommendation ITU-T G.722 [20]. The output of the Recommendation ITU-T G.722 [20] SWB coder has a bandwidth of 50-14 000 Hz."*

  - *"Annex D describes a stereo extension of the wideband codec G.722 and its superwideband extension, G.722 Annex B. It is optimized for the transmission of stereo signals with limited additional bitrate, while keeping full compatibility with both codecs. Annex D operates from 64 to 128 kbit/s with four superwideband stereo bitrates at 80, 96, 112 and 128 kbit/s and two wideband stereo bitrates at 64 and 80 kbit/s".*

5 Test considerations

The terminals within the scope of the present document are not only dedicated to speech communication but are also mixing speech and audio contents and may implement stereo and multichannel transmissions.

5.1 Test Set-ups

5.1.1 Test interfaces

Handsfree

Recommendation ITU-T P.58 [3] indicates:

- *"The artificial ears ... support super-wideband ... applications. It should be noted that the acoustical impedance of the artificial ears has some limitations in realistically simulating human ears".*
- *"The artificial mouth supports super-wideband applications, however it should be noted that the directionality of the artificial mouth is limited in its ability to simulate the human mouth in the super-wideband frequency range."*

All the send and receive characteristics shall be tested with the HATS, it shall be indicated what type of ear was used.

The preferred way of testing a terminal is to connect it to a network simulator with exact defined settings and access points. The test sequences are fed in either electrically, using a reference codec (at least implementing the bit rate offering the best quality for the coder) or using the direct signal processing approach or acoustically.

When, a coder with variable bite rate is used, we should adopt, for testing terminal electro acoustical parameters, the highest bit rate which is recognized as providing the best characteristics is selected.

Handset

All the send and receive characteristics shall be tested with the HATS, it shall be indicated what type of ear was used.

The horizontal positioning of the HATS reference plane shall be guaranteed within ±2º.

The HATS shall be equipped with type 3.3 artificial ear. The type 3.3 artificial ear as specified in Recommendation ITU-T P.57 [26] shall be used. The artificial ear shall be positioned on HATS according to Recommendation ITU-T P.58 [3].

The exact calibration and equalization can be found in Recommendation ITU-T P.581 [4].

The artificial mouth has to be equalized between 50 Hz and 16 kHz.

NOTE: For some artificial mouth the equalization in the complete frequency range is not possible. Therefore the equalization is limited to 100 Hz - 14 kHz which will limit the frequency range of all measurements in sending where the test signal is generated by the artificial mouth to this frequency range.

For receive, if not stated otherwise, the HATS shall be diffuse-field equalized. The reverse nominal diffuse-field curve as found in table 3 of Recommendation ITU-T P.58 [3] shall be used.

5.1.2 Setup for terminals

As the scope of the present document includes all the potential types of handsfree terminals this clause defines the set up for each type of terminal.

5.1.2.1 Handset terminal

The preferred acoustical access to terminals is the most realistic simulation of the "average" subscriber. This can be made by using Head And Torso Simulator (HATS) with appropriate ear simulation and appropriate means to fix handset terminals in a realistic and reproducible way to the HATS. HATS is described in Recommendation ITU-T P.58 [3], appropriate ears are described in Recommendation ITU-T P.57 [26] (type 3.3 and type 3.4 ear), a proper positioning of handsets under realistic conditions is to be found in Recommendation ITU-T P.64 [25].

The preferred way of testing a terminal is to connect it to a network simulator with exact defined settings and access points. The test sequences are fed in either electrically, using a reference codec or using the direct signal processing approach or acoustically using ITU-T specified devices.

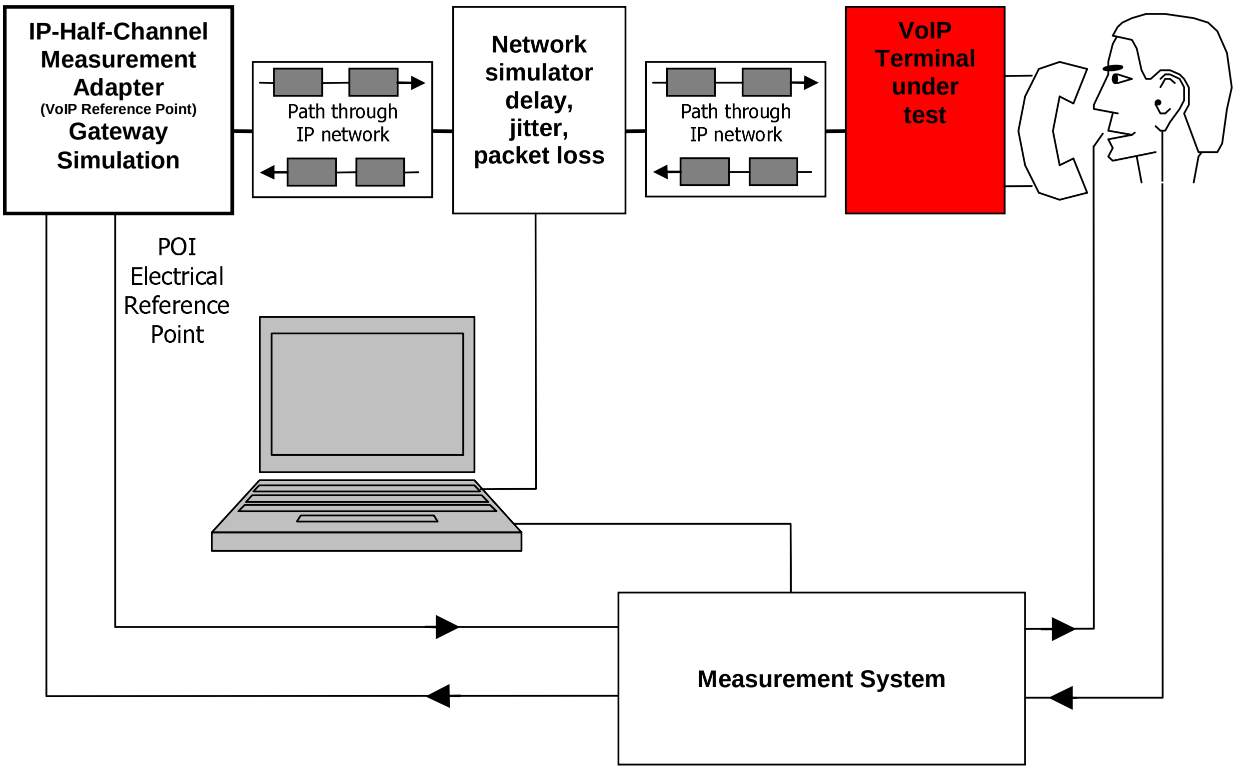


Figure 5.1.2.1: Half channel terminal measurement

5.1.2.2 Handsfree terminal

This kind of terminal could implement SWB, The test configuration is defined on figure 5.1.2.2.

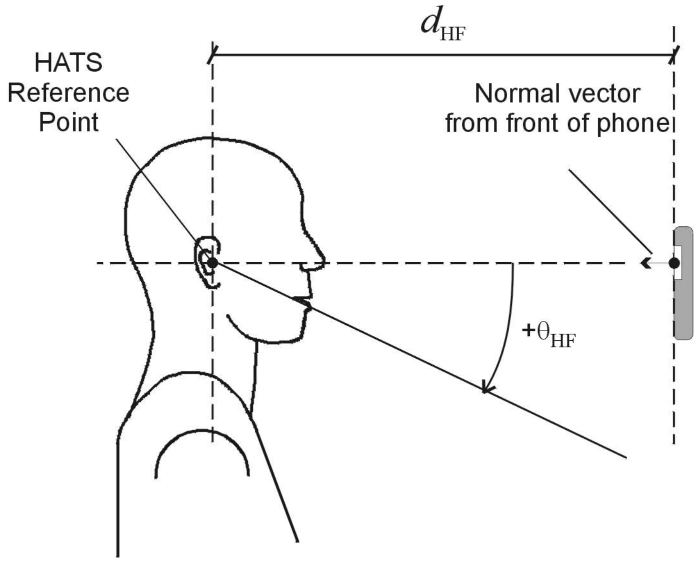


Figure 5.1.2.2: Configuration of Handheld loudspeaker relative to the HATS side view

5.1.3 Test signals

The test signals are defined according to Recommendation ITU-T P.501 Amendment 1 [1] for test made with speech signals. For some parameters it is needed to combine speech signals with other types of signals (e.g. music, background noise) or the test signal may be an audio signal mixing any type of materials. Such signals are defined in ES 202 396-1 [10].

As the bandwidth of the speech signals defined in Recommendation ITU-T P.501 [1], Amendment 1, is fullband, these test signals shall be used in the present document:

- The test signal to be used for measurements such as frequency response and loudness rating, shall be the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [1], Amendment 1.
- The female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [1], Amendment 1, shall be used as activation signal for measurements such as distortion and send noise.
- The compressed real speech signal described in clause 7.3.3 of Recommendation ITU-T P.501 [1], Amendment 1, shall be used for measurements such as TCLw, switching characteristics.

For double-talk performance:

- A "double-talk" sequence representing typical double talk scenarios in real conversations is shown in figure 6.3.4. This uses the single-talk sequence described in section 7.3.1 of Recommendation ITU-T P.501 [1], Amendment 1, shown in the lower pane, as the main speech and an additional competing speaker sequence, shown in the upper pane.

5.1.4 Test signal levels

The level dependency should be considered and consequently tests should also be done with signal levels lower and higher than the reference level defined in the following clauses.

5.1.4.1 Send

Handset

Unless specified otherwise, the applied test signal level shall be -4,7 dBPa.

Handsfree

Unless specified otherwise, the test signal shall be calibrated (including equalization) at HhHFRP (located at 30 cm from the lip ring) according to Recommendation ITU-T P.581 [4].

When using a reference loudspeaker its centre is positioned at the lip ring position. The loudspeaker is intended to be free-field equalized.

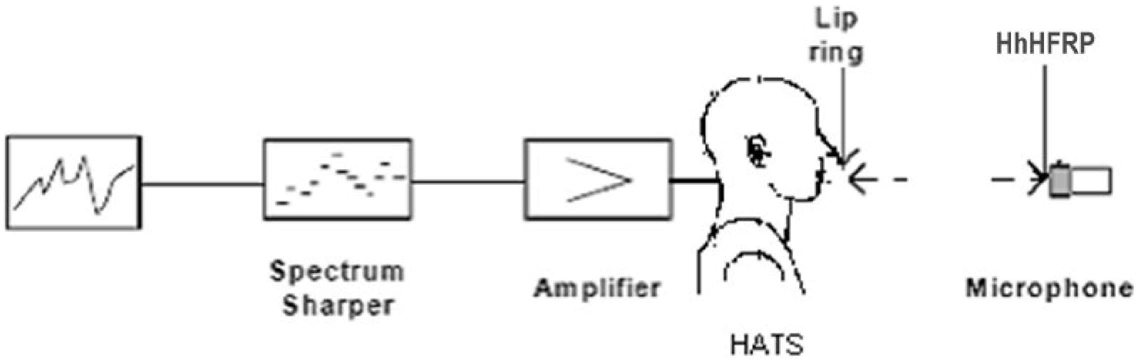


Figure 5.1.4.1: Calibration at HhHFRP (with d_S = 30 cm)

For the distance of 30 cm (HhHFRP) the calibration level shall be adjusted to -24,3 dBPa.

NOTE: As defined in TS 102 925 [24], in order to take into account the difference between the reference test positioning (0,5 meter), defined as HFRP in Recommendation ITU-T P.340 [6], and the actual microphone-talker operating distance (d_S) for which the terminal is adjusted, the following correction factor F_S is defined:

$$F_S \text{ (dB)} = 20 \text{ Log } (d_S/0,5) \qquad (d_S \text{ in meters})$$

The formula may be used to define the relevant level calibration for telepresence systems when using the reference signal level defined for desktop terminal.

5.1.4.2 Receive

Unless specified otherwise, the applied test signal level at the digital input shall be -16 dBm0.

5.1.5 Setup of background noise simulation

A setup for simulating realistic background noises in a lab-type environment is described in ES 202 396-1 [10].

The signals attached to ES 202 396-1 [10] are fullband signals and should be used for background noise simulation.

5.1.6 Acoustic environment

5.1.6.1 Measurement environment

NOTE: The acoustic environment may influence more significantly the results in low and high frequencies. It should be adapted to the terminal bandwidth.

In general two possible approaches need to be taken into account: either room noise and background noise are an inherent part of the test environment or room noise and background noise shall be eliminated to such an extent that their influence on the test results can be neglected.

Unless stated otherwise measurements shall be conducted under quiet and "anechoic" conditions.

In cases where real or simulated background noise is used as part of the testing environment, the original background noise shall not be noticeably influenced by the acoustical properties of the room.

In all cases where the performance of acoustic echo cancellers shall be tested, a realistic room, which represents the typical user environment for the terminal shall be used.

5.1.7 Influence of terminal delay issue for measurements

As delay is introduced by the terminal, care shall be taken for all measurements using an activation signal. Appropriate delay compensation shall be performed in advance to the analysis in order to ensure correct positioning of the analysis window. It shall be checked that the test is performed on the test signal and not on the activation signal.

5.2 Environmental conditions for tests

The following conditions shall apply for the testing environment:

a) Ambient temperature: 15 °C to 35 °C (inclusive).

b) Relative humidity: 5 % to 85 %.

c) Air pressure: 86 kPa to 106 kPa (860 mbar to 1 060 mbar).

d) Unless specified otherwise, the background noise level shall be less than -64 dBPa(A) in conjunction with NC30 (ISO 3745 [23]).

For specified tests, it is desirable to have a background noise level of less than -74 dBPa(A) in conjunction with NC20, but the background noise level of -64 dBPa(A) in conjunction with NC30 shall never be exceeded.

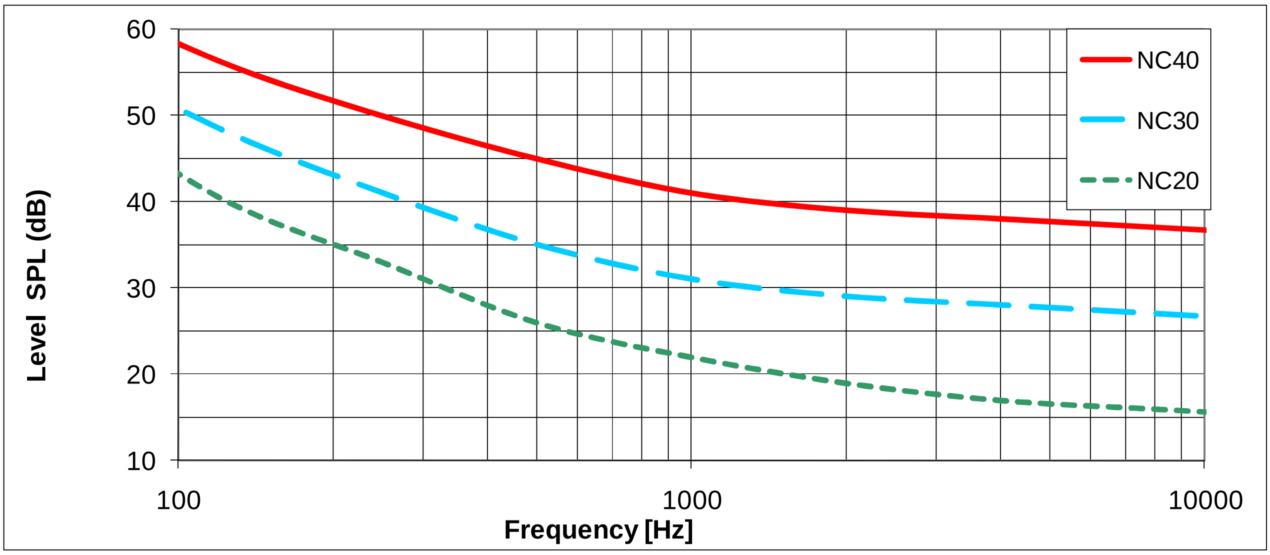


Figure 5.2: NC-criteria for test environment

5.3 Accuracy of measurements and test signal generation

Unless specified otherwise, the accuracy of measurements made by test equipment shall be equal to or better than:

Table 5.3.A: Measurement Accuracy

Item	Accuracy
Electrical signal level	±0,2 dB for levels ≥ -50 dBV ±0,4 dB for levels < -50 dBV
Sound pressure	±0,7 dB
Frequency	±0,2 %
Time	±0,2 %

Unless specified otherwise, the accuracy of the signals generated by the test equipment shall be better than:

Table 5.3.B: Accuracy of test signal generation

Quantity	Accuracy
Sound pressure level at HandheldHandsFree Reference Point (HhHFRP)	0 dB to -6 dB for frequencies from 50 Hz to 100 Hz ±1 dB for frequencies from 100 Hz to 8 000 Hz ±3 dB for frequencies from 8 000 Hz to 16 000 Hz
Electrical excitation levels	±0,4 dB across the whole frequency range
Frequency generation	±2 %
Time	±0,2 %
Specified component values	±1 %
NOTE: This tolerance may be used to avoid measurements at critical frequencies, e.g. those due to sampling operations within the terminal under test.	

NOTE: With some measurement equipment the use of such a bandwidth is not possible and should be limited to 100 Hz to 14 kHz.

For terminal equipment which is directly powered from the mains supply, all tests shall be carried out within ±5 % of the rated voltage of that supply. If the equipment is powered by other means and those means are not supplied as part of the apparatus, all tests shall be carried out within the power supply limit declared by the supplier. If the power supply is a.c., the test shall be conducted within ±4 % of the rated frequency.

5.4 Specific test considerations

Even if the present document is dedicated to conversational services, the signals that are transmitted may combine speech and audio. However the present document is mainly addressing speech communications.

6 Requirement considerations and test methods

When possible, parameter requirements will be derived from requirements defined for the wideband terminals. The recommended test method is also provided in the same clause as requirements.

NOTE: The present document does not provide loudness to define the level perceived by the user. It is expected to add this parameter in a further version.

6.1 Handsfree - Send

All the types of terminals within the scope of the present document shall fulfil the requirements of this clause. Even if these terminals are rather different, the intention of the present document is to guarantee that all the terminals effectively transmit superwideband bandwidths.

6.1.1 Frequency response

Requirements

The objective is to define a flat frequency curve over the whole bandwidth.

The frequency response for handsfree function shall fulfil the mask as defined in table 6.1.1 and figure 6.1.1.

Table 6.1.1: Frequency mask for handsfree terminals - Send

Frequency	Upper Limit	Lower Limit
50 Hz	0 dB	
100 Hz	5 dB	-5 dB
12 500 Hz	5 dB	-5 dB
14 000 Hz	5 dB	-10 dB
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (Hz) scale.		

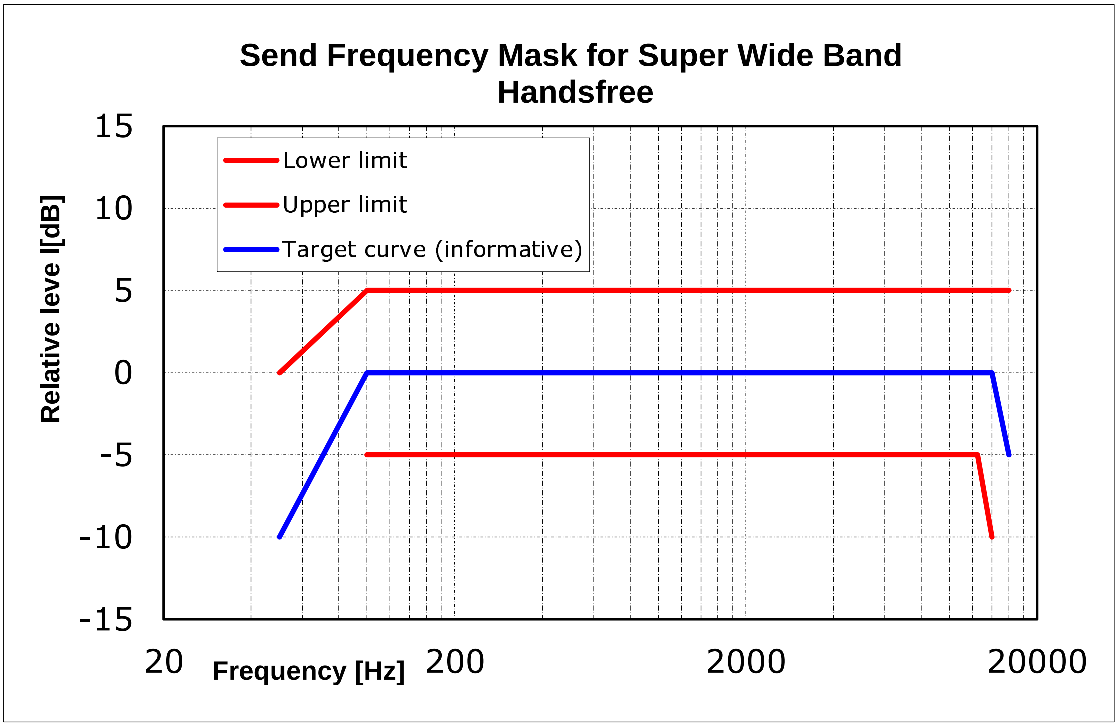


Figure 6.1.1: Frequency mask for handsfree terminals - Send

Measurement method

The test set-ups are described in clause 5.1.

Measurements shall be made at one twelfth-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [22] for frequencies from 100 Hz to 14 kHz inclusive.

For the calculation the averaged measured level at the electrical reference point for each frequency band is referred to the averaged test signal level measured in each frequency band at the HFRP.

The sensitivity is expressed in terms of dBV/Pa.

6.1.2 Loudness rating (SLR)

Requirement

To ensure the compatibility with other terminals or systems a reference SLR needs to be defined.

The requirements refer to wideband handsfree terminals, ES 202 740 [11], or TS 103 740 [12].

Nominal value: +13dB ± 3 dB.

There is no specific requirement for SWB bandwidth.

Measurement method of Wideband Loudness rating.

The test set-ups are described in clause 5.1.

For a correct activation of the system, the test signal to be used for the measurements shall be the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [1], Amendment 1.The spectrum of acoustic signal produced by the artificial mouth is calibrated under free field conditions at the MRP. The test signal level shall be -4,7 dBPa, measured at the MRP. The test signal level is averaged over the complete test signal sequence.

The send sensitivity shall be calculated from each band of the 20 frequencies given in table 1 of Recommendation ITU-T P.79 [5], bands 1 to 20. For the calculation the averaged measured level at the electrical reference point for each frequency band is referred to the averaged test signal level measured in each frequency band at the MRP.

The sensitivity is expressed in terms of dBV/Pa and the SLR shall be calculated according to Recommendation ITU-T P.79 [5], annex A.

6.1.3 Level dependency

The loudness ratings are tested for different input levels (at least the nominal signal level, a 10 dB lower and a 5 dB higher).

Requirements are for further study.

6.1.4 Send noise

Requirements

The limit for the send noise is the following:

- send noise level maximum -64 dBm(A).

No peaks in the frequency domain higher than 10 dB above the average noise spectrum shall occur.

Measurement method

The test set-ups are described in clause 5.1.

The female speaker of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [1], Amendment 1 shall be used for activation. The level of this activation signal will be -4,7 dBPa at the MRP.

The level at the output of the test setup is measured with a A weighting, in the bandwidth from 50 Hz and 14 kHz.

6.1.5 Send distortion

6.1.5.1 Signal to harmonic distortion versus frequency

Requirements

The ratio of signal to harmonic distortion shall be above the following masks.

The following draft requirements are defined for all the terminals within the scope of the present document, as it is needed to ensure that any terminal intended to be used in superwideband sends good quality signals. Care should be taken on the distortion of the HATS or of the loudspeaker used to test the send distortion of the terminal.

Table 6.1.5.1

Frequency	Ratio
100 Hz	25 dB
200 Hz	30 dB
400 Hz	30 dB
1 kHz	30 dB
2 kHz	30 dB
5 kHz	30 dB
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (Hz) scale.	

The signal to harmonic distortion ratio is measured selectively up to 16 kHz.

Measurement method

The test set-ups are described in clause 5.1.

The signal used is an activation signal followed by a series sine wave signal with a frequency at 100 Hz, 200 Hz, 400 Hz, 1 kHz, 2 kHz and 5 kHz. The signal to harmonic distortion ratio is measured selectively up to 14 kHz.

The duration of the sine wave shall be of less than 1 s. The sinusoidal signal level shall be calibrated to -4,7 dBPa at the MRP.

For a correct activation of the system, the female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [1], Amendment 1 shall be used for activation. The level of this activation signal is -4,7 dBPa at the MRP.

NOTE 1: Depending on the type of codec or signal processing the test signal used may need to be adapted.

NOTE 2: When using HATS, due to the distortion limit of the artificial mouth at 100 Hz, as defined in table 10 of Recommendation ITU-T P.58 [3], the measurements with a frequency of 100 Hz and possibly 200 Hz have to take into account the actual distortion of the artificial mouth.

6.1.5.2 Signal to harmonic distortion for higher input level

Requirement

For the signal defined in the measurement method, the signal to harmonic distortion ratio shall be ≥ 30 dB.

Measurement method

The test set-ups are described in clause 5.1.

The signal used is an activation signal followed by a series sine wave signal with a frequency at 1 kHz. The signal to harmonic distortion ratio is measured selectively up to 14 kHz.

The duration of the sine wave shall be of less than 1 s. The sinusoidal signal level shall be calibrated to +10 dBPa at the MRP.

For a correct activation of the system, the female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [1], Amendment 1 shall be used for activation. The level of this activation signal is -4,7 dBPa at the MRP.

NOTE: Depending on the type of codec or signal processing the test signal used may need to be adapted.

6.2 Handsfree - Receive

6.2.1 Equalization

This type of terminal may be used for reproduction of signals other than pure speech (e.g. music) for which user's preference may be different in term of sound signature. So, the terminals may implement an equalization function adjusting frequency response according to user's preference.

When such a function is available it is necessary that the receive frequency response conforms to requirements defined in clause 6.2.2 for at least one setting.

For all settings the conformance with other parameter of the present shall be ensured.

6.2.2 Frequency response

When using HATS (with the restrictions defined in clause 5) HATS shall be equalized according to Recommendation ITU-T P.581 [4].

The size and the applications of the handheld terminal may impact the performance. Class B requirements shall be fulfilled by all the terminals. Terminals ensuring better quality should fulfil Class A requirements.

Requirements

Table 6.2.2.A: Frequency mask for Class A handsfree terminals - Receive

Frequency	Upper Limit	Lower Limit
50 Hz	5 dB	-15 dB
150 Hz	5 dB	-5 dB
400 Hz	5 dB	-5 dB
12 500 Hz	5 dB	-5 dB
14 000 Hz	5 dB	-10 dB
16 000 Hz	5 dB	
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (Hz) scale.		

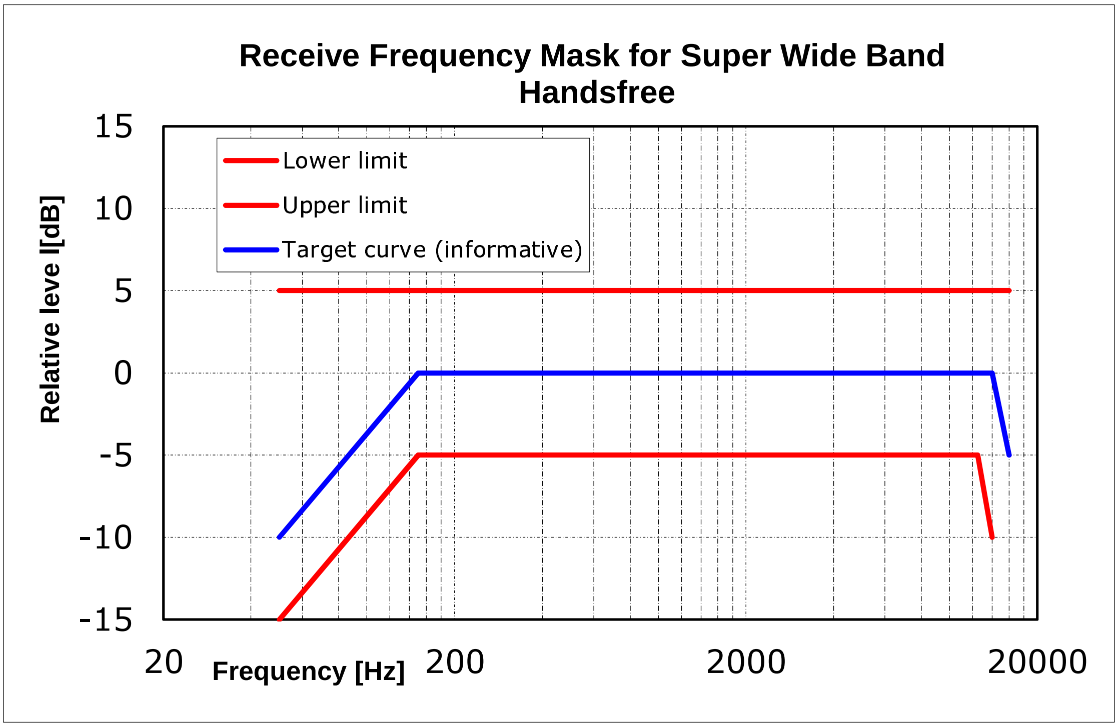


Figure 6.2.2.A: Frequency mask for Class A handsfree terminals - Receive

Table 6.2.2.B: Frequency mask for Class B handsfree terminals - Receive

Frequency	Upper Limit	Lower Limit
50 Hz	5 dB	
400 Hz	5 dB	-5 dB
12 500 Hz	5 dB	-5 dB
14 000 Hz	5 dB	-10 dB
16 000 Hz	5 dB	
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (Hz) scale.		

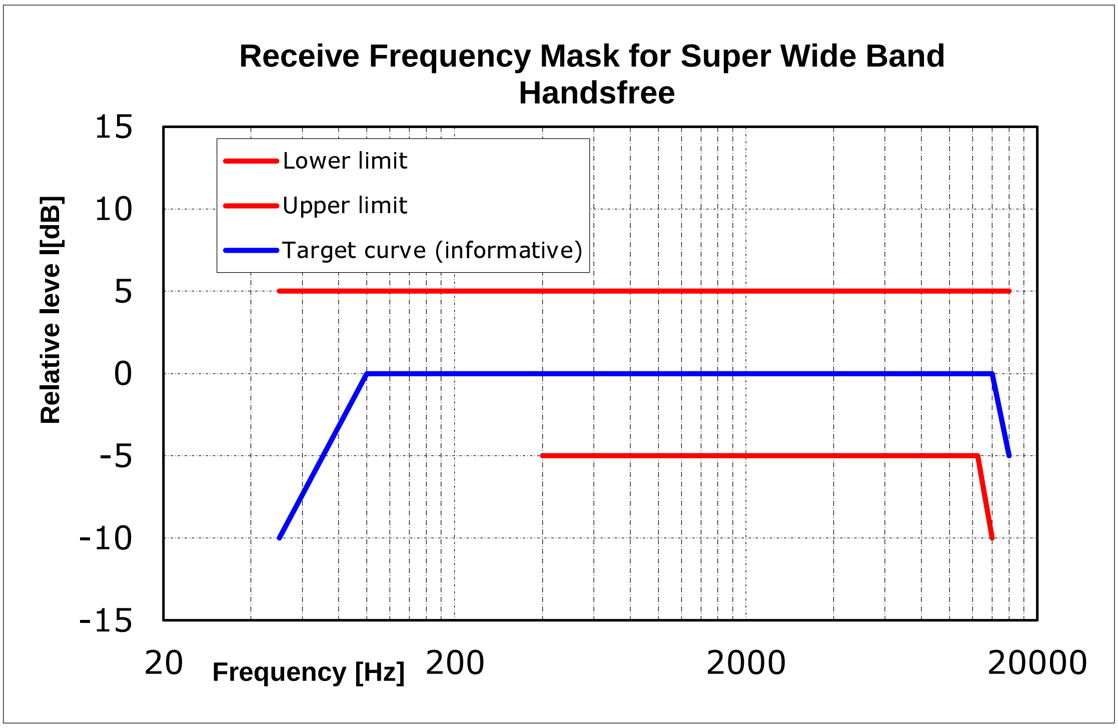


Figure 6.2.2.B: Frequency mask for Class B handsfree terminals - Receive

Measurement methods

The test set-ups are described in clause 5.1.

The test signal to be used for the measurements shall be British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [1], Amendment 1. The test signal level shall be -16 dBm0, measured according to Recommendation ITU-T P.56 [21] at the digital reference point or the equivalent analogue point.

The equalized output signal is power-averaged on the total time of analysis. The 1/3 octave band data are considered as the input signal to be used for calculations or measurements.

Measurements shall be made at one third-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [22] for frequencies from 400 Hz to 14 kHz inclusive. For the calculation the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dBPa/V.

6.2.3 Loudness Rating (RLR)

6.2.3.1 Loudness Rating

Requirements

When terminal implements wideband speech functions or when the superwideband functions may interact with wideband terminals, the handsfree terminal shall fulfil the requirements on RLR as defined in ES 202 740 [11], clause 7.1.7, or TS 103 740 [12]:

- Handheld terminal
  - Nominal value of RLR will be 9 dB ± 3 dB. This value has to be fulfilled for one position of volume range.
  - Value of RLR at upper part of volume range shall be less than (louder) or equal to 5 dB: RLR ≤ 5 dB.
  - Range of volume control shall be ≥ 15 dB.

NOTE 1: Due to the lack of experience in the application of wide band loudness rating calculation as defined in annex G of Recommendation ITU-T P.79 [5] the loudness rating calculation as described in annex A is used.

NOTE 2: Loudness Rating measurement corresponding to level with speech signal, it can be considered that a measurement in wideband may be sufficient. Indeed, energy of speech beyond bandwidth of wideband is rather small.

Measurement method

The test set-ups are described in clause 5.1.

The measurement is conducted at nominal volume control setting.

Receive frequency response is the ratio of the measured sound pressure and the input level.
(dB relative Pa/V)

$$S_{Jeff} = 20 \log (p_{eff} / v_{RCV}) \text{ dB rel 1 Pa / V} \quad (1)$$

S_{Jeff} Receive Sensitivity; Junction to HATS Ear with free field correction.

p_{eff} DRP Sound pressure measured by ear simulator Measurement data are converted from the Drum Reference Point to free field.

v_{RCV} Equivalent RMS input voltage.

The test signal to be used for the measurements shall be British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [1], Amendment 1. The test signal level shall be -20 dBm0, measured according to Recommendation ITU-T P.56 [21] at the digital reference point or the equivalent analogue point.

The HATS is free field equalized as described in Recommendation ITU-T P.581 [4]. The equalized output signal is power-averaged on the total time of analysis. The 1/3 octave band data are considered as the input signal to be used for calculations or measurements.

Measurements shall be made at one third-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [22] for frequencies from 100 Hz to 8 kHz inclusive. For the calculation the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dBPa/V.

6.2.4 Receive noise

Requirements

- A-weighted.

The noise level measured until 10 kHz shall not exceed -54 dBPa(A) at nominal setting of the volume control.

- Third-octave band spectrum.

The level in any 1/3-octave band, between 50 Hz and 12,5 kHz shall not exceed a value of -64 dBPa.

NOTE: No peaks in the frequency domain higher than 10 dB above the average noise spectrum should occur.

Measurement method

The test set-ups are described in clause 5.1.

The female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [1], Amendment 1 shall be used for activation. Level of this activation signal will be -16 dBm0.

For the A-weighted noise level measurement the noise level is measured until 14 kHz.

For the 1/3 octave band spectrum the level is measured in all the 1/3-octave bands, between 50 Hz and 12,5 kHz.

The noise shall be measured just after interrupting the activation signal.

6.2.5 Receive distortion

Requirements

The ratio of signal to harmonic distortion shall be above the following mask.

Table 6.2.5.A: Requirements for Class A handsfree terminals

Frequency	Signal to distortion ratio limit, receive for handsfree terminal
100 Hz	15 dB
200 Hz	17 dB
315 Hz	22dB
400 Hz	30 dB
500 Hz	30 dB
800 Hz	30 dB
1 kHz	30 dB
2 kHz	30 dB
5 kHz	30 dB
8 kHz	30 dB
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (Hz) scale.	

Table 6.2.5.B: Requirements for Class B handsfree terminals

Frequency	Signal to distortion ratio limit, receive for handsfree terminal
100 Hz	ffs (for further study)
200 Hz	ffs
315 Hz	ffs
400 Hz	ffs
500 Hz	15 dB
800 Hz	20 dB
1 kHz	25 dB
2 kHz	25 dB
5 kHz	30 dB
8 kHz	30 dB
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (Hz) scale.	

Measurement method

The test set-ups are described in clause 5.1.

The signal used is an activation signal followed by a sine wave signal with a frequency at 100 Hz, 200 Hz, 315 Hz, 400 Hz, 500 Hz, 800 Hz, 1 kHz, 2 kHz, 5 kHz et 8 kHz. The duration of the sine wave shall be of less than 1 s. Appropriate signals for activation and signal combinations can be found in Recommendation ITU-T P.501 [1], Amendment 1. The sinusoidal signal level shall be calibrated to -16 dBm0.

The female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [1], Amendment 1 shall be used for activation. Level of this activation signal will be -16 dBm0.

The signal to harmonic distortion ratio is measured selectively up to 14 kHz.

NOTE: Depending on the type of codec the test signal used may need to be adapted.

6.3 Handsfree - Other parameters

The parameters to be defined in this clause are intended to guarantee that this expected quality is effectively offered to the user(s).

6.3.1 Round-trip Delay

Requirement

Send and receive delays are tested separately, however the requirement is defined for the combination of send and receive delays (round-trip delay).

It is recognized that the end to end delay should be as small as possible in order to ensure high quality of the communication.

The sum of delay in send direction $\mathbf{T_s}$ and in receive direction $\mathbf{T_r}$ shall be less than $\mathbf{T_{rtd\text{-}req}}$ where $\mathbf{T_{rtd\text{-}req}}$ is the Round Trip Delay Requirement.

Send direction

The delay in send direction is measured from the MRP to POI. The delay measured in send direction is:

$$\mathbf{T_s + t_{System}}$$

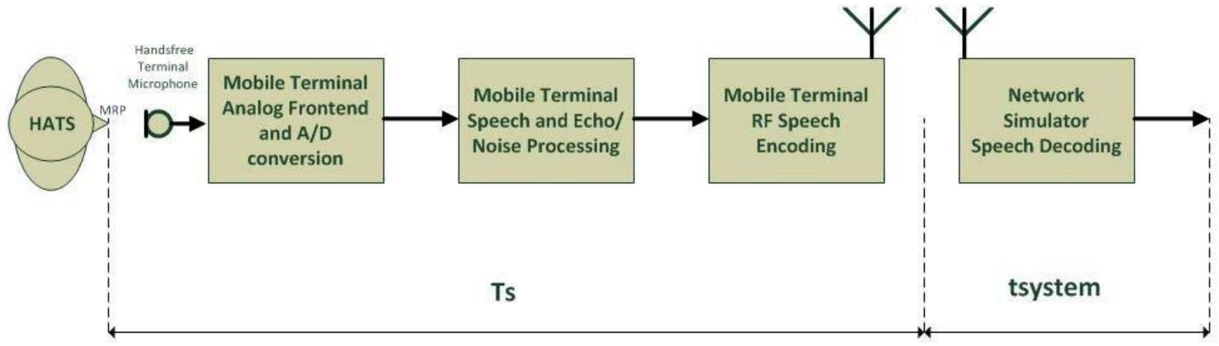


Figure 6.3.1.A: Different blocks contributing to the delay in send direction

The system delay $\mathbf{t_{System}}$ depends on the transmission method used and the network simulator. The delay $\mathbf{t_{System}}$ shall be known.

1) For the measurements a Composite Source Signal (CSS) according to Recommendation ITU-T P.501 [1] is used. The pseudo random noise (pn)-part of the CSS has to be longer than the maximum expected delay. It is recommended to use a pn sequence of 16 k samples (with 48 kHz sampling rate). The test signal level is -24,3 dBPa at the HhHFRP.

   The reference signal is the original signal (test signal).

   The setup of the hands-free terminal is in line with clause 5.2.

2) The delay is determined by cross-correlation analysis between the measured signal at the electrical access point and the original signal. The measurement is corrected by substracting the delays which are caused by the test equipment.

3) The maximum value of the cross-correlation function is used for the determination of the delay.

Receive direction

The delay in receive direction is measured from POI to the Drum Reference Point (DRP). The delay measured in receive direction is:

$$\mathbf{T_r + t_{System}}$$

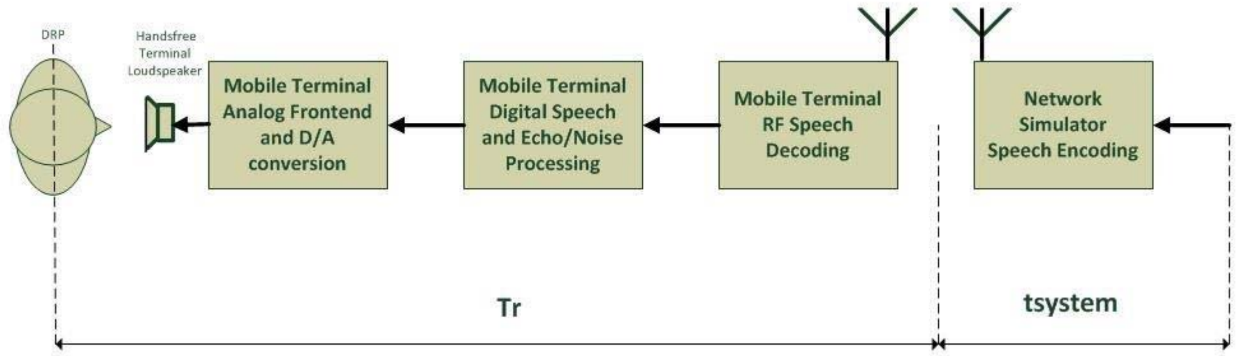


Figure 6.3.1.B: Different blocks contributing to the delay in receive direction

The system delay $\mathbf{t}_{\mathbf{System}}$ depends on the transmission system and on the network simulator used. The delay $\mathbf{t}_{\mathbf{System}}$ shall be known.

1) For the measurements a Composite Source Signal (CSS) according to Recommendation ITU-T P.501 [1] is used. The pseudo random noise (pn)-part of the CSS has to be longer than the maximum expected delay. It is recommended to use a pn sequence of 16 k samples (with 48 kHz sampling rate). The test signal level is -16 dBm0 at the electrical interface (POI).

   The reference signal is the original signal (test signal).

2) The test arrangement is in line with clause 5.2.

3) The delay is determined by cross-correlation analysis between the measured signal at the DRP and the original signal. The measurement is corrected by substracting the delays which are caused by the test equipment.

4) The maximum value of the cross-correlation function is used for the determination of the delay.

6.3.2 Terminal Echo Loss

Requirement

The TCL measured as unweighted Echo Loss shall be ≥ 46 dB for all positions of the volume control (if supplied).

Measurement method

The handsfree terminal is setup as described in clause 5.1. The ambient noise level shall be < -64 dBPa(A). The attenuation from electrical reference point input to electrical reference point output shall be measured using the compressed real speech signal described in clause 7.3.3 of Recommendation ITU-T P.501 Amendment 1 [1].

TCL is calculated as unweighted echo loss from 100 Hz to 8 kHz. For the calculation the averaged measured echo level at each frequency band is referred to the averaged test signal level measured in each frequency band. The first 17,0 s of the test signal (6 sentences) are discarded from the analysis to allow for convergence of the acoustic echo canceller. The analysis is performed over the remaining length of the test sequence (last 6 sentences).

Test methods and requirements for the full bandwidth are for further study.

6.3.3 Objective listening quality

"Recommendation ITU-T P.863 [14] describes an objective method for predicting overall listening speech quality from narrowband (300 to 3 400 Hz) to superwideband (50 to 14'000 Hz) telecommunication scenarios as perceived by the user in an Recommendation ITU-T P.800 [18] or Recommendation ITU-T P.830 [19] ACR listening only test."

The speech sequences shall be carefully selected.

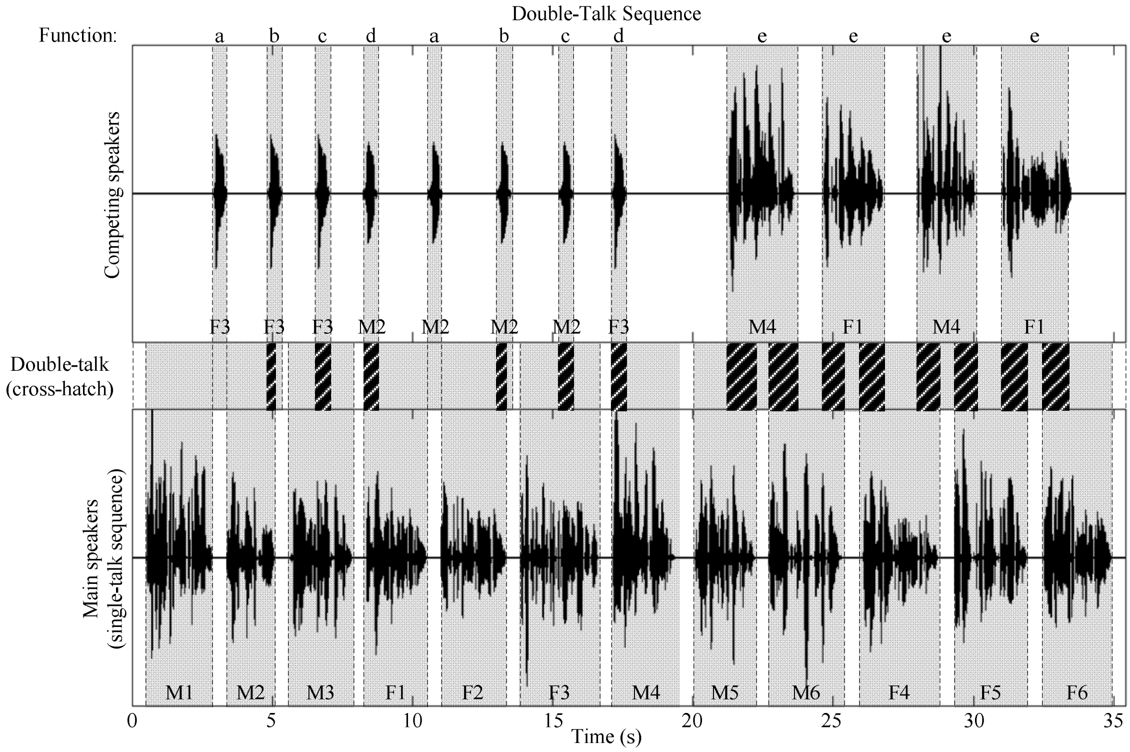
The requirements and the detailed measurement method are for further study.

6.3.4 Double talk performance

Requirements are for further study.

Measurement method

To assess double talk performance, the signals to be used are defined in Recommendation ITU-T P.501 [1], Amendment 1: A "double-talk" sequence representing typical double talk scenarios in real conversations is shown in figure 6.3.4. This uses the single-talk sequence described in section 7.3.1 of Recommendation ITU-T P.501 [1], Amendment 1, shown in the lower pane, as the main speech and an additional competing speaker sequence, shown in the upper pane.



NOTE: Cross-hatched areas between the upper and lower panes show periods of double talk.

Figure 6.3.4: Double-talk test sequence using the single-talk sequence and competing speech serving different functions (a - e)

The competing-speaker sequence includes single words (the word "five") spoken by speakers F3 and M2 during the first half of the sequence followed by full sentences by speakers F1 and M4 during the second half of the sequence. No speaker is competing with themselves during the sequence.

The competing samples serve different double-talk functions, defined as functions "a" to "e" above the upper pane of figure 6.3.4. The functions are:

a) Competing word within a speech pause.

b) Competing word partially masked.

c) Competing word fully masked within a sentence.

d) Competing word fully masked coincident with the start of a sentence.

e) Sentence masking another sentence.

These are meant to represent possible double-talk situations in normal conversation. The area between the upper and lower pane of figure 6.3.4 shows the periods during which double-talk happens as cross-hatched patches. The competing sequence can be used either as a send signal or a receive signal in testing.

6.3.5 Speech quality in presence of noise

For further study.

6.4 Handset- Send

All the types of terminals within the scope of the present document shall fulfil the requirements of this clause. Even if these terminals are rather different, the intention of the present document is to guarantee that all the terminals effectively transmit superwideband bandwidths.

6.4.1 Frequency response

Requirements

The objective is to define a flat frequency curve over the whole bandwidth.

The send frequency response of the handset shall be within a mask as defined in table 6.4.1 and shown in figure 6.4.1. This mask shall be applicable for all types of handsets.

Table 6.4.1: Frequency mask for handset terminals - Send

Frequency	Upper Limit	Lower Limit
50 Hz	0 dB	
100 Hz	5 dB	-5 dB
12 500 Hz	5 dB	-5 dB
14 000 Hz	5 dB	-10 dB
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (Hz) scale.		

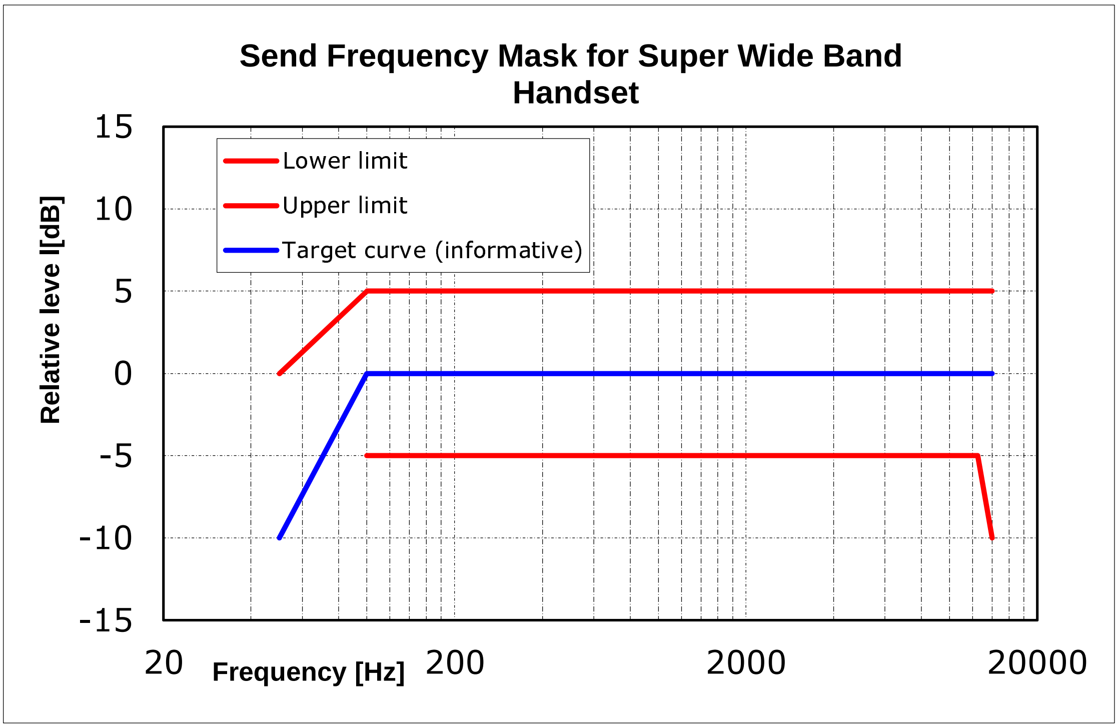


Figure 6.4.1: Send frequency response for handset terminals

Measurement method

The test set-ups are described in clause 5.1.

The handset is mounted in the HATS position (see Recommendation ITU-T P.64 [25]). The application force used to apply the handset against the artificial ear is noted in the test report.

Measurements shall be made at one twelfth-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [22] for frequencies from 100 Hz to 14 kHz inclusive.

For the calculation the averaged measured level at the electrical reference point for each frequency band is referred to the averaged test signal level measured in each frequency band at the MRP.

The sensitivity is expressed in terms of dBV/Pa.

6.4.2 Loudness Rating (SLR)

Requirement

To ensure the compatibility with other terminals or systems a reference SLR needs to be defined.

The requirements refer to wideband handset terminals, ES 202 739 [13] or TS 103 739 [16].

Nominal value : +8 dB ± 3 dB.

There is no specific requirement for SWB bandwidth.

Measurement method of Wideband Loudness rating.

The test set-ups are described in clause 5.1.

For a correct activation of the system, the test signal to be used for the measurements shall be the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [1], Amendment 1.The spectrum of acoustic signal produced by the artificial mouth is calibrated under free field conditions at the MRP. The test signal level shall be -4,7 dBPa, measured at the MRP. The test signal level is averaged over the complete test signal sequence.

The handset is mounted in the HATS position (see Recommendation ITU-T P.64 [25]). The application force used to apply the handset against the artificial ear is noted in the test report.

The send sensitivity shall be calculated from each band of the 20 frequencies given in table 1 of Recommendation ITU-T P.79 [5], bands 1 to 20. For the calculation the averaged measured level at the electrical reference point for each frequency band is referred to the averaged test signal level measured in each frequency band at the MRP.

The sensitivity is expressed in terms of dBV/Pa and the SLR shall be calculated according to Recommendation ITU-T P.79 [5], Annex A.

6.4.1 Level dependency

Requirements

Requirements are for further study.

Measurement method

The loudness ratings are tested for different input levels (at least the nominal signal level, a 10 dB lower and a 5 dB higher).

Same method as SLR with adapted level.

6.4.2 Send noise

Requirements

The maximum noise level produced by the VoIP terminal at the POI under silent conditions in the send direction shall not exceed -68 dBm0 (A).

No peaks in the frequency domain higher than 10 dB above the average noise spectrum shall occur.

Measurement method

The test set-ups are described in clause 5.1.

The handset is mounted at the HATS position (see Recommendation ITU-T P.64 [25]).

For the actual measurement no test signal is used. In order to reliably activate the terminal an activation signal is introduced before the actual measurement. The female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [1] shall be used for activation. The level of this activation signal is -4,7 dBPa at the MRP. The level at the output of the test setup is measured with a A weighting, in the bandwidth from 50 Hz and 14 kHz.

The send noise is measured at the POI in the frequency range from 50 Hz to 14 kHz for Superwideband. The analysis window is applied directly after stopping the activation signal but taking into account the influence of all acoustical components (reverberations). The averaging time is 1 s. The test house has to ensure (e.g. by monitoring the time signal) that during the test the terminal remains in activated condition. If the terminal is deactivated during the measurement, the measurement time has to be reduced to the period where the terminal remains in activated condition.

The noise level is measured in dBm0(A).

6.4.3 Send distortion

The following draft requirements are defined for all the terminals within the scope of the present document, as it is needed to ensure that any terminal intended to be used in superwideband sends good quality signals. Care should be taken on the distortion of the HATS or of the loudspeaker used to test the send distortion of the terminal.

6.4.3.1 Signal to harmonic distortion versus frequency

Requirements

The ratio of signal to harmonic distortion shall be above the following masks.

Table 6.4.3.1

Frequency	Ratio
100 Hz	25 dB
200 Hz	30 dB
400 Hz	30 dB
1 kHz	30 dB
2 kHz	30 dB
5 kHz	30 dB
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (Hz) scale.	

Measurement method

The test set-ups are described in clause 5.1.

The signal used is an activation signal followed by a series sine wave signal with a frequency at 100 Hz, 200 Hz, 400 Hz, 1 kHz, 2 kHz and 5 kHz. The signal to harmonic distortion ratio is measured selectively up to 14 kHz.

The duration of the sine wave shall be of less than 1 s. The sinusoidal signal level shall be calibrated to -4,7 dBPa at the MRP.

For a correct activation of the system, the female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [1], Amendment 1 shall be used for activation. The level of this activation signal is -4,7 dBPa at the MRP.

NOTE: Depending on the type of codec or signal processing the test signal used may need to be adapted.

6.4.3.2 Signal to harmonic distortion for higher input level

Requirement

For the signal defined in the measurement method, the signal to harmonic distortion ratio shall be ≥ 30 dB.

Measurement method

The test set-ups are described in clause 5.1.

The signal used is an activation signal followed by a series sine wave signal with a frequency at 1 kHz. The signal to harmonic distortion ratio is measured selectively up to 14 kHz.

The duration of the sine wave shall be of less than 1 s. The sinusoidal signal level shall be calibrated to +10 dBPa at the MRP.

For a correct activation of the system, the female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [1], Amendment 1 shall be used for activation. The level of this activation signal is -4,7 dBPa at the MRP.

NOTE: Depending on the type of codec or signal processing the test signal used may need to be adapted.

6.5 Handset - Receive

6.5.1 Equalization

This type of terminal may be used for reproduction of signals other than pure speech (e.g. music) for which user's preference may be different in term of sound signature. So, the terminals may implement an equalization function adjusting frequency response according to user's preference.

When such a function is available it is necessary that the receive frequency response conforms to requirements defined in clause 6.5.2 for at least one setting.

For all settings the conformance with other parameter of the present shall be ensured.

6.5.2 Frequency response

When using HATS (with the restrictions defined in clause 5) HATS shall be equalized according to Recommendation ITU-T P.581 [4].

Requirements

The receive frequency response of the handset shall be within a mask as defined in table 6.5.2 and shown in figure 6.5.2. The application force for handsets is 2 N, 8 N and 13 N.

Table 6.5.2: Frequency mask for handset terminals - Receive

Frequency	Upper limit 8N	Lower limit 8N	Upper limit 2N	Lower limit 2N	Upper limit 13N	Lower limit 13N
50 Hz	3 dB	-5 dB	3 dB	-5 dB	3 dB	-5 dB
400 Hz	3 dB	-5 dB	3 dB	-5 dB	3 dB	-5 dB
1 010 Hz	See note	-5 dB	See note	-5 dB	See note	-5 dB
1 200 Hz	See note	-8 dB	See note	-8 dB	See note	-8 dB
1 500 Hz	See note	-8 dB	See note	-8 dB	See note	-8 dB
2 000 Hz	9 dB	-3 dB	9 dB	-3 dB	9 dB	-3 dB
3 200 Hz	9 dB	-3 dB	9 dB	-3 dB	9 dB	-3 dB
14 000 Hz	9 dB	-13 dB	9 dB	-13 dB	9 dB	-13 dB
NOTE: The limit curves shall be determined by straight lines joining successive co-ordinates given in the table, where frequency response is plotted on a linear dB scale against frequency on a logarithmic scale. The mask is a floating or "best fit" mask.						

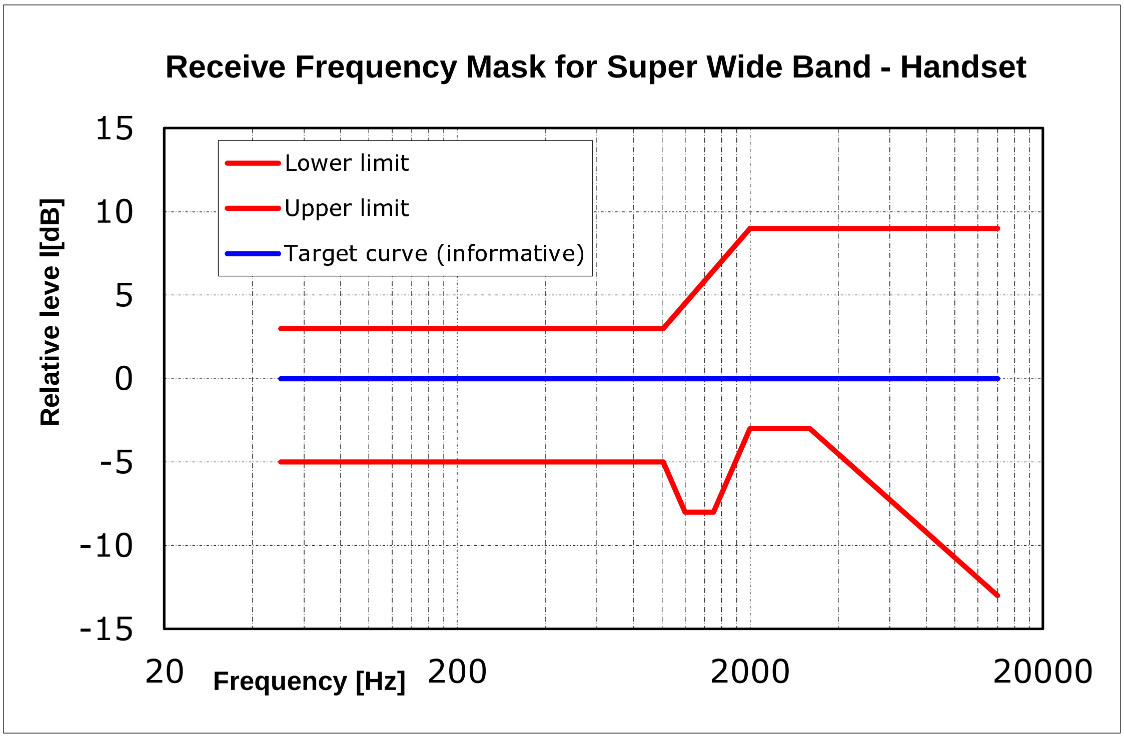


Figure 6.5.2: Frequency mask for handset terminals - Receive valid for handset application forces 2 N, 8 N and 13 N

Measurement methods

Receive frequency response is the ratio of the measured sound pressure and the input level (dB relative Pa/V).

$$S_{Jeff} = 20 \log (p_{eff} / v_{RCV}) \text{ dB rel 1 Pa / V}$$

SJeff Receive Sensitivity; Junction to HATS Ear with free field correction.

peff DRP Sound pressure measured by ear simulator Measurement data are converted from the Drum Reference Point to free field.

vRCV Equivalent RMS input voltage.

The test set-ups are described in clause 5.1.

The handset is mounted in the HATS position (see Recommendation ITU-T P.64 [25]). The application forces used to apply the handset against the artificial ear is 2 N, 8 N and 13 N.

The test signal to be used for the measurements shall be British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [1], Amendment 1. The test signal level shall be -16 dBm0, measured according to Recommendation ITU-T P.56 [21] at the digital reference point or the equivalent analogue point.

The HATS is diffuse field equalized as described in Recommendation ITU-T P.581 [4]. The diffuse field correction as defined in Recommendation ITU-T P.58 [3] is applied. The equalized output signal is power-averaged on the total time of analysis. Until 10 kHz inclusive, the 1/12 octave band data are considered as the input signal to be used for calculations measurements. Above 10 kHz, the third octave band data are considered as the input signal to be used for calculations measurements. The sensitivity is expressed in terms of dBPa/V.

6.5.3 Loudness Rating (RLR)

6.5.3.1 Loudness Rating

Requirements

When terminal implements wideband speech functions or when the superwideband functions may interact with wideband terminals, the handsfree terminal shall fulfil the requirements on RLR as defined in ES 202 739 [13] or TS 103 739 [16], clause 7.2.12.

Nominal value of RLR shall be -2 dB ± 3 dB. This value has to be fulfilled for one position of volume range.

NOTE 1: Due to the lack of experience in the application of wide band loudness rating calculation as defined in annex G of Recommendation ITU-T P.79 [5] the loudness rating calculation as described in annex A is used.

NOTE 2: Loudness Rating measurement corresponding to level with speech signal, it can be considered that a measurement in wideband may be sufficient. Indeed, energy of speech beyond bandwidth of wideband is rather small.

Measurement method

The test set-ups are described in clause 5.1.

The measurement is conducted at nominal volume control setting.

The test signal to be used for the measurements shall be British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [1], Amendment 1. The test signal level shall be -20 dBm0, measured according to Recommendation ITU-T P.56 [21] at the digital reference point or the equivalent analogue point.

The handset terminal or the headset terminal is setup as described in clause 7.1. The handset is mounted in the HATS position (see Recommendation ITU-T P.64 [25]). The application force used to apply the handset against the artificial ear is noted in the test report. The HATS is NOT diffuse field equalized as described in Recommendation ITU-T P.581 [4]. The DRP-ERP correction as defined in Recommendation ITU-T P.57 [26] is applied. The application force used to apply the handset against the artificial ear is noted in the test report. By default, 8 N will be used.

The receive sensitivity shall be calculated from each band of the 20 frequencies given in table 1 of Recommendation ITU-T P.79 [5], bands 1 to 20. For the calculation the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dBPa/V and the RLR shall be calculated according to Recommendation ITU-T P.79 [5], annex A. No leakage correction shall be applied for the measurement.

6.5.4 Receive noise

Requirements

Telephone sets with adjustable receive levels shall be adjusted so that the RLR is as close as possible to the nominal RLR.

The receive noise shall be less than -57 dBPa(A).

Where a volume control is provided, the measured noise shall not be greater than -54 dBPa(A) at the maximum setting of the volume control.

Measurement method

The test set-ups are described in clause 5.1.

The female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [1], Amendment 1 shall be used for activation. Level of this activation signal will be -16 dBm0.

For the A-weighted noise level measurement the noise level is measured until 14 kHz.

The A-weighted noise level shall be measured at DRP of the artificial ear with the diffuse field equalization active.

6.5.5 Receive distortion

Requirements

The ratio of signal to harmonic distortion shall be above the following mask.

Table 6.5.5

Frequency	Signal to distortion ratio limit, receive for handset terminal
100 Hz	ffs (for further study)
200 Hz	ffs
315 Hz	26 dB
400 Hz	30 dB
500 Hz	30 dB
800 Hz	30 dB
1 kHz	30 dB
2 kHz	30 dB
5 kHz	30 dB
8 kHz	30 dB
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (Hz) scale.	

Measurement method

The test set-ups are described in clause 5.1.

The signal used is an activation signal followed by a sine wave signal with a frequency at 100 Hz, 200 Hz, 315 Hz, 400 Hz, 500 Hz, 800 Hz, 1 kHz, 2 kHz, 5 kHz et 8 kHz. The duration of the sine wave shall be of less than 1 s. Appropriate signals for activation and signal combinations can be found in Recommendation ITU-T P.501 [1], Amendment 1. The sinusoidal signal level shall be calibrated to -16 dBm0.

The female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [1], Amendment 1 shall be used for activation. Level of this activation signal will be -16 dBm0.

The signal to harmonic distortion ratio is measured selectively up to 14 kHz.

The ratio of signal to harmonic distortion shall be measured at the DRP of the artificial ear with the diffuse field equalization active.

NOTE: Depending on the type of codec the test signal used may need to be adapted.

6.6 Handset - Other parameters

The parameters to be defined in this clause are intended to guarantee that this expected quality is effectively offered to the user(s).

6.6.1 Round-trip Delay

Requirement

Send and receive delays are tested separately but the requirement is defined from the combination of send and receive delays (round-trip delay).

It is recognized that the end to end delay should be as small as possible in order to ensure high quality of the communication.

The sum of delay in send direction $\mathbf{T_s}$ and receive direction $\mathbf{T_r}$ shall be less than $\mathbf{T_{rtd\text{-}req}}$ where $\mathbf{T_{rtd\text{-}req}}$ is the Round Trip Delay Requirement.

Measurement method

- **Send direction**

The delay in send direction is measured from the MRP to POI. The delay measured in send direction is:

$$\mathbf{T_s + t_{System}}$$

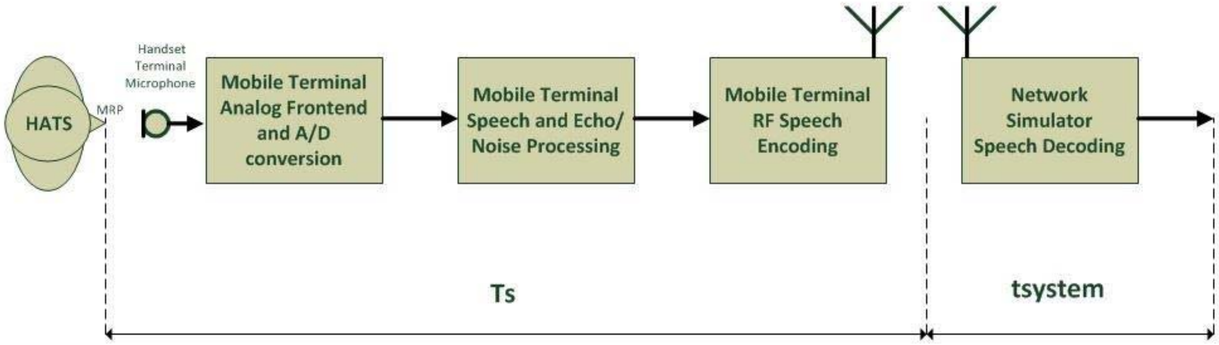


Figure 6.6.1.A: Different blocks contributing to the delay in send direction

The system delay $\mathbf{t_{System}}$ depends on the transmission method used and the network simulator. The delay $\mathbf{t_{System}}$ shall be known.

1) For the measurements a Composite Source Signal (CSS) according to Recommendation ITU-T P.501 [1] is used. The pseudo random noise (pn)-part of the CSS has to be longer than the maximum expected delay. It is recommended to use a pn sequence of 16 k samples (with 48 kHz sampling rate). The test signal level is -4,7 dBPa at the MRP.

   The reference signal is the original signal (test signal).

   The setup of the handset terminal is in line with clause 5.2.

2) The delay is determined by cross-correlation analysis between the measured signal at the electrical access point and the original signal. The measurement is corrected by substracting the delays which are caused by the test equipment.

3) The maximum value of the cross-correlation function is used for the determination of the delay.

- **Receive direction**

The delay in receive direction is measured from POI to the Drum Reference Point (DRP). The delay measured in receive direction is:

$$\mathbf{T_r + t_{System}}$$

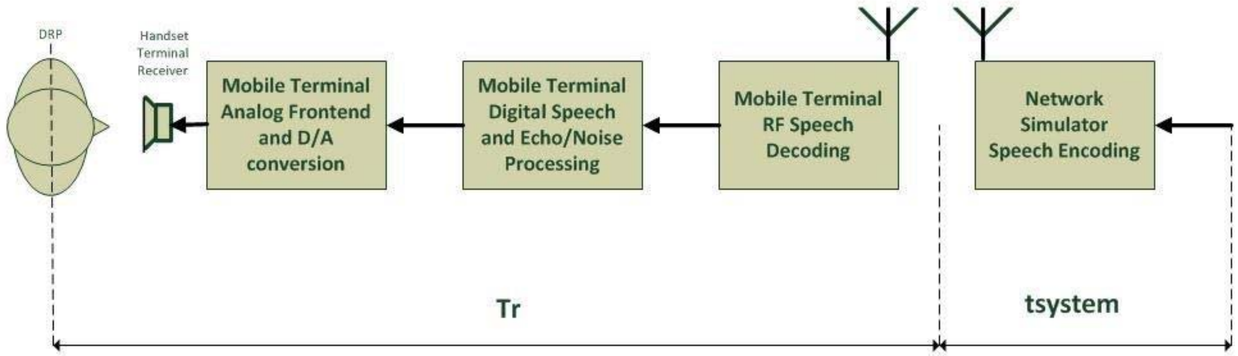


Figure 6.6.1.B: Different blocks contributing to the delay in receive direction

The system delay $\mathbf{t_{System}}$ depends on the transmission system and on the network simulator used. The delay $\mathbf{t_{System}}$ shall be known.

1) For the measurements a Composite Source Signal (CSS) according to Recommendation ITU-T P.501 [1] is used. The pseudo random noise (pn)-part of the CSS has to be longer than the maximum expected delay. It is recommended to use a pn sequence of 16 k samples (with 48 kHz sampling rate). The test signal level is -16 dBm0 at the electrical interface (POI).

   The reference signal is the original signal (test signal).

2) The test arrangement is in line with clause 5.2.

3) The delay is determined by cross-correlation analysis between the measured signal at the DRP and the original signal. The measurement is corrected by substracting the delays which are caused by the test equipment.

4) The maximum value of the cross-correlation function is used for the determination of the delay.

6.6.2 Terminal Echo Loss

Requirement

The TCL measured as unweighted Echo Loss shall be ≥ 46 dB for all positions of the volume control (if supplied).

Measurement method

The handset terminal is setup as described in clause 5.1. The ambient noise level shall be < -64 dBPa(A). The attenuation from electrical reference point input to electrical reference point output shall be measured using the compressed real speech signal described in clause 7.3.3 of Recommendation ITU-T P.501 Amendment 1 [1].

TCL is calculated as unweighted echo loss from 100 Hz to 8 kHz. For the calculation the averaged measured echo level at each frequency band is referred to the averaged test signal level measured in each frequency band. The first 17,0 s of the test signal (6 sentences) are discarded from the analysis to allow for convergence of the acoustic echo canceller. The analysis is performed over the remaining length of the test sequence (last 6 sentences).

6.6.3 Objective listening quality

Recommendation ITU-T P.863 [14] describes an objective method for predicting overall listening speech quality from narrowband (300 to 3 400 Hz) to superwideband (50 to 14 000 Hz) telecommunication scenarios as perceived by the user in an Recommendation ITU-T P.800 [18] or Recommendation ITU-T P.830 [19] ACR listening only test.

The speech sequences shall be carefully selected.

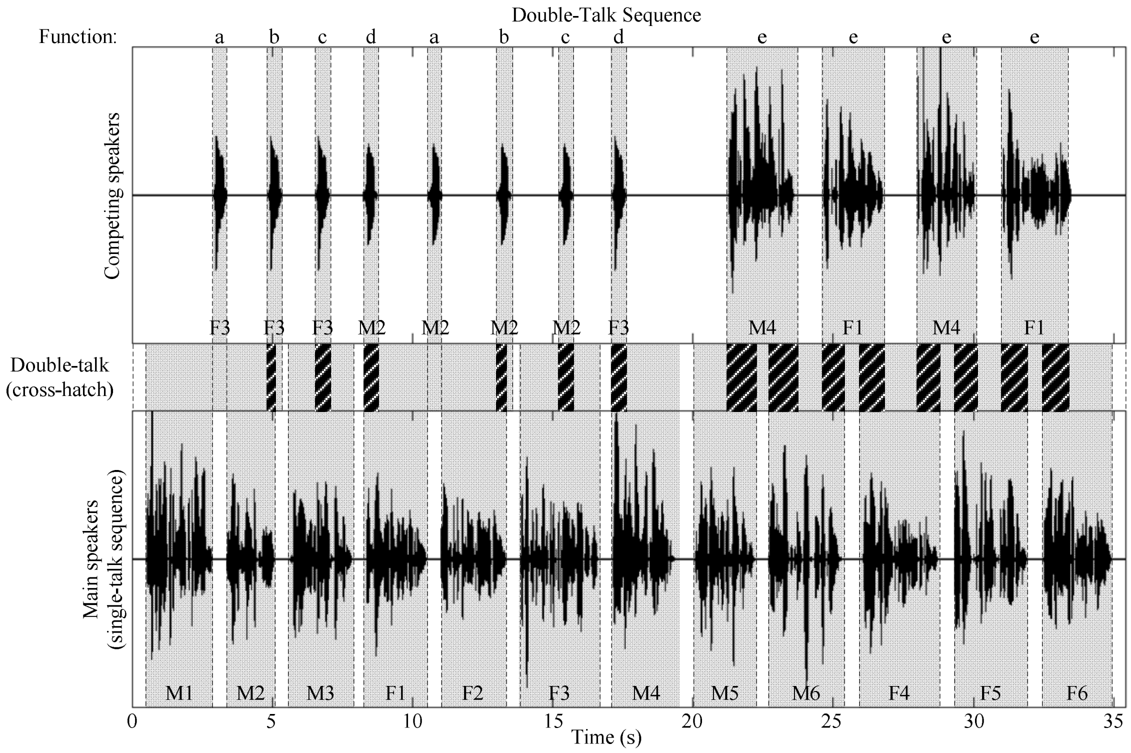
The requirements and the detailed measurement method are for further study.

6.6.4 Double talk performance

Requirements are for further study.

Measurement method

To assess double talk performance, the signals to be used are defined in Recommendation ITU-T P.501 [1], Amendment 1: A "double-talk" sequence representing typical double talk scenarios in real conversations is shown in figure 6.3.4. This uses the single-talk sequence described in section 7.3.1 of Recommendation ITU-T P.501 [1], Amendment 1, shown in the lower pane, as the main speech and an additional competing speaker sequence, shown in the upper pane.



NOTE: Cross-hatched areas between the upper and lower panes show periods of double talk.

Figure 6.6.4: Double-talk test sequence using the single-talk sequence and competing speech serving different functions (a - e)

The competing-speaker sequence includes single words (the word "five") spoken by speakers F3 and M2 during the first half of the sequence followed by full sentences by speakers F1 and M4 during the second half of the sequence. No speaker is competing during the sequence.

The competing samples serve different double-talk functions, defined as functions "a" to "e" above the upper pane of figure 6.5.3. The functions are:

a) Competing word within a speech pause.

b) Competing word partially masked.

c) Competing word fully masked within a sentence.

d) Competing word fully masked coincident with the start of a sentence.

e) Sentence masking another sentence.

These are meant to represent possible double-talk situations in normal conversation. The area between the upper and lower pane of figure 6.3.4 shows the periods during which double-talk happens as cross-hatched patches. The competing sequence can be used either as a send signal or a receive signal in testing.

6.6.5 Stability loss

Requirement

With the handset lying on and) facing a hard surface, the attenuation from the digital input to the digital output shall be at least 6 dB at all frequencies in the range of 50 Hz to 16 kHz.

Measurement method

Before the actual test a training sequence of the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [1] is applied. The training sequence level shall be -16 dBm0 in order not to overload the codec.

The test signal is a PN sequence complying with Recommendation ITU-T P.501 [1] with a length of 4 096 points (for the 48 kHz sampling rate) and a crest factor of 6 dB. The duration of the test signal is 250 ms. With an input signal of 3 dBm0, the attenuation from digital input to digital output shall be measured for frequencies from 100 Hz to 8 kHz under the following conditions:

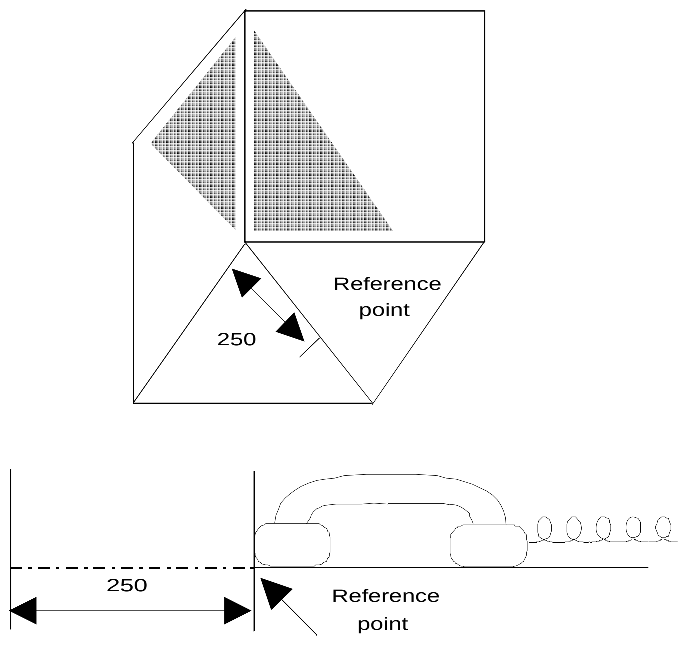
a) the handset, with the transmission circuit fully active, shall be positioned on one inside surface that is of three perpendicular planes, smooth, hard surfaces forming a corner. Each surface shall extend 0,5 m from the apex of the corner. One surface shall be marked with a diagonal line, extending from the corner formed by the three surfaces, and a reference position 250 mm from the corner, as shown in figure 4;

b) the handset, with the transmission circuit fully active, shall be positioned on the defined surface as follows:

    1) the mouthpiece and ear cup shall face towards the surface;

    2) the handset shall be placed centrally, the diagonal line with the ear cup nearer to the apex of the corner;

    3) the extremity of the handset shall coincide with the normal to the reference point, as shown in figure 4;



NOTE: All dimensions in mm.

Figure 6.6.5

6.6.6 Sidetone delay

Requirement

The maximum sidetone-round-trip delay shall be ≤ 5 ms, measured in an echo-free setup.

Measurement method

The handset terminal is setup as described in clause 5.1.

The test signal is a CS-signal complying with Recommendation ITU-T P.501 [1] using a PN sequence with a length of 4 096 points (for the 48 kHz sampling rate) which equals to the period T. The duration of the complete test signal is as specified in Recommendation ITU-T P.501 [1].

The level of the signal shall be -4,7 dBPa at the MRP.

The cross-correlation function Φxy(τ) between the input signal $S_x(t)$ generated by the test system in send direction and the output signal $S_y(t)$ measured at the artificial ear is calculated in the time domain:

$$\Phi_{xy}(\tau) = \lim_{T\to\infty} \sum_{t=-T/2}^{T/2} S_x(t) S_y(t+\tau) \tag{1}$$

The measurement window T shall be exactly identical with the time period T of the test signal, the measurement window is positioned to the pn-sequence of the test signal.

The sidetone delay is calculated from the envelope E(τ) of the cross-correlation function Φxy(τ). The first maximum of the envelope function occurs in correspondence with the direct sound produced by the artificial mouth, the second one occurs with a possible delayed sidetone signal. The difference between the two maxima corresponds to the sidetone delay. The envelope E(τ) is calculated by the Hilbert transformation H {xy(τ)} of the cross-correlation:

$$H\{xy(\tau)\} = \sum_{-\infty}^{\infty} \frac{\Phi_{xy}(u)}{\pi(\tau-u)} \tag{2}$$

$$E(\tau) = \sqrt{\left[\Phi_{xy}(\tau)\right]^2 + \left[H\{xy(\tau)\}\right]^2} \tag{3}$$

It is assumed that the measured sidetone delay is less than T/2.

6.6.7 SideTone Masking Rating STMR (mouth to ear)

Requirement

The STMR shall be 18 dB ± 2 dB for nominal setting of the volume control.

For all other positions of the volume control, the STMR shall not be below 8 dB.

NOTE 1: It is preferable to have a constant STMR independent of the volume control setting.

NOTE 2: STMR measurement in P.79 [5] is not defined above 8 kHz, but sidetone signal is not supposed to have such limitation.

Measurement method

The test signal is defined in 5.1.4. The test signal level shall be -4,7 dBPa, measured at the MRP. The headset terminal is setup as described in clause 5.1.

Where a user operated volume control is provided, the measurements shall be carried out the nominal setting of the volume control. In addition the measurement is repeated at the maximum volume control setting.

Measurements shall be made at one twelfth-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [22] for frequencies from 100 Hz to 8 kHz inclusive. For the calculation the averaged measured level at each frequency band (Recommendation ITU-T P.79 [5] table 3, bands 1 to 20) is referred to the averaged test signal level measured in each frequency band.

The Sidetone path loss (LmeST), as expressed in dB, and the SideTone Masking Rate (STMR) (in dB) shall be calculated from the formula 5-1 of Recommendation ITU-T P.79 [5], using m = 0,225 and the weighting factors as in table 3 of Recommendation ITU-T P.79 [5].

6.6.8 Speech quality in presence of noise

For further study.

6.7 Additional requirements

When switching from handset to handsfree function (and vice-versa), local and distant users should perceive similar performance, and in particular equivalent loudness.

Annex A (informative): Bibliography

- ETSI SR 002 959: "Electronic Working Tools; Roadmap including recommendations for the deployment and usage of electronic working tools in the ETSI standardization process".
- STQ(13)42-30: "Superwideband and fullband testing. Performance characteristics of the Head Acoustics HMS II.3 Artificial Head".

History

Document history		
V1.1.1	March 2014	Publication