ETSI TS 103 740 V1.4.1 (2021-10)

TECHNICAL SPECIFICATION

Speech and multimedia Transmission Quality (STQ);
Transmission requirements for wideband
mobile wireless terminals (hands-free)
from a QoS perspective as perceived by the user

Reference

RTS/STQ-287-4

Keywords

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 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are 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 (https://ipr.etsi.org/).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, 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.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

BLUETOOTH® is a trademark registered and owned by Bluetooth SIG, Inc.

Foreword

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

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Introduction

The present document covers mobile wireless speech terminals. It aims to enhance the interoperability and end-to-end quality with all other types of terminals.

The advanced signal processing of terminals is targeted to speech signals. Therefore, wherever possible speech signals are used for testing in order to achieve the most realistic test conditions and meaningful results.

It is the aim to optimize the listening and talking quality, conversational performance, as well as the use in noisy environments. Related requirements and test methods are defined in the present document.

In contrast to other standards which define minimum performance requirements, it is the intention of the present document to specify terminal equipment requirements, which enable manufacturers and service providers to ensure good end-to-end speech quality as perceived by the user.

Most of the requirements available in the present document are codec-independent and ensure a high compatibility across access networks with all types of terminals.

For all the functions, the present document considers the limitations in audio performance due to different form factors (e.g. size, shape).

1 Scope

The present document provides speech transmission performance requirements for wireless terminals; it addresses several types of mobile wireless terminals, including softphones. The present document addresses hands-free functionality of wideband wireless terminals.

Test methods and performance requirements apply (but are not limited) to wireless terminals equipped with the following mobile network access functionalities:

- Circuit-switched telephony in mobile networks like e.g. GSM/2G, 3G/UMTS.
- Packet-switched mobile networks like e.g. LTE/4G, NR/5G, WLAN and WIMAX™ in conjunction with the associated telephony services like e.g. VoLTE, VoNR or VoWiFi.

The terminal may include additional analogue (e.g. electrical headset connection) or digital (e.g. Bluetooth®) interfaces (wired or wireless) between POI and acoustic test equipment. Requirements and test methods only apply to the whole terminal, i.e. the full electro-acoustic path between talker/listener and network access. Specific requirements and test methods exclusively on such additional interfaces are for further study and out of scope for the present document.

Terminals equipped with the following network access functionalities are out of scope:

- DECT - corresponding tests and requirements can be found in ETSI EN 300 176-2 [i.4]
- Wired VoIP telephony - corresponding tests and requirements can be found in ETSI ES 202 740 [i.5]
- Wireless VoIP telephony like e.g. OTT applications - corresponding tests and requirements can be found in ETSI ES 202 740 [i.5]

2 References

2.1 Normative 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 referenced document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at https://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.

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

[1] Void.

[2] ETSI TS 103 224: "Speech and multimedia Transmission Quality (STQ); A sound field reproduction method for terminal testing including a background noise database".

[3] Recommendation ITU-T G.122: "Influence of national systems on stability and talker echo in international connections".

[4] Void.

[5] Void.

[6] Recommendation ITU-T G.131: "Talker echo and its control".

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

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

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

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

[11] Recommendation ITU-T P.341: "Transmission characteristics for wideband digital loudspeaking and hands-free telephony terminals".

[12] Recommendation ITU-T P.501: "Test signals for use in telephony and other speech-based applications".

[13] Recommendation ITU-T P.502: "Objective test methods for speech communication systems using complex test signals".

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

[15] IEC 61260-1: "Electroacoustics - Octave-band and fractional-octave-band filters - Part 1: Specifications".

[16] Void.

[17] Recommendation ITU-T G.722.2: "Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)".

[18] ETSI TS 126 171 (V6.0.0): "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); AMR speech codec, wideband; General description (3GPP TS 26.171 version 6.0.0 Release 6)".

[19] IEC 61672 (all parts): "Electroacoustics - Sound level meters".

[20] ETSI TS 103 106: "Speech and multimedia Transmission Quality (STQ); Speech quality performance in the presence of background noise: Background noise transmission for mobile terminals-objective test methods".

[21] Recommendation ITU-T P.1110: "Wideband hands-free communication in motor vehicles".

[22] ETSI TS 126 441: "Universal Mobile Telecommunications System (UMTS); LTE; Codec for Enhanced Voice Services (EVS); General overview (3GPP TS 26.441)".

[23] Recommendation ITU-T P.1010: "Fundamental voice transmission objectives for VoIP terminals and gateways".

[24] Recommendation ITU-T P.863: "Perceptual objective listening quality prediction".

[25] Recommendation ITU-T P.863.1: "Application guide for Recommendation ITU-T P.863".

[26] Recommendation ITU-T P.700: "Calculation of loudness for speech communication".

[27] ETSI TS 126 132: "Universal Mobile Telecommunications System (UMTS); LTE; 5G; Speech and video telephony terminal acoustic test specification (3GPP TS 26.132)".

2.2 Informative 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 referenced document (including any amendments) applies.

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

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.

[i.1] Void.

[i.2] ETSI EG 201 377-1: "Speech and multimedia Transmission Quality (STQ); Specification and measurement of speech transmission quality; Part 1: Introduction to objective comparison measurement methods for one-way speech quality across networks".

[i.3] Void.

[i.4] ETSI EN 300 176-2: "Digital Enhanced Cordless Telecommunications (DECT); Test specification; Part 2: Audio and speech".

[i.5] 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".

[i.6] ETSI ES 202 737: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for narrowband VoIP terminals (handset and headset) from a QoS perspective as perceived by the user".

3 Definitions of terms, symbols and abbreviations

3.1 Terms

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

artificial ear: device for the calibration of earphones incorporating an acoustic coupler and a calibrated microphone for the measurement of the sound pressure and having an overall acoustic impedance similar to that of the median adult human ear over a given frequency band

codec: combination of an analogue-to-digital encoder and a digital-to-analogue decoder operating in opposite directions of transmission in the same equipment

ear-Drum Reference Point (DRP): point located at the end of the ear canal, corresponding to the ear-drum position

freefield equalization: artificial head equalized for frontal sound incidence in anechoic conditions

freefield reference point: point located in the free sound field, at least in 1,5 m distance from a sound source radiating in free air (in case of a head and torso simulator [HATS] in the centre of the artificial head with no artificial head present)

hands-free telephony terminal: telephony terminal using a loudspeaker associated with an amplifier as a telephone receiver and which can be used without a handset

HATS Hands-Free Reference Point (HATS HFRP): reference point "n" from Recommendation ITU-T P.58 [8]: "n" is one of the points numbered from 11 to 17 and defined in table 6a of Recommendation ITU-T P.58 [8] (coordinates of far field front point)

NOTE: The HATS HFRP depends on the location(s) of the microphones of the terminal under test: the appropriate axis lip-ring/HATS HFRP is to be as close as possible to the axis lip-ring/HFT microphone under test.

Head And Torso Simulator (HATS) for telephonometry: manikin extending downward from the top of the head to the waist, designed to simulate the sound pick-up characteristics and the acoustic diffraction produced by a median human adult and to reproduce the acoustic field generated by the human mouth

loudspeaking function: function of a handset telephone using a loudspeaker associated with an external amplifier as a telephone receiver

Mouth Reference Point (MRP): point located on axis and 25 mm in front of the lip plane of a mouth simulator

nominal setting of the volume control: when a receive volume control is provided, setting which is closest to the nominal RLR

NOTE: If no user operable volume control is available, this should be noted in the test report.

softphone: speech communication system based upon a computer

3.2 Symbols

Void.

3.3 Abbreviations

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

a.c.	alternative current
$A_{H,S,dt}$	attenuation range in send direction during double talk
AM-FM	Amplitude and Frequency Modulated
AMR-WB	Adaptive MultiRate codec - Wideband
CS	Composite Source
CSS	Composite Source Signal
DECT	Digital Enhanced Cordless Telecommunications
DRP	ear Drum Reference Point
DUT	Device Under Test
EC	Echo Cancellation
EL	Echo Loss
EVS	Enhanced Voice Services
EVS-WB	Enhanced Voice Services Wideband mode
FFT	Fast Fourier Transformation
G-MOS-LQOw	Overall transmission quality for wideband systems
GSM	Global System for Mobile telecommunication
HATS	Head And Torso Simulator
HF	Hands-Free
HFRP	Hands-Free Reference Point
HFT	Hands-Free Terminal
IEC	International Electrotechnical Commission
ITU-T	International Telecommunication Union - Telecommunication standardization sector
LQO	Listening Quality Objective
LTE	Long Term Evolution (3GPP)
MOS	Mean Opinion Score
MRP	Mouth Reference Point
NLP	Non Linear Processor
N-MOS-LQOw	Transmission quality of the background noise for wideband systems
NR/5G	New Radio/5G
OTT	Over The Top solutions
PDA	Personal Digital Assistant
PN	Pseudo random Noise
POI	Point Of Interconnect
QoS	Quality of Service
RF	Radio Frequency
RLL	Receive Loudness Level
RLR	Receive Loudness Rating
RMS	Root Mean Square
SLL	Send Loundness Level
SLR	Send Loudness Rating
S-MOS-LQOw	Transmission quality of the speech for wideband systems
TCL	Terminal Coupling Loss
TOSQA	Telecom Objective Speech Quality Assessment
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
VAD	Voice Activity Detector
VoLTE	Voice over LTE
VoNR	Voice over New Radio
VoWiFi	Voice over WiFi
WB	Wideband

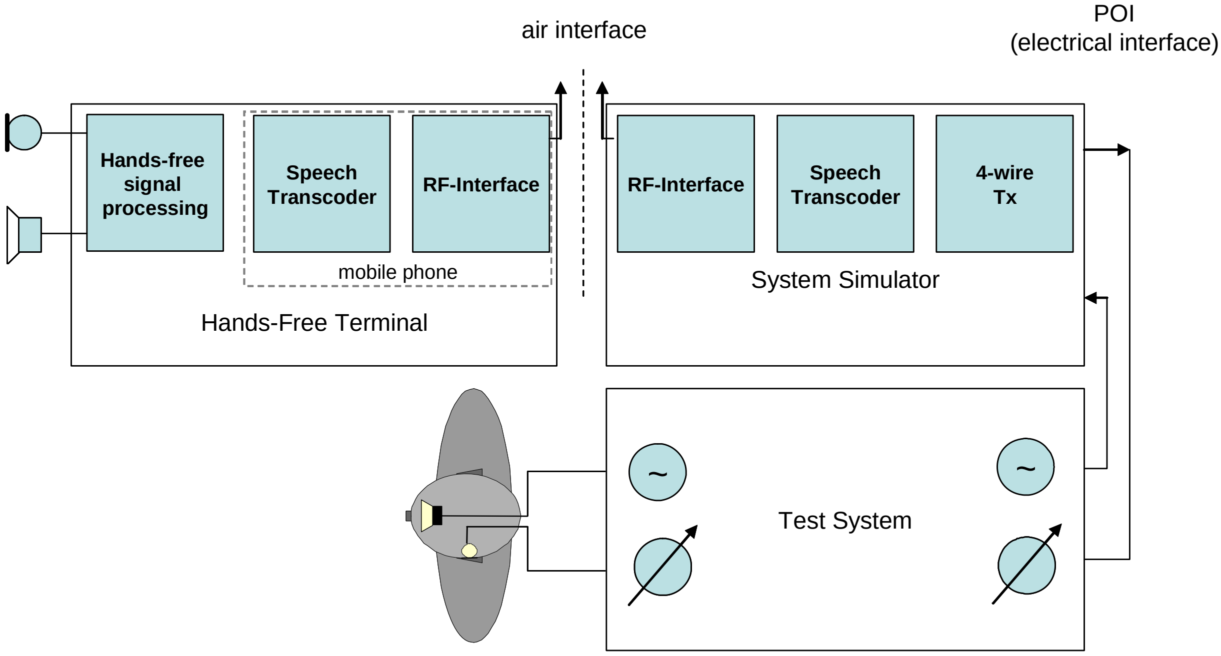
WCDMA	Wideband Code Division Multiple Access
WiFi®	Wireless Fidelity
WIMAX™	Worldwide Interoperability for Microwave ACCess
WLAN	Wireless Local Area Network

4 Void

5 Test Configurations

5.1 Set-up interface

The generic schematic as defined in figure 5.1-1 is applicable to any wireless link.



NOTE: The "whole" terminal includes all the components from "RF interface" to the transducers and may include an additional (radio) link. The air interface considered in the figure is not the additional radio link.

Figure 5.1-1: Set-up interface

5.2 Set-up for terminals

5.2.0 Overview

For electroacoustical testing, HATS as described in Recommendation ITU-T P.58 [8] shall be used, which is equipped with a mouth simulator and two artificial ears. In general, most measurements described in the present document can be conducted with one artificial ear. However, for some measurements, the usage of both artificial ears is required.

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.

When a codec with variable bit rate is used for testing terminal electroacoustical parameters, the bit rate giving the best characteristics or the most commonly used should be selected, e.g.:

- AMR-WB [18]: 12,65 kbit/s;
- EVS-WB [22]: 13,2 kbit/s.

For packet-switched network access, prior to the actual measurements, the clock skew between terminal and test system shall be compensated by adjusting the clock of the test equipment to match the clock of the terminal. The inaccuracy of the clock skew adjustment shall be less than 1 ppm measured according to the procedure in Annex D of ETSI TS 126 132 [27].

5.2.1 Handheld terminal

HATS measurement equipment shall be configured to the Handheld hands-free UE according to figure 5.2.1-1. The HATS should be positioned so that the HATS Reference Point is at a distance d_{HF} from the centre point of the visual display of the Mobile Station. The distance d_{HF} is specified by the manufacturer. A vertical angle θ_{HF} may be specified by the manufacturer. In case it is not specified the distance d_{HF} shall be 42 cm and θ_{HF} shall be 0.

NOTE: The nominal distance of 42 cm corresponds to lip plane-HATS reference point distance (12 cm) with an additional 30 cm giving a realistic figure as a reference usage of handheld terminals.

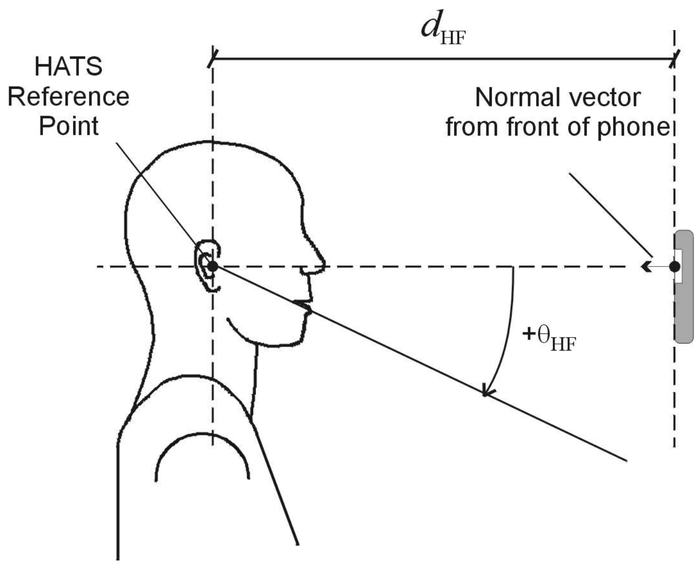


Figure 5.2.1-1: Configuration of Handheld Hands-free UE relative to the HATS

5.2.2 Vehicle mounted hands-free

Test arrangement, test methods and performance requirements are according to Recommendation ITU-T P.1110 [21].

5.2.3 Desktop hands-free terminal

For HATS test equipment, definition of hands-free terminals and setup for desktop hands-free terminals is based on Recommendation ITU-T P.581 [14].

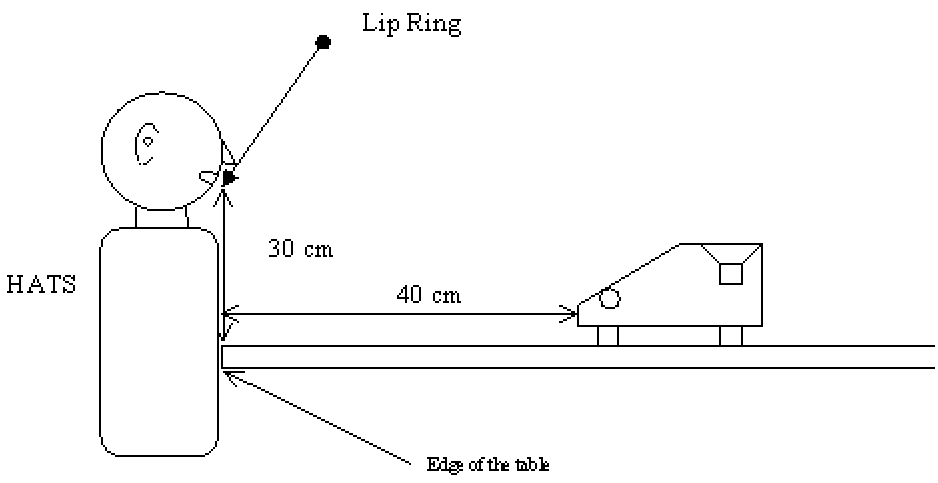


Figure 5.2.3-1: Position for test of desktop hands-free terminal side view

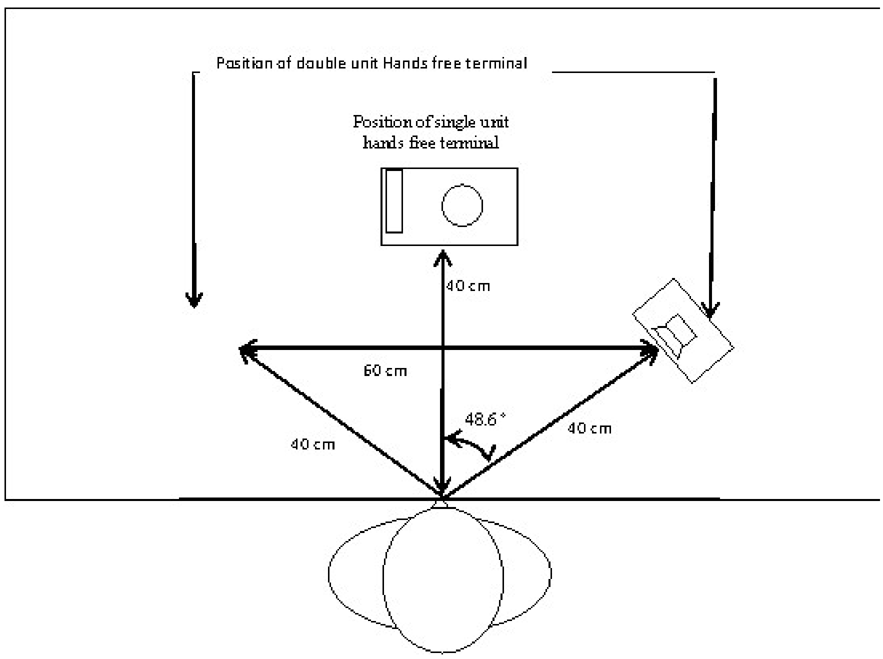


Figure 5.2.3-2: Position for test of desktop hands-free terminal top sight

5.2.4 Additional test setup for handsfree function with softphone

5.2.4.0 Overview

Two types of softphones are to be considered:

- Type 1 is to be used as a desktop type (e.g. notebook).
- Type 2 is to be used as a handheld type (e.g. PDA).

When manufacturer gives conditions of use, they will apply for test. If no other requirement is given by manufacturer softphone will be positioned according the following conditions.

5.2.4.1 Softphone including speakers and microphone

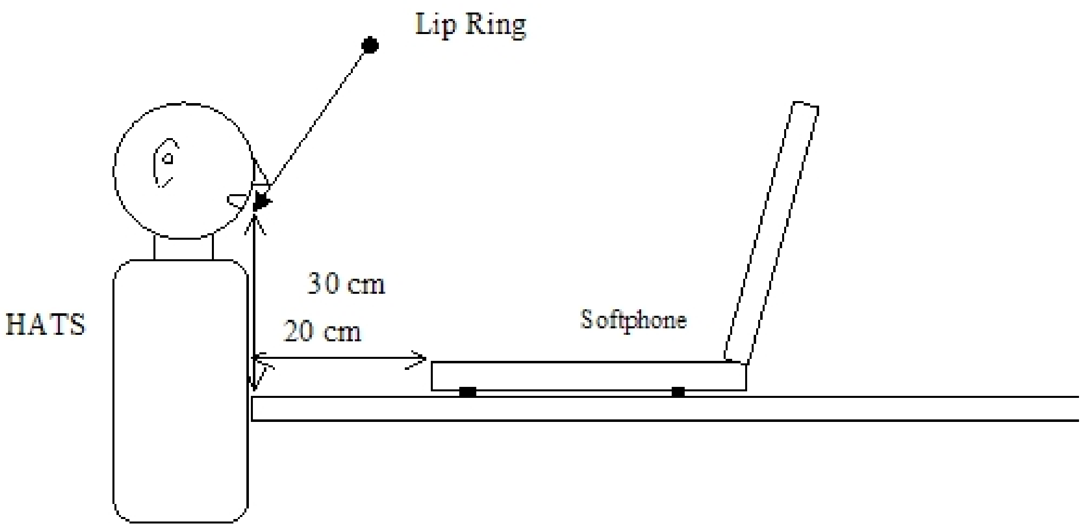


Figure 5.2.4.1-1: Configuration of softphone relative to the HATS side view

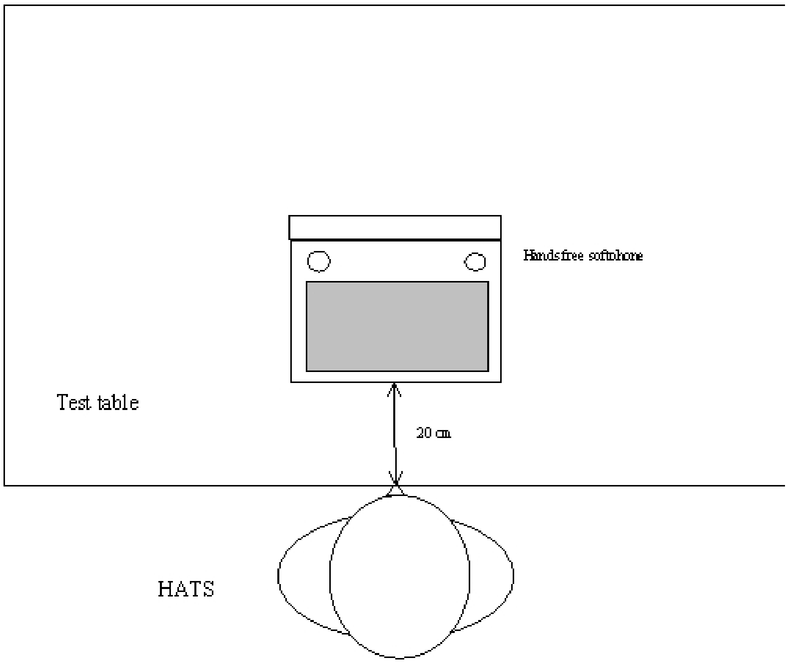


Figure 5.2.4.1-2: Configuration of softphone relative to the HATS-top sight

5.2.4.2 Softphone with separate speakers

When separate loudspeakers are used, system will be positioned as in figure 5.2.4.2-1.

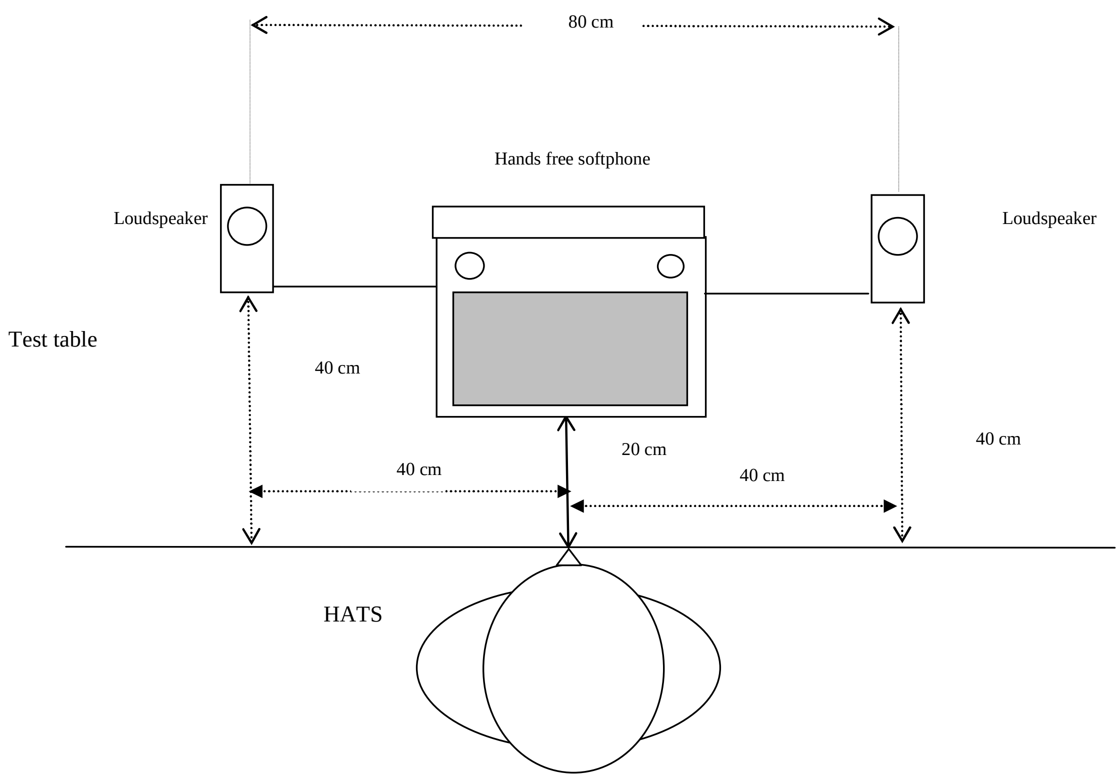


Figure 5.2.4.2-1: Configuration of softphone using external speakers relative to the HATS-top sight

When external microphone and speakers are used, system will be positioned as in figure 5.2.4.2-2.

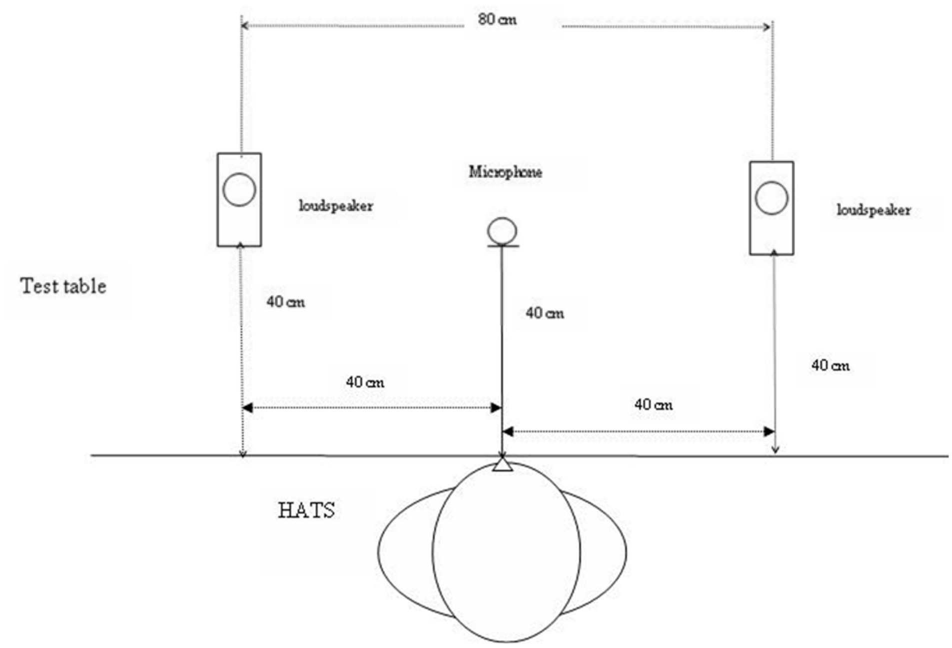


Figure 5.2.4.2-2: Configuration of softphone using external speakers and microphone relative to the HATS-top sight

5.2.5 Test setup for variable echo path

Test setup for desktop hands-free terminals: A notebook is positioned at least 20 cm in front of the device (or devices) with the transducers, as shown in figure 5.2.5-1. The notebook lid is moved during the measurement.

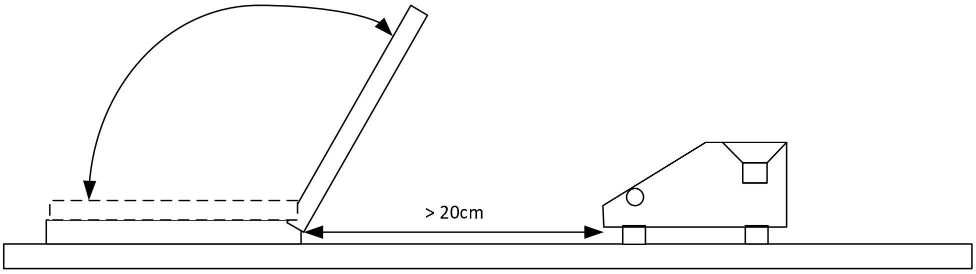


Figure 5.2.5-1: Positioning of DUT

Test setup for softphone: The test setup is described in clause 6.2. The notebook lid is moved during the measurement, as shown in figure 5.2.5-2. This setup is valid for all combinations of notebook with or without external speakers or microphone.

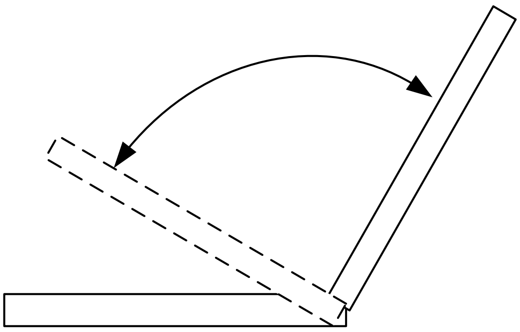


Figure 5.2.5-2: Positioning of DUT

Test setup for other handsfree devices is for further study.

5.3 Acoustical environment

In general, different acoustical environments have 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. Considering this, the test laboratory, in the case where its test room does not conform to anechoic conditions as given in Recommendation ITU-T P.341 [11], has to present difference in results for measurements due to its test room. In case where an anechoic room is not available the test room has to be an acoustically treated room with few reflections and a low noise level.

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.

5.4 Test signals

Modern wireless terminals often deploy non linear and time-varying processing. As such terminals are designed for speech transmission, the most appropriate test signal is real speech. Appropriate test signals (general description) are defined in Recommendation ITU-T P.501 [12].

For testing the wideband telephony service provided by a terminal the test signal used shall be band limited between 100 Hz and 8 kHz with a bandpass filter providing a minimum of 24 dB/Oct. filter roll off, when feeding into the receive direction.

Unless specified otherwise, the test signal levels are referred to the average level of the (band limited in receive direction) test signal, averaged over the complete test sequence. Unless specified otherwise, the active speech level calculation according to Recommendation ITU-T P.56 [7] shall be used.

5.5 Calibration

5.5.1 Send

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

The various steps for calibration of the artificial mouth are described in Recommendation ITU-T P.581 [14]. The spectrum of acoustic signal produced by the artificial mouth is equalized under freefield conditions at the MRP.

The test setup shall be in conformance with figure 5.5.1-1 but, depending on the type of terminal, the appropriate distance and level will be used. When using this calibration method, send sensitivity shall be calculated as follows:

$$S_{\mathrm{mJ}} = 20\log\mathrm{V}_s - 20\log\mathrm{P}_{\mathrm{MRP}} + \mathrm{Corr} - D_{corr} \qquad (1)$$

Where:

V_s is the measured voltage across the appropriate termination (unless stated otherwise, a 600 ohm termination).

P_{MRP} is the applied sound pressure at the MRP.

Corr is 20 log (P_{MRP}/P_{HFRP}) of the used artificial mouth.
The value of Corr is the value required to calibrate the artificial mouth to the exact value of D_{corr} (e.g. 24,0 dB for 50 cm distance).

D_{corr} is the correction to achieve the target sound pressure level at the intended distance (see below).

NOTE: Reason for this procedure of calibration in two steps is to take into account the different variation of signal with distance by using different implementations of HATS.

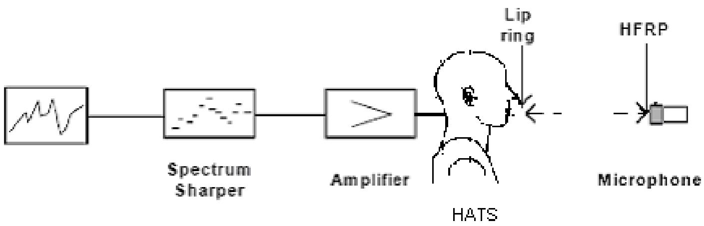


Figure 5.5.1-1: Calibration at HFRP for HATS

The distance used for level calibration corresponds to the following values:

Desktop terminal: 50 cm and level to adjust - 28,7 dBPa, D_{corr} = 24 dB.

Handheld terminal: 30 cm with -24,3 dBPa, D_{corr} = 19,6 dB.

Softphone: 36 cm with -25,8 dBPa, D_{corr} = 21,1 dB.

5.5.2 Receive

Unless specified otherwise, the applied test signal level at the digital input shall be -16 dBm0, measured according to Recommendation ITU-T P.56 [7] at the digital reference point or the equivalent analogue point.

All measurement values produced by HATS are intended to be free-field equalized according Recommendation ITU-T P.581 [14].

5.5.3 Setup of background noise simulation

A setup for simulating realistic background noises in a lab-type environment is described in ETSI TS 103 224 [2].

ETSI TS 103 224 [2] contains a description of the recording arrangement for realistic background noises, a description of the setup for a loudspeaker arrangement suitable to simulate a background noise field in a lab-type environment and a database of realistic background noises, which can be used for testing the terminal performance with a variety of different background noises.

The principle loudspeaker setup for the simulation arrangement is shown in figure 5.5.3-1.

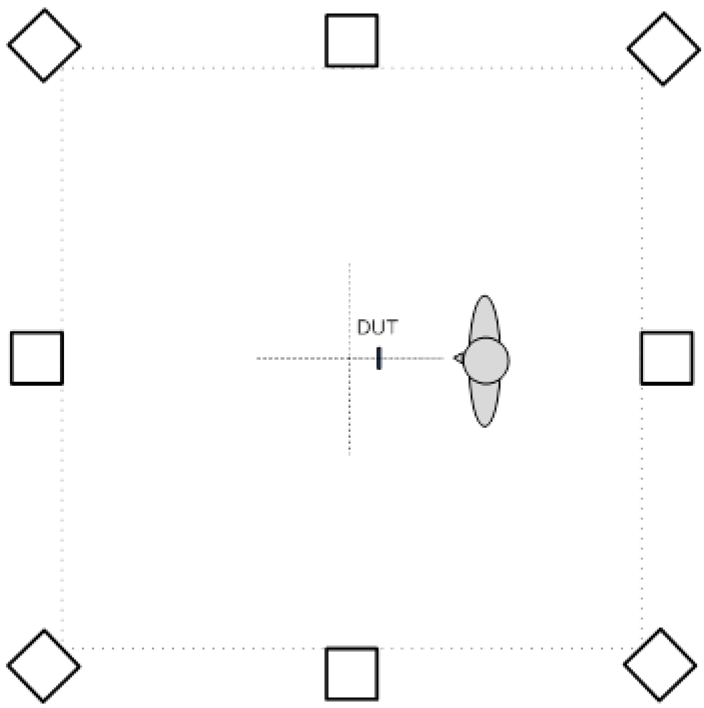


Figure 5.5.3-1: Loudspeaker arrangement for background noise simulation

The equalization and calibration procedure for the setup is described in detail in ETSI TS 103 224 [2].

If not stated otherwise this setup is used in all measurements where background noise simulation is required.

The following noises of ETSI TS 103 224 [2] in table 5.5.3-1 shall be used.

Table 5.5.3-1: Noises used for background noise simulation

Name	Description	Length	Handsfree Levels
Full-size car 130 km/h (FullSizeCar_130)	HATS and microphone array at co-drivers position	30 s	1: 69,5 dB 2: 68,6 dB 3: 68,6 dB 4: 68,7 dB 5: 68,8 dB 6: 68,8 dB 7: 69,2 dB 8: 69,7 dB
Cafeteria (Cafeteria)	HATS and microphone array inside a cafeteria	30 s	1: 69,0 dB 2: 69,7 dB 3: 69,6 dB 4: 69,8 dB 5: 69,5 dB 6: 69,5 dB 7: 69,7 dB 8: 70,0 dB
Roadnoise (Roadnoise)	HATS and microphone array standing outside near a road	30 s	1: 69,9 dB 2: 70,7 dB 3: 70,9 dB 4: 71,0 dB 5: 70,8 dB 6: 70,8 dB 7: 70,9 dB 8: 71,0 dB
Pub Noise (Pub)	HATS and microphone array in a pub	30 s	1: 75,2 dB 2: 75,1 dB 3: 74,9 dB 4: 75,1 dB 5: 74,8 dB 6: 74,8 dB 7: 74,8 dB 8: 75,0 dB
Airport departure	HATS and microphone array in an airport gate area	30 s	1: 77,2 dB 2: 77,4 dB 3: 77,6 dB 4: 77,7 dB 5: 78,1 dB 6: 77,9 dB 7: 77,8 dB 8: 77,9 dB

5.6 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).

5.7 Accuracy of test equipment

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

Table 5.7-1: Accuracy of measurements

Item	Accuracy
Electrical Signal Level	±0,2 dB for levels ≥ -50 dBV
Electrical Signal Level	±0,4 dB for levels < -50 dBV
Sound pressure	±0,7 dB
Frequency	±0,2 %
Time	±0,2 %
Measured maximum frequency	20 kHz
Clock Accuracy	< 2 ppm
NOTE: The measured maximum frequency is required for free-field correction in Recommendation ITU-T P.58 [8].	

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

Table 5.7-2: Accuracy of test signal generation

Quantity	Accuracy
Sound pressure level at MRP	±3 dB for 100 Hz to 200 Hz ±1 dB for 200 Hz to 4 kHz ±3 dB for 4 kHz to 14 kHz
Electrical excitation levels	±0,4 dB across the whole frequency range
Frequency generation	±2 % (see note)
Time	±0,2 %
NOTE: This tolerance may be used to avoid measurements at critical frequencies, e.g. those due to sampling and coding operations within the terminal under test.	

The measurements results shall be corrected for the measured deviations from the nominal level.

The sound level measurement equipment shall comply with class 1 accuracy according to IEC 61672 [19].

5.8 Power feeding conditions

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.

5.9 Influence of terminal delay on measurements

As delay is introduced by the terminal, care shall be taken for all measurements where exact position of the analysis window is required. It shall be checked that the test is performed on the test signal and not any other signal.

6 Codec independent requirements and associated Measurement Methodologies

6.1 Send and receive frequency response

6.1.1 Send frequency response

Requirements

The send sensitivity frequency response from the MRP to the measurement output (digital or analogue output according measurement system used) shall within the mask which can be drawn with straight lines between the breaking points in table 6.1.1-1 on a logarithmic (frequency) - linear (dB sensitivity) scale.

Table 6.1.1-1: Hands-free send sensitivity/frequency response

Frequency (Hz)	Upper limit (dB)	Lower limit (dB)
100	4	
125	4	-10
200	4	-4
1 000	4	-4
5 000	(see note)	-4
6 300	9	-7
8 000	9	
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (Hz) scale.		

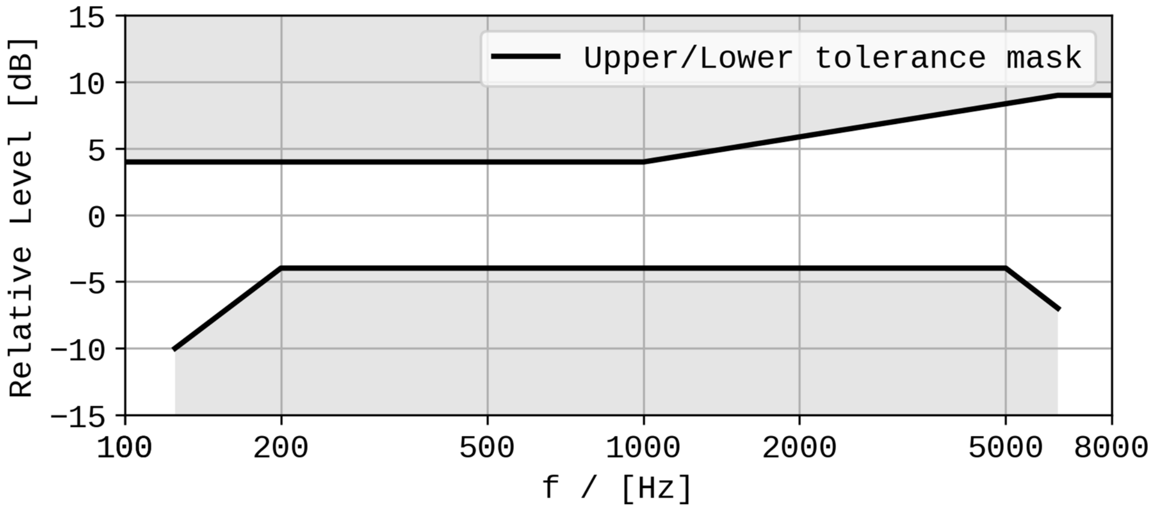


Figure 6.1.1-1: Hands-free send sensitivity/frequency response

Measurement method

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 [12].

The terminal is set up as described in clause 5.2 and calibration is realized as explained in clause 5.5.1.

Measurements shall be made at 1/3-octave intervals as given by IEC 61260-1 [15] 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 dBV/Pa.

6.1.2 Receive frequency response

6.1.2.1 Handheld terminal

Requirements

The receive sensitivity frequency response from the measurement input (digital or analogue input according measurement system used) to ear of HATS free field corrected shall be within the mask which can be drawn with straight lines between the breaking points in table 6.1.2.1-1 on a logarithmic (frequency) - linear (dB sensitivity) scale.

Table 6.1.2.1-1: Handheld terminal receive sensitivity/frequency response

Frequency (Hz)	Upper limit (dB)	Lower limit (dB)
125	6	
400	6	-12
500	6	-6
4 000	6	-6
5 000	6	-9
6 300	6	-12
8 000	6	
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (Hz) scale.		

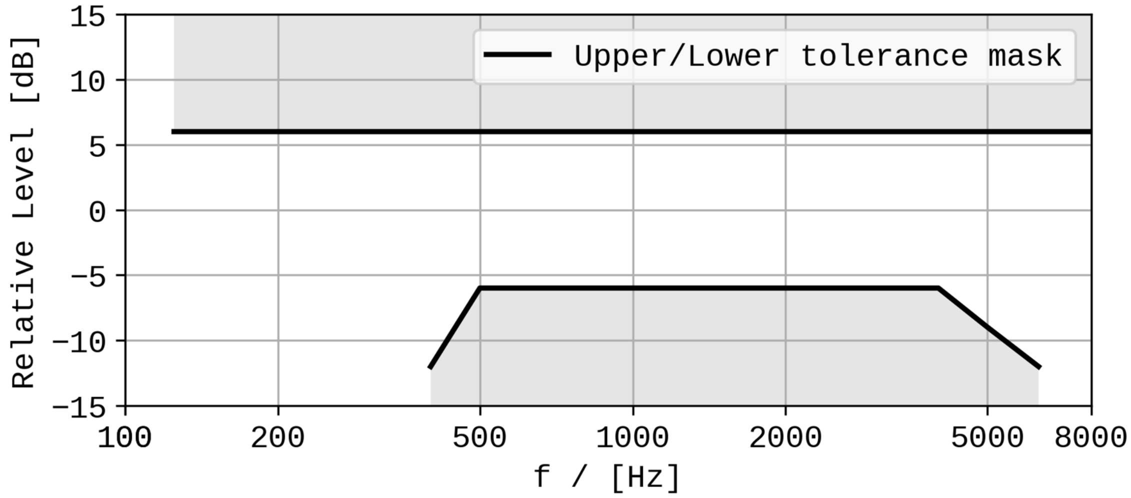


Figure 6.1.2.1-1: Handheld receive sensitivity/frequency response

Measurement method

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

$$S_{Jeff} = 20 \log (pe_{ff} / v_{RCV}) \text{ dB rel 1 Pa / V} \quad (2)$$

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

pe_{ff} DRP Sound pressure measured by ear simulator Measurement data are converted from the Drum Reference Point to freefield.

v_{RCV} Equivalent RMS input voltage.

The test signal used for the measurements shall be the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [12].

The terminal is set up as described in clause 5.2.

The measurement is conducted at nominal volume control setting.

Measurements shall be made at 1/3-octave intervals as given by IEC 61260-1 [15] 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.

6.1.2.2 Void

6.1.2.3 Softphone (computer-based terminals)

Type 1 or softphone with external speakers: The requirement as for desktop terminals applies.

Type 2: The requirement as for handheld terminals applies.

6.1.2.4 Desktop Terminal

Requirements

The receive sensitivity frequency response from the measurement input (digital or analogue input according measurement system used) to ear of HATS free field corrected shall be within the mask which can be drawn with straight lines between the breaking points in table 6.1.2.4-1 on a logarithmic (frequency) - linear (dB sensitivity) scale.

Table 6.1.2.4-1: Desktop receive sensitivity/frequency response

Frequency	Upper limit	Lower limit
125 Hz	8 dB	
200 Hz	8 dB	-12 dB
250 Hz	8 dB	-9 dB
315 Hz	7 dB	-6 dB
400 Hz	6 dB	-6 dB
5 000 Hz	6 dB	-6 dB
6 300 Hz	6 dB	-9 dB
8 000 Hz	6 dB	

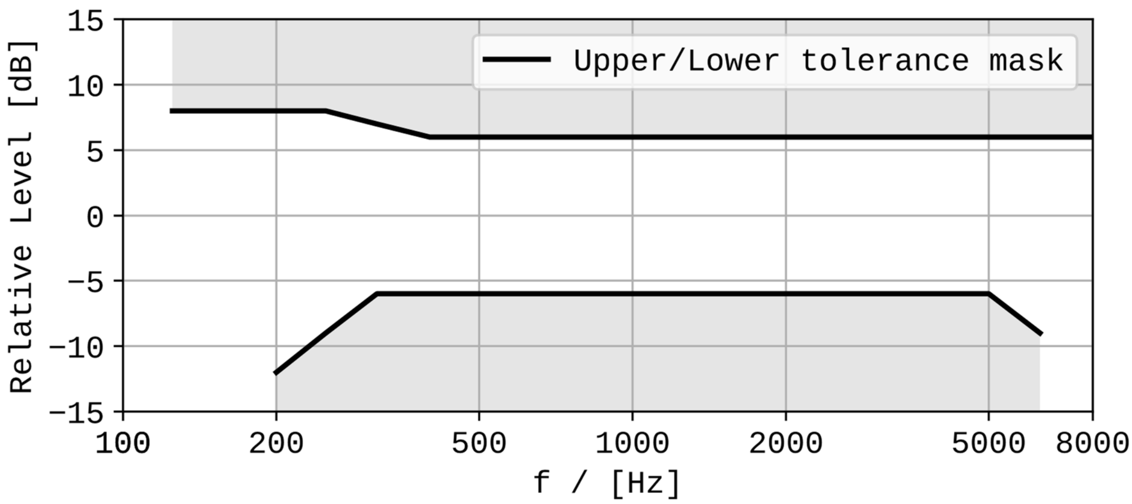


Figure 6.1.2.4-1: Desktop receive sensitivity/frequency response

Measurement method

The terminal is set up as described in clause 5.2.

The measurement method described in clause 6.1.2.1 shall apply.

6.2 Send and receive loudness ratings

6.2.1 Send Loudness Ratings

Requirement

The nominal values of SLR shall be:

$$SLR = +13\ dB \pm 3\ dB \quad (3)$$

Measurement method

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 [12].

The terminal is positioned as described in clause 5.2 and calibration is realized as explained in clause 5.5.1.

The send sensitivity shall be calculated for each of the 20 frequency bands given in table A.2 of Recommendation ITU-T P.79 [9] (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 [9], formula (A - 23b), over bands 1 to 20 and the send weighting factors from Recommendation ITU-T P.79 [9], Annex A, table A.2.

6.2.2 Microphone (Mic) mute

Requirement

The SLR (Send Loudness Rating) with mic mute on shall be at least 50 dB higher than with mic mute off.

Measurement method

The terminal is set up as described in clause 6.2.1, but its microphone shall be configured to be muted.

Measurement and calculation method are the same as in clause 6.2.1.

6.2.3 Receive Loudness Ratings

Requirements

Handheld terminal

Nominal value of RLR shall be 9 dB ± 3 dB. This value has to be fulfilled for at least one volume setting.

Value of RLR at maximum volume setting shall be less than (louder) or equal to +5 dB.

Range of volume control (RLR measured at minimum and maximum volume setting) shall be larger than or equal to 15 dB.

Desktop terminal

Nominal value of RLR shall be 5 dB ± 3 dB. This value has to be fulfilled for at least volume setting.

The value of RLR at the upper part of the volume range shall be less than (louder) or equal to -2 dB.

The range of volume control (RLR measured at minimum and maximum volume setting) shall be larger than or equal to 15 dB.

Softphone (computer-based terminal)

Type 1 or softphone with external speakers: requirement as for desktop terminal.

Type 2: requirement as for handheld terminal.

Measurement method

The terminal is set up as described in clause 5.2.

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 [12].

The receive sensitivity shall be calculated for each of the 20 frequency bands given in table A.2 of Recommendation ITU-T P.79 [9] (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 [9], formula (A-23c), over bands 1 to 20 and the receive weighting factors from table A.2 of Recommendation ITU-T P.79 [9]. No leakage correction shall be applied.

For binaural measurements, the individual sensitivities for left and right ears are energetically summed up. The hands-free RLR based on this overall sensitivity is then calculated with a correction factor of -8 dB.

The test shall be repeated for maximum and minimum volume control setting.

6.2.4 Send Loudness Level

Requirements

The nominal value of Send Loudness Level (SLL) shall be:

$$SLL = 71 \text{ phon} \pm 4 \text{ phon}$$

Measurement method

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 [12].

The terminal is set up as described in clause 5.2 and calibration is realized as explained in clause 5.5.1.

The loudness (in sone) of the recorded signal is calculated according to Recommendation ITU-T P.700 [26] and noted in the test report. The loudness level (in phon) is calculated according to clause 9 of Recommendation ITU-T P.700 [26].

6.2.5 Receive Loudness Level

Requirements

Handheld terminal

The nominal value of Receive Loudness Level (RLL) shall be 67 ± 4 phon. This value has to be fulfilled for at least one volume setting.

Value of RLL at maximum volume setting shall be louder than or equal to 71 phon.

Range of volume control (RLL measured at minimum and maximum volume setting) shall be larger than or equal to 15 phon.

Desktop terminal

The nominal value of Receive Loudness Level (RLL) shall be 71 ± 4 phon. This value has to be fulfilled for at least one volume setting.

Value of RLL at maximum volume setting shall be louder than or equal to 78 phon.

Range of volume control (RLL measured at minimum and maximum volume setting) shall be larger than or equal to 15 phon.

Softphone (computer-based terminal)

Type 1 or softphone with external speakers: requirement as for desktop terminal.

Type 2: requirement as for handheld terminal.

Measurement method

The terminal is set up as described in clause 5.2.

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 [12].

For each recorded artificial ear signal, the loudness (in sone) of the recorded signal is calculated according to Recommendation ITU-T P.700 [26] and noted in the test report. The loudness level (in phon) is determined as follows for binaural measurements: the resulting loudnesses for left and right ears are first halved (individual loudness per ear). Both loudnesses are added (assuming perfect loudness summation). With this overall loudness, the overall loudness level is finally determined according to clause 8.2 of Recommendation ITU-T P.700 [26].

6.3 Send and receive noise

6.3.1 Send Noise

Requirements

The send noise shall not exceed -64 dBm0(A).

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

Requirement as for other tests is identical for all types of terminals.

NOTE: Softphones with cooling devices (fans) can produce a rather high level of noise, furthermore largely dependent of activity of system.

Measurement method

The terminal will be positioned as described in clause 5.2.

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 activation signal shall be the female speaker signal of the short conditioning sequence described in clause 7.3.7 of in Recommendation ITU-T P.501 [12]. The level of this activation signal shall be -4,7 dBPa, measured at the MRP. The activation signal level is averaged over the complete activation signal sequence.

The send noise is measured at the POI in the frequency range from 100 Hz to 8 kHz. 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).

Spectral peaks are measured in the frequency domain from 100 Hz to 6,3 kHz. The frequency spectrum of the A-weighted idle channel noise is measured by a spectral analysis having a noise bandwidth of 8,79 Hz (determined using FFT 8 k samples/48 kHz sampling rate with Hanning window or equivalent). The idle channel noise spectrum is stated in dB. A smoothed average idle channel noise spectrum is calculated by a moving average (arithmetic mean) 1/3-octave wide across the idle noise channel spectrum stated in dB (linear average in dB of all FFT bins in the range from 2^-1/6)f to 2^(+1/6)f). Peaks in the idle channel noise spectrum are compared against a smoothed average idle channel noise spectrum.

6.3.2 Receive Noise

Requirements

The receive noise level shall not exceed -54 dBPa(A) at **nominal setting of the volume control**.

The level in any 1/3-octave band, between 100 Hz and 10 kHz shall not exceed a value of -64 dBPa.

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

NOTE 1: Void.

NOTE 2: For softphone the test condition excludes fan noise.

Measurement method

The terminal will be positioned as described in clause 5.2.

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 activation signal shall be the female speaker signal of the short conditioning sequence described in clause 7.3.7 of in Recommendation ITU-T P.501 [12]. The level of this activation signal shall be -16 dBm0.

The A-weighted noise level shall be measured at DRP of the artificial ear with the free field equalization active. The noise level is measured until 10 kHz.

The level in any 1/3-octave band, between 100 Hz and 10 kHz shall be measured at DRP of the artificial ear, including free field equalization.

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

Spectral peaks are measured in the frequency domain in the frequency range from 100 Hz to 6,3 kHz. The frequency spectrum of the A-weighted idle channel noise is measured by a spectral analysis having a noise bandwidth of 8,79 Hz (determined using FFT 8 k samples/48 kHz sampling rate with Hanning window or equivalent). The idle channel noise spectrum is stated in dB. A smoothed average idle channel noise spectrum is calculated by a moving average (arithmetic mean) 1/3-octave wide across the idle noise channel spectrum stated in dB (linear average in dB of all FFT bins in the range from 2^(-1/6)f to 2^(+1/6)f). Peaks in the idle channel noise spectrum are compared against a smoothed average idle channel noise spectrum.

NOTE 3: Care should be taken that only the noise is windowed out by the analysis and the analysis window is not impaired by any remaining reverberance or room noise.

6.4 Send and receive distortion

6.4.0 General

It is not intended to provide codec-dependant requirements but to assess the electroacoustic performances of the terminal.

6.4.1 Send distortion

Requirements

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

Table 6.4.1-1: Limits for harmonic distortion ratio for send

Frequency (Hz)	Signal to harmonic distortion ratio limit, send (dB)
200	25
315	26
400	30
1 000	30
2 000	30
NOTE: The limits for intermediate frequencies lie on straight lines drawn between the given values on a linear (dB) - logarithmic (Hz) scale.	

Measurement method

The terminal will be positioned as described in clause 5.2.

After the correct activation of the system, a sinewave signal at frequencies of 200 Hz, 315 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz, 1 000 Hz and 2 000 Hz is applied. The duration of the sine-wave shall be less than 1 s. The sinusoidal signal level shall be calibrated to -4,7 dBPa at the MRP.

The signal to harmonic distortion ratio is measured selectively up to 7 kHz.

The female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [12] shall be used for activation. The level of this activation signal shall be -4,7 dBPa at the MRP.

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

6.4.2 Receive distortion

Requirement

- **Handheld terminal**

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

Table 6.4.2-1: Limits for harmonic distortion ratio for receive

Frequency	Signal to distortion ratio limit, receive for vehicle mounted or desktop terminal at nominal volume	Signal to distortion ratio limit, receive for handheld terminal at nominal volume	Signal to distortion ratio limit, receive for all terminals at maximum volume
315 Hz	26 dB		
400 Hz	30 dB		
500 Hz	30 dB	20 dB	
800 Hz	30 dB	30 dB	20 dB
1 kHz	30 dB	30 dB	
2 kHz	30 dB	30 dB	
3 kHz	30 dB	30 dB	
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (kHz) scale.			

- **Desktop terminal**

The ratio of signal to harmonic distortion is given in table 6.4.2-1.

- **Softphone (computer-based terminal)**

For a type 1 or softphone with external speakers the requirements given in table 6.4.2-1 (desktop terminal) shall apply.

For a Type 2 terminal the requirements given in table 6.4.2-1 (handheld terminal) shall apply.

Measurement method

Test setup is described in clause 5.2.

After a correct activation of the system, a sinewave signal at frequencies of 315 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz, 1 000 Hz, 2 000 Hz, 3 000 Hz is applied at the digital interface. The duration of the sine-wave shall be less than 1 s. The sinusoidal signal level shall be -16 dBm0.

The female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [12] shall be used for activation. The level of this activation signal shall be -16 dBm0.

The signal to harmonic distortion ratio is measured selectively up to 10 kHz.

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

6.5 Terminal Coupling Loss (TCL)

Requirements

In order to meet talker echo objective requirements [6], the recommended terminal coupling loss during single talk (TCL) should be greater than 55 dB at **nominal setting of the volume control**.

The TCL shall be ≥ 46 dB for all settings of the volume control (if supplied).

NOTE 1: A TCL ≥ 55 dB is recommended as a performance objective. Depending on the idle channel noise in the send direction, it may not always be possible to measure an echo loss ≥ 50 dB.

Measurement method

The setup for terminal is described in clause 5.2.

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 [12]. The signal level shall be -10 dBm0.

The TCL is calculated as the difference between the averaged test signal level and the averaged echo level in the frequency range from 100 Hz to 8 000 Hz. 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). For the measurement, a time window has to be applied which is adapted to the duration of the actual test signal. The echo loss is calculated by the equations:

$$L_e = C - 10\log_{10}\sum_{i=1}^{N}(A_i + A_{i-1})(\log_{10} f_i - \log_{10} f_{i-1}) \quad (4)$$

and

$$C = 10\log_{10}(2\,(\log_{10} f_N - \log_{10} f_0)) \quad (5)$$

where:

- A_0 is the output/input power ratio at frequency $f_0 = 100\ Hz$;
- A_1 the ratio at frequency f_i; and
- A_N the ratio at frequency $f_N = 8\,000\ Hz$.

Equation (4) is a generalized form of the equation defined in clause B.4 of Recommendation ITU-T G.122 [3] for calculating echo loss based on tabulated data, which allows the calculation of echo loss within any frequency range between f_0 and f_N.

The ambient noise level shall be < -64 dBPa(A).

NOTE 2: The extension of the frequency range is for further study.

NOTE 3: Care should be taken when measuring TCL: the echo return not to be masked by the residual noise or the comfort noise when implemented.

6.6 Stability Loss (or similar parameters)

Requirements

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 attenuation shall exceed 6 dB for all frequencies and for all settings of volume control.

Measurement method

Test set-up is identical as for TCL as described in clause 6.5.

Before the actual test a training sequence consisting of the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [12] shall be 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 [12] 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.

6.7 Double talk performance

6.7.0 Overview

During double talk the speech is mainly determined by 2 parameters: impairment caused by echo during double talk and level variation between single and double talk (attenuation range).

In order to guarantee sufficient quality under double talk conditions the Talker Echo Loudness Rating should be high and the attenuation inserted should be as low as possible. Terminals which do not allow double talk in any case should provide a good echo attenuation which is realized by a high attenuation range in this case.

The most important parameters determining the speech quality during double talk are (see Recommendations ITU-T P.340 [10] and P.502 [13]):

- Attenuation range in send direction during double talk $A_{H,S,dt}$.
- Attenuation range in receive direction during double talk $A_{H,R,dt}$.
- Echo attenuation during double talk.

The categorization of a terminal is based on the three categories defined in clauses 6.7.1, 6.7.2 and 6.7.3 and this categorization is given by the "lowest" of the three parameters e.g. if $A_{H,S,dt}$ provides 2a, $A_{H,R,dt}$ 2b and echo loss 1, the categorization of the terminal is 2b.

6.7.1 Attenuation Range in Send Direction during Double Talk $A_{H,S,dt}$

Requirements

Based on the level variation in send direction during double talk $A_{H,S,dt}$ the behaviour of the terminal can be classified according to table 6.7.1-1.

The terminal should have full duplex capability (category 1). For desktop type terminals, category 1 shall be achieved. For handheld type terminals, category 2a shall be achieved.

Table 6.7.1-1: Double talk categories for send direction

Category (according to Recommendation ITU-T P.340 [10])	**1**	**2a**	**2b**	**2c**	**3**
	Full Duplex Capability	Partial Duplex Capability			No Duplex Capability
$A_{H,S,dt}$ [dB]	≤ 3	≤ 6	≤ 9	≤ 12	> 12

The requirement applies with nominal setting of the volume control for nominal signal levels in send and receive directions as well as for the level combinations +6 dB (relative to nominal level) in send/-6 dB (relative to nominal level) in receive and +6 dB (relative to nominal level) in receive/-6 dB (relative to nominal level) in send. The requirement also applies for maximum setting of the volume control with nominal signal levels in send and receive directions.

In general, table 6.7.1-1 provides a quality classification of terminals regarding double talk performance. However, this does not mean that a terminal which is category 1 based on the double talk performance is of high quality concerning the overall quality as well.

The measurement shall be repeated for desktop type terminals with variable echo path.

Measurement method

The terminal is positioned as described in clause 5.2.

The long conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [12] shall be used for conditioning of the terminal, with the female speaker in the receive direction. The test signal to determine the attenuation range during double talk is the double talk speech sequence as defined in clause 7.3.5 of Recommendation ITU-T P.501 [12] as shown in figure 6.7.1-1. The competing speaker (upper signal in figure 6.7.1-1) is always inserted as the double talk sequence $s_{dt}(t)$ in send direction and is used for analysis. The test signals are synchronized in time at the acoustical interface.

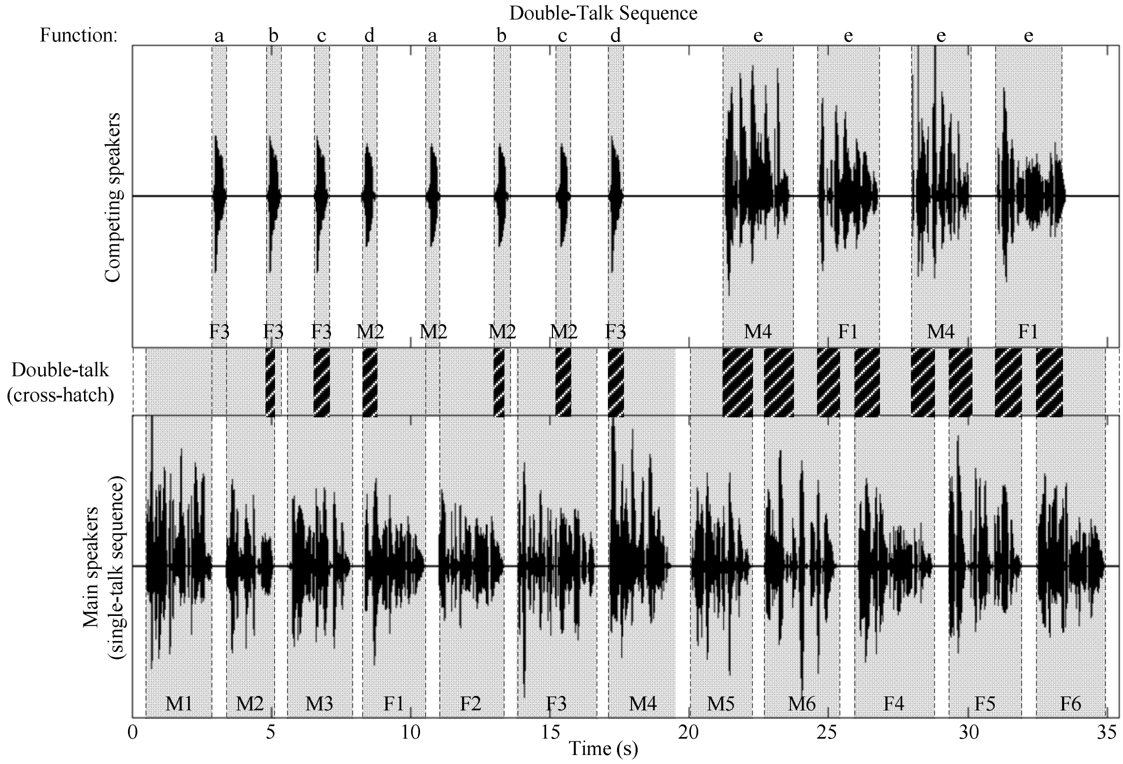


Figure 6.7.1-1: Double Talk Test Sequence with overlapping speech sequences in send and receive direction

Table 6.7.1-2: Void

The attenuation range during double talk is determined as described in Appendix III of Recommendation ITU-T P.502 [13]. The double talk performance is analysed for the sequences of words (first 20 s) and sentences (from 20 s to 35 s) produced by the competing speaker. The requirements have to be met for both sequences.

6.7.2 Attenuation Range in Receive Direction during Double Talk $A_{H,R,dt}$

Requirements

Based on the level variation in receive direction during double talk $A_{H,R,dt}$ the behaviour of the terminal can be classified according to table 6.7.2-1.

The terminal shall have full duplex capability (category 1).

Table 6.7.2-1: Double talk categories for receive direction

Category (according to Recommendation ITU-T P.340 [10])	1	2a	2b	2c	3
	Full Duplex Capability	*Partial Duplex Capability*			*No Duplex Capability*
***$A_{H,R,dt}$* [dB]**	*≤ 3*	*≤ 5*	*≤ 8*	*≤ 10*	*> 10*

The requirement applies with nominal setting of the volume control for nominal signal levels in send and receive directions as well as for the level combinations +6 dB (re. nominal level) in send/-6 dB (re. nominal level) in receive and +6 dB (re. nominal level) in receive/-6 dB (re. nominal level) in send. The requirement also applies for maximum setting of the volume control with nominal signal levels in send and receive directions.

In general, table 6.7.2-1 provides a quality classification of terminals regarding double talk performance. However, this does not mean that a terminal which is category 1 based on the double talk performance is of high quality concerning the overall quality as well.

Measurement method

The test setup is described in clause 5.2.

The long conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [12] shall be used for conditioning the terminal, with the female speaker in the receive direction. The test signal to determine the attenuation range during double talk is shown in figure 6.7.1-1. The competing speaker (upper signal in figure 6.7.1-1) is always inserted as the double talk sequence in receive direction and is used for analysis. The test signals are synchronized in time at the acoustical interface.

Table 6.7.2-2: Void

The attenuation range during double talk is determined as described in Appendix III of Recommendation ITU-T P.502 [13]. The double talk performance is analysed for the sequences of words (first 20 s) and sentences (from 20 s to 35 s) produced by the competing speaker. The requirements have to be met for both sequences.

6.7.3 Detection of echo components during double Talk

Requirement

"Echo Loss" (EL) is the echo suppression provided by the terminal measured at the electrical reference point.

Table 6.7.3-1: Echo loss during double talk categories

Category (according to Recommendation ITU-T P.340 [10])	1	2a	2b	2c	3
	Full Duplex Capability	*Partial Duplex Capability*			*No Duplex Capability*
Echo Loss [dB]	*≥ 27*	*≥ 23*	*≥ 17*	*≥ 11*	*< 11*

NOTE: The echo attenuation during double talk is based on the parameter Talker Echo Loudness Rating ($TELR_{dt}$). It is assumed that the terminal at the opposite end of the connection provides nominal Loudness Rating (SLR + RLR = 10 dB).

The terminal shall have full duplex capability (category 1).

In general table 6.7.3-1 provides a quality classification of terminals regarding double talk performance. However, this does not mean that a terminal which is category 1 or 2 based on the double talk performance is of high quality concerning the overall quality as well.

The requirement applies with nominal setting of the volume control for nominal signal levels in send and receive directions as well as for the level combinations +6 dB (re. nominal level) in send/-6 dB (re. nominal level) in receive and +6 dB (re. nominal level) in receive/-6 dB (re. nominal level) in send. The requirement also applies for maximum setting of the volume control with nominal signal levels in send and receive directions.

The measurement shall be repeated for desktop type terminals with variable echo path.

Measurement method

The terminal is set up as described in clause 5.2.

The double talk signal consists of a sequence of orthogonal signals which are realized by voice-like modulated sine waves spectrally shaped similar to speech. A detailed description can be found in Recommendation ITU-T P.501 [12]. The signal settings used are given in table 6.7.3-2.

The settings for the signals are as follows.

Table 6.7.3-2: Parameters of the two Test Signals for Double Talk Measurement based on AM-FM modulated sine waves

Send direction		Receive direction	
$f_0^{(1)}$ (Hz)	$\pm\Delta f^{(1)}$ (Hz)	$f_0^{(2)}$ (Hz)	$\pm\Delta f^{(2)}$ (Hz)
125	±2,5	180	±2,5
250	±5	270	±5
500	±10	540	±10
750	±15	810	±15
1 000	±20	1 080	±20
1 250	±25	1 350	±25
1 500	±30	1 620	±30
1 750	±35	1 890	±35
2 000	±40	2 160	±35
2 250	±40	2 400	±35
2 500	±40	2 650	±35
2 750	±40	2 900	±35
3 000	±40	3 150	±35
3 250	±40	3 400	±35
3 500	±40	3 650	±35
3 750	±40	3 900	±35
4 000	±40	4 150	±35
4 250	±40	4 400	±35
4 500	±40	4 650	±35
4 750	±40	4 900	±35
5 000	±40	5 150	±35
5 250	±40	5 400	±35
5 500	±40	5 650	±35
5 750	±40	5 900	±35
6 000	±40	6 150	±35
6 250	±40	6 400	±35
6 500	±40	6 650	±35
6 750	±40	6 900	±35
7 000	±40		
NOTE Parameters of the Shaping Filter: f ≥ 250 Hz: Low Pass Filter, 5 dB/oct.			

The signal generation is according to Recommendation ITU-T P.501 [12].

The signals are fed simultaneously in send and receive direction. The level in send direction is -4,7 dBPa at the MRP (nominal level), the level in receive direction is -16 dBm0 at the electrical reference point (nominal level).

The test signal is measured at the electrical reference point (send direction). The measured signal consists of the double talk signal which was fed in by the artificial mouth and the echo signal. The echo signal is filtered by comb filter using mid-frequencies and bandwidth according to the signal components of the signal in receive direction (see Recommendation ITU-T P.501 [12]). The filter will suppress frequency components of the double talk signal.

In each frequency band which is used in receive direction the echo attenuation can be measured separately. The requirement for category 1 is fulfilled if in any frequency band the echo signal is either below the signal noise or below the required limit. If echo components are detectable, the classification is based on table 6.7.3-1. The echo attenuation is to be achieved for **each individual frequency band** according to the different categories.

6.7.4 Minimum activation level and sensitivity of double talk detection

For further study.

6.8 Switching parameters

6.8.0 Overview

Additional requirements may be needed in order to further investigate the effect of NLP implementations on the users' perception of speech quality.

6.8.1 Activation in Send Direction

The activation in send direction is mainly determined by the built-up time $T_{r,S,min}$ and the minimum activation level ($L_{S,min}$). The minimum activation level is the level required to remove the inserted attenuation in send direction during idle mode. The built-up time is determined for the test signal burst which is applied with the minimum activation level.

The activation level described in the following is always referred to the test signal level at the Mouth Reference Point (MRP).

Requirements

The minimum activation level $L_{s,min}$ shall be ≤ -20 dBPa.

The built-up time $T_{r,S,min}$ (measured with minimum activation level) shall be ≤ 15 ms.

Measurement method

The terminal is set up as described in clause 5.2.

The test signal is the "short words for activation" sequence described in clause 7.3.4 of Recommendation ITU-T P.501 [12] with increasing level for each single word.

The settings of the test signal are described in table 6.8.1-1.

Table 6.8.1-1: Settings for the signal

	CSS Duration/ Pause Duration	Level of the first CS Signal (active Signal Part at the MRP)	Level Difference between two Periods of the Test Signal
Single word to Determine Switching Characteristic in Send Direction	~600 ms / ~400 ms	-24 dBPa (see notes)	1 dB
NOTE 1: The level of the active signal part corresponds to an average level of -24,7 dBPa at the MRP for the CSS according to Recommendation ITU-T P.501 [12]. NOTE 2: The signal level is determined for each utterance individually according to Recommendation ITU-T P.56 [7].			

It is assumed that the pause length of about 400 ms is longer than the hang-over time so that the test object is back to idle mode after each single word.

The level of the transmitted signal is measured at the electrical reference point. The test signal is filtered by the transfer function of the test object. The measured signal level is referred to the filtered test signal level and displayed vs. time. The levels are calculated from the time domain using an integration time of 5 ms.

The minimum activation level is determined from the single word which indicates the first activation of the test object. The time between the beginning of the single word and the complete activation of the test object is measured.

6.8.2 Minimum activation level and sensitivity in Receive direction

For further study.

6.8.3 Automatic level control

For further study.

6.8.4 Silence Suppression and Comfort Noise Generation

For further study.

6.9 Background noise performance

6.9.1 Performance in send direction in the presence of background noise

Requirement

The level of comfort noise, if implemented, shall be within in a range of +2 dB and -5 dB compared to the original (transmitted) background noise. The noise level is calculated with A-weighting.

NOTE 1: It is advisable that the comfort noise matches the original signal as good as possible (from a perception point of view).

NOTE 2: Input for further specification necessary (e.g. on temporal matching).

The spectral difference between comfort noise and original (transmitted) background noise shall be within the mask given through straight lines between the breaking points on a logarithmic (frequency) - linear (dB sensitivity) scale as given in table 6.9.1-1.

Table 6.9.1-1: Requirements for Spectral Adjustment of Comfort Noise (Mask)

Frequency	Upper Limit	Lower Limit
200 Hz	12 dB	-12 dB
800 Hz	12 dB	-12 dB
800 Hz	10 dB	-10 dB
2 000 Hz	10 dB	-10 dB
2 000 Hz	6 dB	-6 dB
4 000 Hz	6 dB	-6 dB
8 000 Hz	6 dB	-6 dB
NOTE: All sensitivity values are expressed in dB on an arbitrary scale.		

Measurement method

The terminal is set up as described in clause 5.2.

The background noise simulation as described in clause 5.5.3 is used.

First the background noise transmitted in send is recorded at the POI for a period of at least 20 s.

In a second step a test signal is applied in receive direction consisting of an initial pause of 10 s and a periodical repetition of the female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [12] in receive direction (duration 10 s) with nominal level to enable comfort noise injection simultaneously with the background noise. For the measurement the background noise sequence has to be started at the same point as it was started in the previous measurement.

The transmitted signal is recorded in send direction at the POI.

The power density spectra measured in send direction without far end speech simulation averaged between 10 s and 20 s is referred to the power density spectrum measured in send direction determined during the period with far end speech simulation in receive direction averaged between 10 s and 20 s. Level and spectral differences between both power density spectra are analysed and compared to the requirements.

6.9.2 Speech Quality in the Presence of Background Noise

Requirement

Speech Quality for wideband systems shall be tested based on ETSI TS 103 106 [20]. The test method leads to three MOS-LQO quality numbers:

- N-MOS-LQOw: Transmission quality of the background noise.
- S-MOS-LQOw: Transmission quality of the speech.
- G-MOS-LQOw: Overall transmission quality.

For the background noises defined in clause 5.5.3 the following requirements apply:

- N-MOS-LQOw ≥ 3,0.
- S-MOS-LQOw ≥ 3,0.
- G-MOS-LQOw ≥ 3,0.

NOTE: It is recommended to test the terminal performance with other types of background noises if the terminal is likely to be exposed to other noises than specified in clause 5.5.3.

Measurement method

The background noise simulation as described in clause 5.5.3 is used. The hands-free terminal is set-up as described in clause 5.2.

The background noise shall be applied for at least 5 s in order to adapt noise reduction algorithms in advance the test.

The near end speech signal consists of 16 sentences of speech (2 male and 2 female talkers, 4 sentences each). An appropriate measurement sequence in American English is provided in Annex C of ETSI TS 103 106 [20]. The test signal level is +1,3 dBPa at the MRP.

Three signals are required for the tests:

1) The clean speech signal is used as the undisturbed reference.
2) The speech plus undisturbed background noise signal is recorded at the terminal's microphone position using an omni directional measurement microphone with a linear frequency response between 50 Hz and 12 kHz.
3) The send signal is recorded at the electrical reference point.

N-MOS-LQOw, S-MOS LQOw and G-MOS LQOw are calculated as described in ETSI TS 103 106 [20].

6.9.3 Quality of Background Noise Transmission (with Far End Speech)

Requirement

The test is carried out applying a speech signal in receive direction and comparing the noise level transmitted in the send direction under reference conditions with no far end speech, to the noise level transmitted in the send direction under test conditions including far end speech. During and after the end of the speech signal the signal level in send direction shall not vary more than 10 dB (during transition to transmission of background noise without far end speech). The measurement is conducted for all types of background noise as defined in clause 5.5.3.

NOTE: The intention of this measurement is to detect impairments (modulations, switching and others) influencing the background noise transmitted from the terminal under test when a signal from the distant end (Receive side of the terminal under test) is present. Under these test conditions no modulation of the transmitted signal should occur. Modulation, switching or other type of impairments might be caused by an improper behaviour of a non linear processor working in conjunction with the echo canceller and erroneously switching or modulating the transmitted background noise.

Measurement method

The terminal is set up is described in clause 5.2.

The background noises are generated as described in clause 5.5.3.

First the reference measurement is conducted without inserting the signal at the far end. At least 10 s of noise are analysed. The background signal level versus time is calculated using a time constant of 35 ms. This is the reference signal.

In a second step the same measurement is conducted, but with inserting the speech signal at the far end. The exactly identical background noise signal is applied. The background noise signal should start at the same point in time as was used for the reference measurement without the far end signal. The background noise shall be applied for at least 5 s in order to allow adaptation of the noise reduction algorithms. After at least 5 s a series of the female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [12] is applied in receive direction with duration of at least 10 s. The test signal level in the Receive direction is -16 dBm0 at the electrical reference point.

For both, reference and test conditions the send signal is recorded at the electrical reference point and the test signal level versus time is calculated using a time constant of 35 ms.

The level variation in send direction is determined during the time interval when the speech signal is applied and after it stops. The level difference is determined from the difference of the recorded signal levels versus time between reference signal and the signal measured with far end signal.

6.10 Quality of echo cancellation

6.10.1 Temporal echo effects

Requirements

This test is intended to verify that the system will maintain sufficient echo attenuation during single talk. The measured echo attenuation during single talk shall not decrease by more than 6 dB from the maximum echo attenuation measured.

Measurement method

The test arrangement is according to clause 5.2.

The test signal consists of periodically repeated Composite Source Signal according to Recommendation ITU-T P.501 [12] with an average level of -5 dBm0 as well as an average level of -25 dBm0. The echo signal is analysed during a period of at least 2,8 s which represents 8 periods of the CS signal. The integration time for the level analysis shall be 35 ms, the analysis is referred to the level analysis of the reference signal.

The measurement result is displayed as attenuation vs. time. The exact synchronization between input and output signal has to be guaranteed.

The difference between the maximum attenuation and the minimum attenuation is measured.

NOTE 1: In addition, tests with more speech like signals should be made, e.g. Recommendation ITU-T P.501 [12] to see time variant behaviour of EC. However, for such tests the simple broadband attenuation based test principle as described above cannot be applied due to the time varying spectral content of the speech like signals.

NOTE 2: The analysis is conducted only during the active signal part, the pauses between the Composite Source Signals are not analysed. The analysis time is reduced (35 ms) by the integration time of the level analysis.

NOTE 3: Care should be taken not to confuse noise or comfort noise with residual echo. In cases of doubt the measured echo signal should be compared to the residual noise signal measured under the same conditions without inserting the receive signal. If the level vs. time analysis leads to the identical result it can be assumed that no echo but just comfort noise is present.

6.10.2 Spectral Echo Attenuation

Requirement

The echo attenuation vs. frequency shall be below the tolerance mask given in table 6.10.2-1.

Table 6.10.2-1: Spectral echo loss limits

Frequency	Limit
100 Hz	-41 dB
1 300 Hz	-41 dB
3 450 Hz	-46 dB
5 200 Hz	-46 dB
7 500 Hz	-37 dB
8 000 Hz	-37 dB
NOTE: The limit at intermediate frequencies lies on a straight line drawn between the given values on a log (frequency) - linear (dB) scale.	

During the measurement it should be ensured that the measured signal is really the echo signal and not the Comfort Noise which possibly may be inserted in send direction in order to mask the echo signal.

Measurement method

The test setup is described in clause 5.2.

Before the actual measurement a training sequence is fed in consisting of the compressed real speech signal described in clause 7.3.3 of Recommendation ITU-T P.501 [12]. The level of the training sequence shall be -16 dBm0.

The test signal is the compressed real speech signal described in clause 7.3.3 of Recommendation ITU-T P.501 [12]. The average test signal level shall be -16 dBm0, averaged over the complete test signal. The power density spectrum of the measured echo signal is referred to the power density spectrum of the original test signal. The analysis is conducted using FFT analysis with 8 k points (48 kHz sampling rate, Hanning window).

The spectral echo attenuation is analysed in the frequency domain in dB.

6.10.3 Occurrence of Artefacts

For further study.

6.10.4 Variable echo path

Requirement

This test is intended to verify that the system will maintain sufficient echo attenuation during single talk with dynamic changing echo paths. The measured echo level over time during single talk shall not be more than 10 dB above the minimum noise level during the measurement.

Measurement method

The test setup is described in clause 5.2.5.

NOTE: Care should be taken to not generate noise during the movement of the notebook lid. Because of this, this measurement is not applicable for a softphone without external microphone.

As test signal the compressed real speech signal described in clause 7.3.3 of Recommendation ITU-T P.501 [12] is used. The signal level shall be -10 dBm0. The terminal volume control is set to nominal RLR. The first 4 sentences of the test signal are used to allow full convergence of the echo canceller. The next 4 sentences (from 10,75 s to 22,5 s) are used for the analysis. The echo signal level is analysed over time The echo signal level is analysed for 11,75 s, using a time constant of 35 ms.

The measurement result is displayed as echo level versus time.

6.11 Send and receive delay - round-trip delay

Requirement

Send and receive delays are tested separately but 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 round-trip delay T_{rtd} (sum of send and receive delay) shall be less than 100 ms (category B in Recommendation ITU-T P.1010 [23]).

From the user's perspective, a value less than 50 ms (category A in Recommendation ITU-T P.1010 [23]) is preferred.

NOTE: The delay should in general be minimized. This can be accomplished by e.g. designing the speech decoder output, and the signal processing in a way, that sample-based processing and frame-based processing interoperate by using common buffers at their interfaces.

Measurement method

The terminal is set up as described in clause 5.2.

- **Send direction**

The delay in send direction is measured from the MRP (Mouth Reference Point) to POI (reference speech codec of the system simulator, output). The delay measured in send direction is:

$$T_S + t_{system} \tag{6}$$

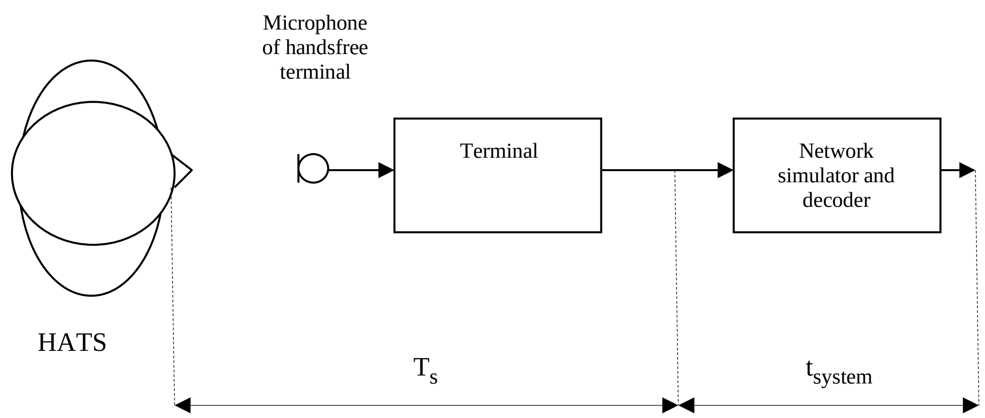


Figure 6.11-1: Different blocks contributing to the delay in send direction

The system delay t_{system} is depending on the transmission method used and the network simulator. The delay t_{system} shall be known and considered in the calculation of the delay, which is determined as follows:

1) For the measurements a Composite Source Signal (CSS) according to Recommendation ITU-T P.501 [12] 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 test signal level is adjusted to -28,7 dBPa at the HATS-HFRP (see Recommendation ITU-T P.581 [14]).

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 delays which are caused by the test equipment.

3) The delay is measured in ms and the maximum of the envelope of the cross-correlation function is used for the determination.

- **Receive direction**

The delay in receive direction is measured from POI (input of the reference speech codec of the system simulators) to the Drum Reference Point (DRP). The delay measured in receive direction is:

$$T_r + t_{System} \tag{7}$$

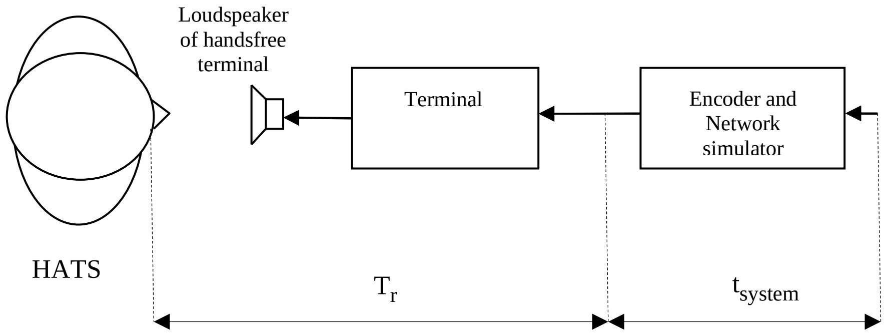


Figure 6.11-2: Different blocks contributing to the delay in receive direction

The system delay t_{System} is depending on the transmission system and on the network simulator used. The delay t_{System} shall be known and considered in the calculation of the delay, which is determined as follows:

1) For the measurements a Composite Source Signal (CSS) according to Recommendation ITU-T P.501 [12] 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).

2) Artificial head is free-field adjusted according to Recommendation ITU-T P.581 [14]. The equalized output signal of the right ear is used for the measurement.

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 delays which are caused by the test equipment.

4) The delay is measured in ms and the maximum of the envelope of the cross-correlation function is used for the determination.

6.12 Void

7 Codec dependent requirements and associated Measurement Methodologies

7.1 Speech Codecs

The present document is intended to be applicable for different speech codecs implemented in access networks.

Table 7.1-1 defines a list of speech codecs implemented (non exhaustive).

Table 7.1-1: List of speech codecs

System	Codec
UMTS (WCDMA)	AMR-WB (Recommendation ITU-T G.722.2) @12,65 kbit/s [17]
VoLTE	AMR-WB (Recommendation ITU-T G.722.2) @12,65 kbit/s [17] EVS-WB @ 13,2 kbit/s [22]

The objective is to minimize the impact of transcodings on the quality. Care should also be taken to avoid as far as possible to cascade different speech processing.

7.2 Void

7.3 Objective listening Quality in send and receive direction

7.3.0 Overview

The measurements of this clause are only applicable if the hands-free terminal supports at least one of the codecs listed in table 7.1-1.

7.3.1 Objective listening speech quality MOS-LQO in send direction

The listening speech quality tests are conducted without any packet impairments (clean network conditions).

Requirements

The requirements for the listening speech quality according to table 7.3.1-1 apply.

Table 7.3.1-1: Requirements for speech quality in send direction

Speech codec	MOS-LQOS (P.863)	MOS-LQOM (TOSQA 2001)
AMR-WB @12,65 kbit/s [17]	(ffs)	(ffs)
EVS-WB @ 13,2 kbit/s [22]	(ffs)	(ffs)

NOTE 1: Insufficient experience is available so far with Recommendation ITU-T P.863 [24] and TOSQA 2001 (ETSI EG 201 377-1 [i.2]) for measuring hands-free terminals. Therefore, the numbers for MOS-LQOS and MOS-LQOM are for further study.

Measurement method

The terminal is set up as described in clause 5.2.

For the assessment of objective listening speech quality, fullband mode of Recommendation ITU-T P.863 [24] shall be applied.

The test signal to be used for the measurements shall be 4 sentence pairs (male/female) fulfilling the requirements of Recommendation ITU-T P.863.1 [25]. The four British English sentence pairs are taken from Annex C of Recommendation ITU-T P.501 [12]. The test signal level is averaged over all sentence pairs (4 sentence pairs). The measurement is repeated for each pair of speech sentences. The overall result of the measurement is the averaged value of all four per-sample measurements.

NOTE 2: Recommendation ITU-T P.863 [24] in fullband mode provides results in a fullband context (MOS-LQOF).

NOTE 3: An alternative test method is TOSQA 2001 (ETSI EG 201 377-1 [i.2]). With TOSQA, terminals used in narrowband and wideband mode should be measured based on MOS-LQOM.

7.3.2 Objective listening quality MOS-LQO in receive direction

7.3.2.1 Jitter- and Error-Free Condition

The listening speech quality tests are conducted without any packet impairments (clean network conditions).

Requirements

The requirement for the listening speech quality and the delay under clean network conditions according to table 7.3.2.1-1 apply.

Table 7.3.2.1-1: Requirements for speech quality in receive direction

Speech coder	MOS-LQOS (P.863)	MOS-LQOM (TOSQA 2001)
AMR-WB @12,65 kbit/s [17]	(ffs)	(ffs)
EVS-WB @ 13,2 kbit/s [22]	(ffs)	(ffs)

NOTE 1: Not sufficient experience is available so far with Recommendation ITU-T P.863 [24] and TOSQA 2001 (ETSI EG 201 377-1 [i.2]) for measuring handsfree terminals. Therefore, the numbers for MOS-LQOS and MOS-LQON are for further study.

Measurement method

The terminal is set up as described in clause 5.2.

For the assessment of objective listening speech quality, fullband mode of Recommendation ITU-T P.863 [24] shall be applied.

The test signal to be used for the measurements shall be 4 sentence pairs (male/female) fulfilling the requirements of Recommendation ITU-T P.863.1 [25]. The four British English sentence pairs are taken from Recommendation ITU-T P.501 [12]. The test signal level shall be -16 dBm0, measured according to Recommendation ITU-T P.56 [7] at the digital reference point or the equivalent analogue point. The measurement is repeated for each pair of speech sentences. The overall result of the measurement is the averaged value of all 4 per-sample measurements.

NOTE 2: Recommendation ITU-T P.863 [24] in fullband mode provides results in a fullband context (MOS-LQOF).

NOTE 3: An alternative test method is TOSQA 2001 (ETSI EG 201 377-1 [i.2]). With TOSQA, terminals used in narrowband and wideband mode should be measured based on MOS-LQOM.

7.3.2.2 Packet Impairments

The listening speech quality tests are conducted with simulated packet impairments. In addition to the listening speech quality tests, the delay is measured. The tests of this clause are only applicable to terminals providing an IP-based network access.

Requirement

The degradation between the error- and jitter-free condition (equals network condition 1) and impairment conditions shall not exceed the delta-values provided in table 7.3.2.2-1.

Table 7.3.2.2-1: Requirements for speech codecs per network condition

Codec	Condition	MOS-LQON (P.863)	MOS-LQOM (TOSQA 2001)	Delay T_{rtd}
AMR-WB @12,65 kbit/s [17] EVS WB @ 13,2 kbit/s [22]	0	(ffs)	(ffs)	$T_{rtd,clean}$ + 5 ms
	1	(ffs)	(ffs)	$T_{rtd,clean}$ + 5 ms
	2	(ffs)	(ffs)	$T_{rtd,clean}$ + 5 ms
	3	(ffs)	(ffs)	$T_{rtd,clean}$ + 5 ms
	4	(ffs)	(ffs)	$T_{rtd,clean}$ + 5 ms
	5	(ffs)	(ffs)	$T_{rtd,clean}$ + 50 ms
	6	(ffs)	(ffs)	$T_{rtd,clean}$ + 50 ms
	7	(ffs)	(ffs)	$T_{rtd,clean}$ + 50 ms

NOTE: The delay requirements for conditions with network impairments are based on the measured roundtrip delay of the terminal in the absence of network impairments $T_{rtd,clean}$ (see clause 6.11). A small additional tolerance takes into account the variable behaviour of the delay.

Measurement method

For the performance tests with network impairments the settings according to table 7.3.2.2-2 are used. The test setup is the same as in clause 7.3.2.1.

Table 7.3.2.2-2: Network conditions for electrical-acoustical measurements

Condition	Packet Loss [%]	Delay Variation	Reordering
0	0	No	No
1 (see note 2) (VAD)	0	No	No
2	1	No	No
3	3	No	No
4	0	Yes (see note 1)	No
5	1	Yes (see note 1)	No
6	0	Yes (see note 1)	Yes
7	1	Yes (see note 1)	Yes

NOTE 1: Delay variation produced according to Annex D of ETSI ES 202 737 [i.6].
NOTE 2: VAD on, all other conditions tested with VAD off.
NOTE 3: For some network emulation tools, it is necessary to introduce a constant delay to offer the possibility to generate a delay variation distribution. This delay has to be subtracted from the measured delay before interpreting the results.
NOTE 4: When running tests with the conditions in row, it may be necessary to make one call per condition to avoid the influence of the order of the conditions to the results.

Annex A (informative): Bibliography

- Recommendation ITU-T G.107: "The E-model, a computational model for use in transmission planning".
- Recommendation ITU-T G.108: "Application of the E-model: A planning guide".
- Recommendation ITU-T G.109: "Definition of categories of speech transmission quality".
- Recommendation ITU-T P.862: "Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs".
- Recommendation ITU-T P.1100: "Narrowband hands-free communication in motor vehicles".
- ETSI TS 126 131: "Universal Mobile Telecommunications System (UMTS); LTE; 5G; Terminal acoustic characteristics for telephony; Requirements (3GPP TS 26.131)".
- ETSI ES 202 738: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for narrowband VoIP terminals (handset and headset) from a QoS perspective as perceived by the user".
- 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".
- ETSI EG 202 396-3: "Speech and multimedia Transmission Quality (STQ); Speech Quality performance in the presence of background noise; Part 3: Background noise transmission - Objective test methods".
- Recommendation ITU-T G.722: "7 kHz audio-coding within 64 kbit/s".
- Recommendation ITU-T G.729.1: "G.729-based embedded variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable with G.729".
- Recommendation ITU-T G.711.1: "Wideband embedded extension for ITU-T G.711 pulse code modulation".

History

Document history		
V1.1.1	November 2009	Publication
V1.1.2	September 2010	Publication
V1.2.1	July 2017	Publication
V1.3.1	October 2018	Publication
V1.4.1	October 2021	Publication