ETSI TS 103 738 V1.3.1 (2018-10)

Speech and multimedia Transmission Quality (STQ); Transmission requirements for narrowband wireless terminals (handsfree) from a QoS perspective as perceived by the user

Reference

RTS/STQ-272-2

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

Important notice

The present document can be downloaded from:
http://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (PDF) version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

If you find errors in the present document, please send your comment to one of the following services:
https://portal.etsi.org/People/CommiteeSupportStaff.aspx

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables 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 (https://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.

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.

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 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 mostly realistic test conditions and meaningful results.

1 Scope

The present document provides speech transmission performance requirements for wireless terminals; it addresses all types of wireless terminals, including softphones. This part addresses handsfree function of narrowband wireless terminals.

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 enable good quality end-to-end speech performance as perceived by the user, whatever be the radio link (terminals may implement different radio links with the access network).

When an additional radio link between the terminal and external electroacoustical devices is used (e.g. Bluetooth® link), the standard will address the overall quality.

In the present document, objective measurement methodologies and requirements for wireless speech terminals are given.

In addition to basic testing procedures, the present document describes advanced testing procedures taking into account further quality parameters as perceived by the user.

The requirements available in the present document will ensure a high compatibility across access networks with all types of terminals.

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

For all the functions, the standard will consider the limitations in audio performance due to different form factors (e.g. size, shape).

Terminals which are not intended to be connected to public networks are outside the scope of the present document.

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] 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)".

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

[3] Void.

[4] Recommendation ITU-T G.711: "Pulse code modulation (PCM) of voice frequencies".

[5] Recommendation ITU-T G.726: "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)".

[6] Recommendation ITU-T G.729: "Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)".

[7] 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".

[8] Void.

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

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

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

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

[13] Recommendation ITU-T P.342: "Transmission characteristics for narrow-band digital loudspeaking and hands-free telephony terminals".

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

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

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

[17] Recommendation ITU-T O.41: "Psophometer for use on telephone-type circuits".

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

[19] ETSI TS 146 010: "Digital cellular telecommunications system (Phase 2+); Full-rate speech; Transcoding (3GPP TS 46.010 Release 9)".

[20] ETSI TS 146 060: "Digital cellular telecommunications system (Phase 2+); Enhanced Full Rate (EFR) speech transcoding (3GPP TS 46.060 Release 9)".

[21] ETSI TS 103 106 (03-2013) (V1.2.1): "Speech and multimedia Transmission Quality (STQ); Speech quality performance in the presence of background noise: Background noise transmission for mobile terminals-objective test methods".

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

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

[24] ETSI TS 126 441 (V12.0.0): "Universal Mobile Telecommunications System (UMTS); LTE; EVS Codec General Overview (3GPP TS 26.441 version 12.0.0 Release 12)".

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

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] Recommendation ITU-T P.1100: "Narrowband hands-free communication in motor vehicles".

[i.3] IEC 61672 (Edition 1.0): "Electroacoustics - Sound level meters".

[i.4] Void.

[i.5] Void.

[i.6] 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.7] Recommendation ITU-T P.1010: "Fundamental voice transmission objectives for VoIP terminals and gateways".

3 Definitions of terms 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 flat for frontal sound incidence in anechoic conditions

group audio terminal: handsfree telephony terminal primarily designed for use by several users which will not be equipped with a handset.

handsfree 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 [10]: "n" is one of the points numbered from 11 to 17 and defined in table 6a of Recommendation ITU-T P.58 [10] (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 an external loudspeaker associated with an 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: volume control setting which yields the RLR value 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 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 Modulation Frequency Modulation
AMR	Adaptive Multi-Rate codec
CDMA	Code Division Multiple Access
CS	Composite Source
CSS	Composite Source Signal
DECT	Digital Enhanced Cordless Telecommunications
DFT	Discrete Fourrier Transform
DRP	Drum Reference Point
DUT	Device Under Test
EC	Echo Cancellation
EFR	Enhanced Full Rate
EL	Echo loss
EVS-NB	Enhanced Voice Services - Narrowband
FFT	Fast Fourier Transform
G-MOS-LQOn	Overall transmission quality narrowband
GSM	Global System for Mobile communication
HATS	Head And Torso Simulator
HF	Handsfree
HFRP	Handsfree Reference Point
HFT	Handsfree Telephone
IEC	International Electrotechnical Commission
ITU-T	International Telecommunication Union - Telecommunication standardization sector
LE	Earcap Leakage/Coupling Loss
LQO	Listening Quality Objective
LTE	Long Term Evolution (3GPP)
MOS	Mean Opinion Score
MRP	Mouth Reference Point
NB	Narrow Band
NLP	Non Linear Processing
N-MOS-LQOn	Transmission quality of the background noise narrowband
PDA	Personal Data Analyser
PLC	Packet Loss Concealment
PMRP	Sound Pressure at MRP
PN	Pseudo random Noise
POI	Point Of Interconnect
QoS	Quality of Service
RF	Radio Frequency
RLR max	Receive Loudness Rating corresponding to the maximum setting of the volume control
RLR min	Receive Loudness Rating corresponding to the minimum setting of the volume control
RLR	Receive Loudness Rating
RMS	Root Mean Square
SLR	Send Loudness Rating
S-MOS-LQOn	Transmission quality of the speech narrowband
TCL_W	Terminal Coupling Loss (weighted)
$TELR_{dt}$	Talker Echo Loudness Rating
TOSQA	Telecommunication Objective Speech Quality Assessment
UE	User Equipment
UMTS	Universal Mobil Telecommunications System
VAD	Voice Activity Detector
VoLTE	Voice over LTE
WIFI	Wireless Fidelity
WIMAX™	Worldwide Interoperability for Microwave ACCess

4 Configurations and interfaces

4.0 Introduction

The present document is intended to be applicable for different wireless access networks and for additional radio links.

4.1 Access networks

The present document applies to any wireless terminal whatever the access network, e.g. GSM, UMTS, VoLTE, DECT, Bluetooth®, WIFI, WIMAX™, CDMA, etc.

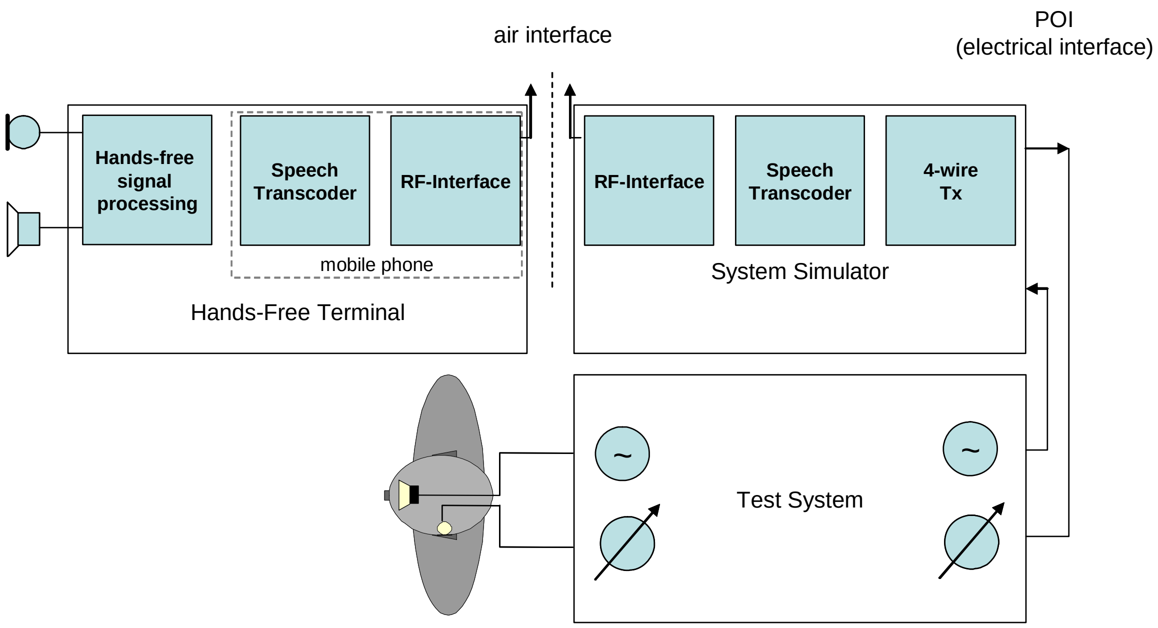
4.2 Additional (radio) links between the terminal and external electroacoustical devices

The present document also applies when an additional radio link exists between the wireless terminal and external electro acoustic devices, e.g. Bluetooth®.

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 figure 5.1-1 is not the additional radio link.

Figure 5.1-1: Set-up interface

5.2 Set-up for terminals

5.2.0 General

For electroacoustical testing, HATS as described in Recommendation ITU-T P.58 [10] shall be 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 or using the direct signal processing approach or acoustically using ITU-T specified devices.

When a coder 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-NB (ETSI TS 126 171 [1]): 12,2 kbit/s
- Recommendation ITU-T G.729.1 [7]: 32 kbit/s

5.2.1 Handheld terminal

HATS measurement equipment shall be configured to the handheld hands-free UE according to figure 5.2-1. 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.

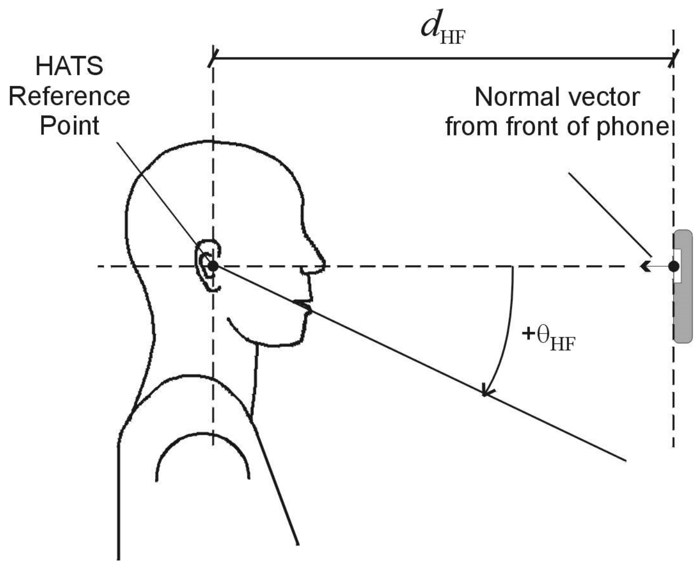


Figure 5.2-1: Configuration of handheld hands-free UE relative to the HATS

5.2.2 Vehicle mounted hands-free

Test arrangement, test methods and performance requirements are according to Recommendation ITU-T P.1100 [i.2].

Figure 5.2.2-1: Void.

5.2.3 Desktop hands-free terminal

Definition of hands-free terminals and setup for desktop hands-free terminals are based on in Recommendation ITU-T P.581 [16].

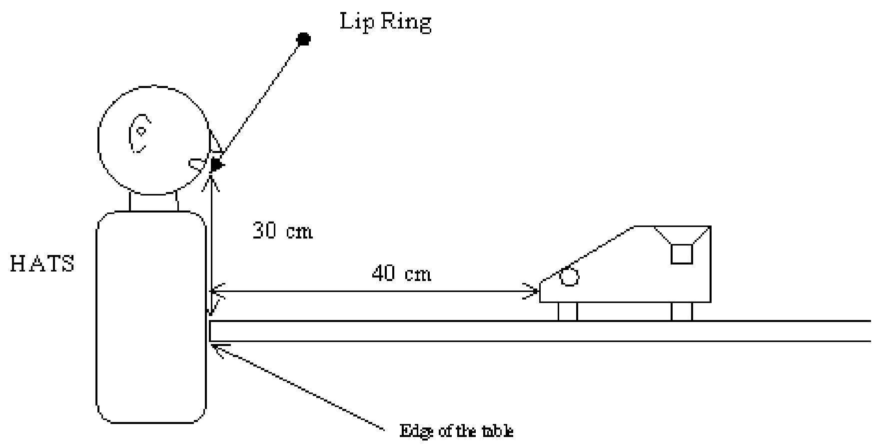


Figure 5.2.3-1: Position for test of desktop hands-free terminal side view

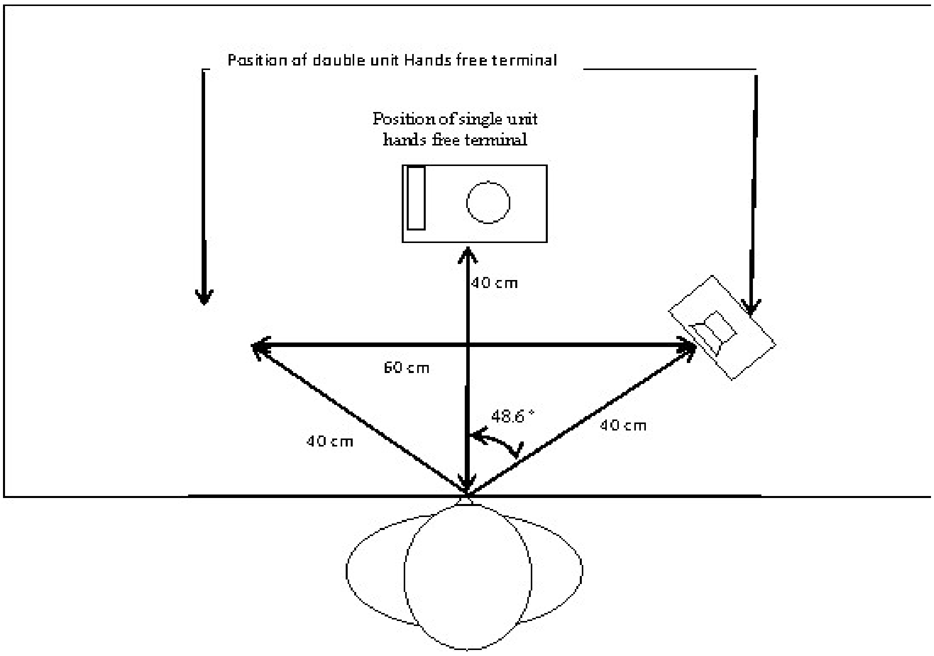


Figure 5.2.3-2: Position for test of desktop hands-free terminal top sight

5.2.4 Additional test setup for hands-free function with softphone

5.2.4.0 General

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

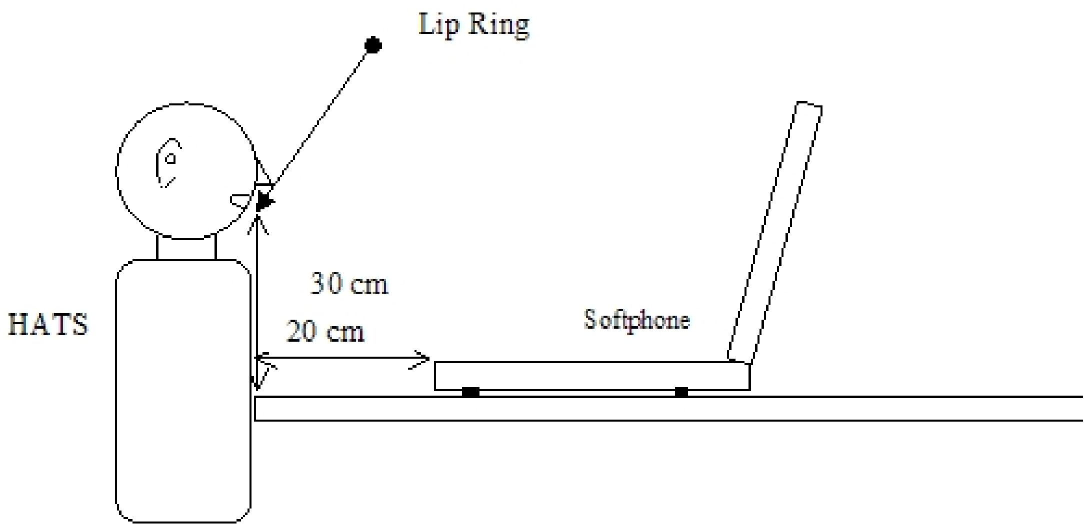


Figure 5.2.4.1-1: Configuration of softphone relative to the HATS side view

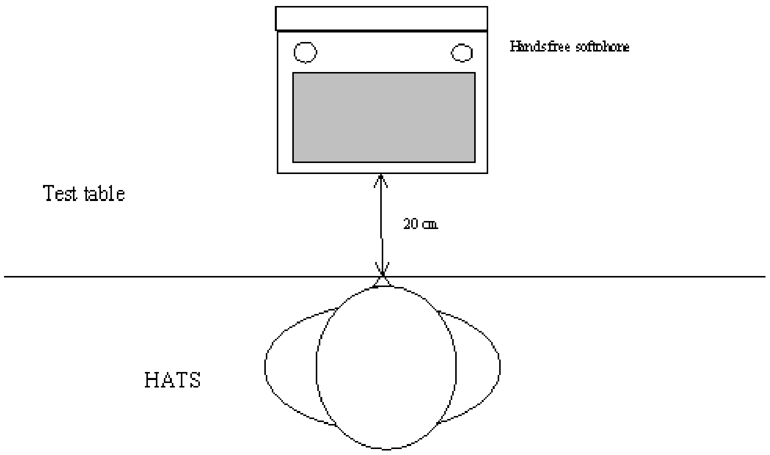


Figure 5.2.4.1-2: Configuration of softphone relative to the HATS-top sight

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.

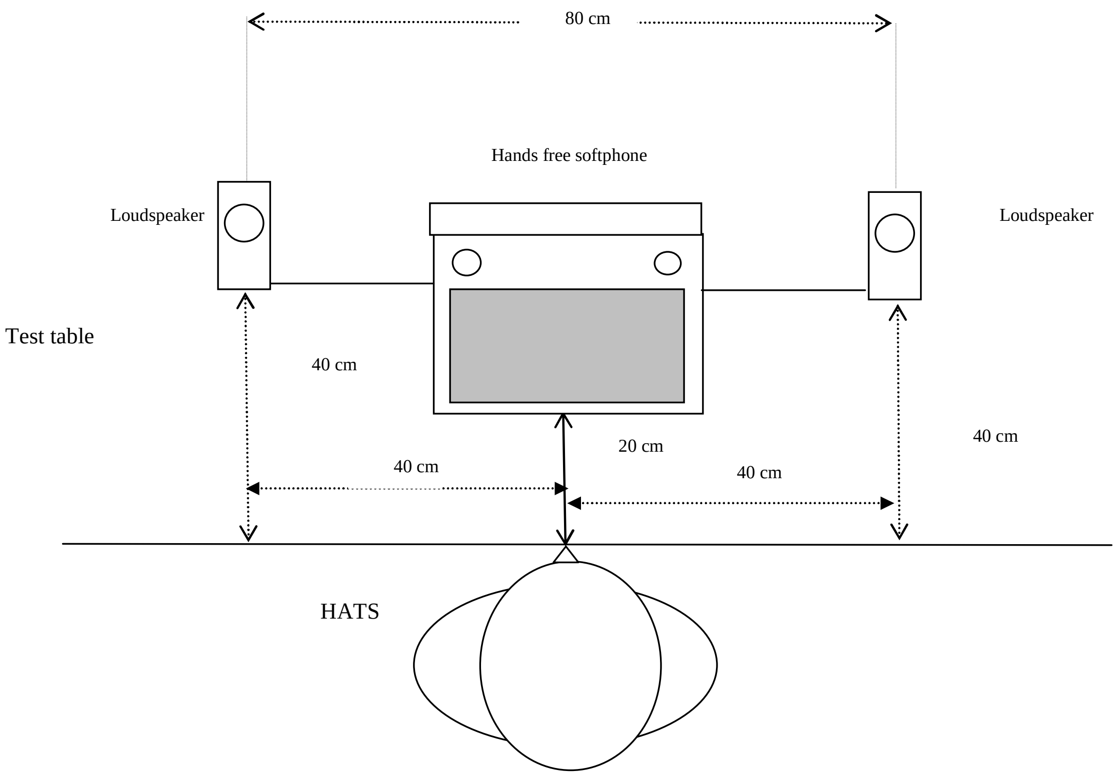


Figure 5.2.4.2-2: Configuration of softphone using external speakers relative to the HATS-top sight

When external microphone and speakers are used, system will be positioned as in figure 5.2.4.2-3.

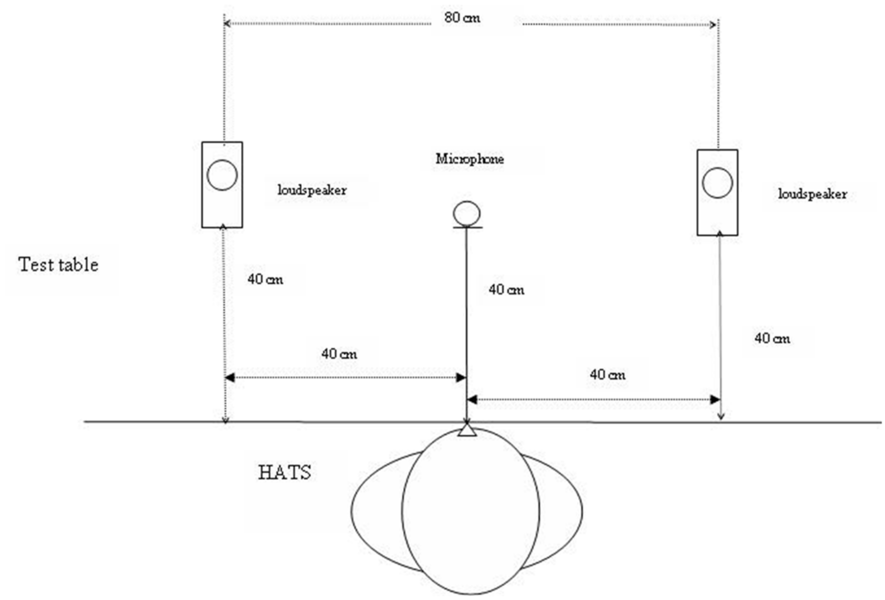


Figure 5.2.4.2-3: Configuration of softphone using external speakers and microphone relative to the HATS-top sight

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.

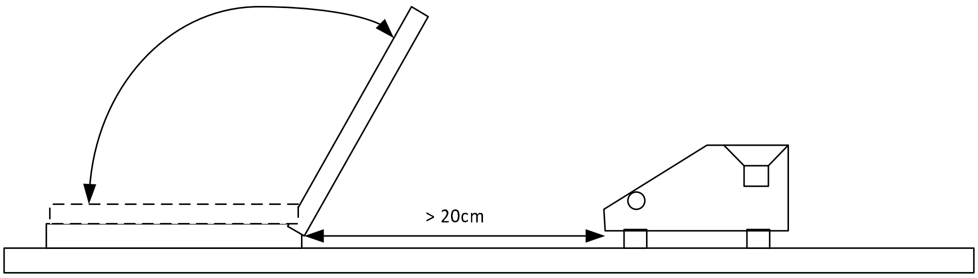


Figure 5.2.5-1: Positioning of DUT

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:

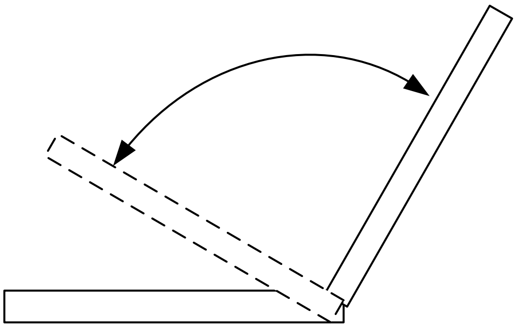


Figure 5.2.5-2: Positioning of DUT

Test setup for other hands-free 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.342 [13], 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

Due to the coding of the speech signals, care should be taken when using sinusoidal test signals for some wireless terminals/networks (e.g. GSM/3G), appropriate test signals (general description) are defined in Recommendation ITU-T P.501 [14].

Unless stated otherwise 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 [14]. More information can be found in the test procedures described below.

For testing the narrowband telephony service provided by a terminal the test signal used shall be band limited between 100 Hz and 4 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.

Unless specified otherwise, the test signal level shall be -4,7 dBPa at the MRP, calibrated at the various positions as defined in clause 5.5.1.

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

5.5 Calibration and test signal level

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 [16].

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:

$$\mathrm{S_{mJ}} = 20\log\mathrm{V_s} - 20\log\mathrm{P_{MRP}} + \mathrm{Corr} - Dcorr$$

where:

- **Vs** is the measured voltage across the appropriate termination (unless stated otherwise, a 600 Ω termination).
- **PMRP** is the applied sound pressure at the MRP.
- Corr is 20 log (PMRP/PHFRP) of the used artificial mouth.
  The value of Corr is the value required to calibrate the artificial mouth to the exact value of Dcorr (e.g. 24,0 dB for 50 cm distance).
- Dcorr 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.

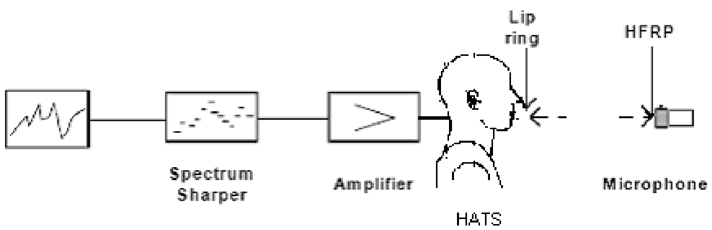


Figure 5.5.1-1: Calibration at HFRP for HATS

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

Desktop terminal:	50 cm and level to adjust -28,7 dBPa, Dcorr = 24 dB
Handheld terminal:	30 cm with -24,3 dBPa, Dcorr = 19,6 dB
Softphone:	36 cm with -25,8 dBPa, Dcorr = 21,1 dB

5.5.2 Receive

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

All measurement values produced by HATS are intended to be free-field equalized according Recommendation ITU-T P.581 [16].

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 [25].

ETSI TS 103 224 [25] 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.

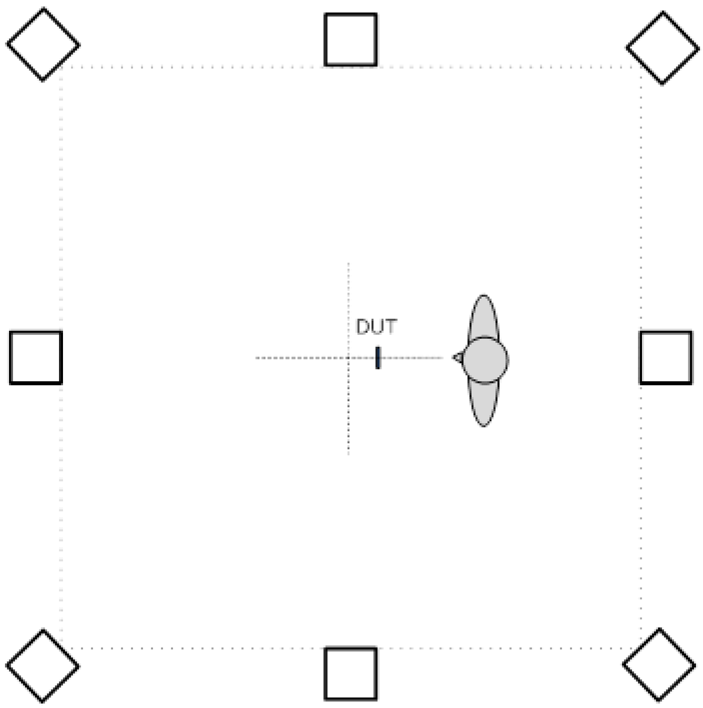


Figure 5.5.3-1: Loudspeaker arrangement for background noise simulation

The equalization and calibration procedure for the setup is described in detail in ETSI TS 103 224 [25].

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 [25] 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 due to P.58 limitations.	

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

Table 5.7-2: Accuracy of generated signals

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 8 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 conform to IEC 61672 [i.3] Type 1.

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

Requirement

The send sensitivity frequency response from the MRP to the measurement output (digital or analogue output according measurement system used) shall be 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	Lower limit
100	-8	-
200	0	-
300	0	-12
1 000	0	-6
2 000	4	-6
3 000	4	-6
3 400	4	-9
4 000	0	
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (Hz) scale.		

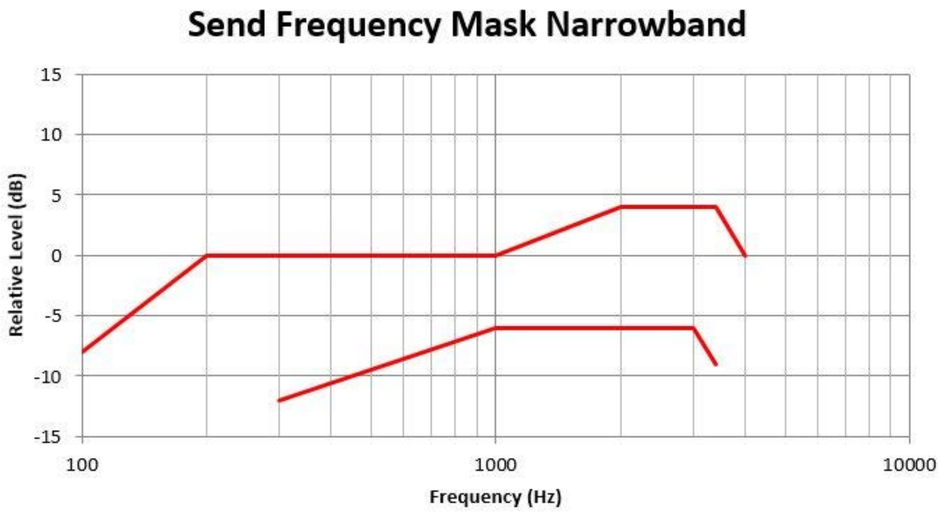


Figure 6.1.1-1: Hands-free send sensitivity/frequency response

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 [14]. The spectrum of acoustic signal produced by the artificial mouth is calibrated under free field conditions at the MRP. The signal level is adjusted according to clause 5.5.1.

The terminal is set up as described in clause 5.2.

Measurements shall be made at 1/3rd octave intervals as given by IEC 61260-1 [18] for frequencies from 100 Hz to 4 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

Requirement

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	Lower limit
100	6	
500	6	-9
630	6	-6
3 150	6	-6
4 000	6	-∞
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (kHz) scale.		

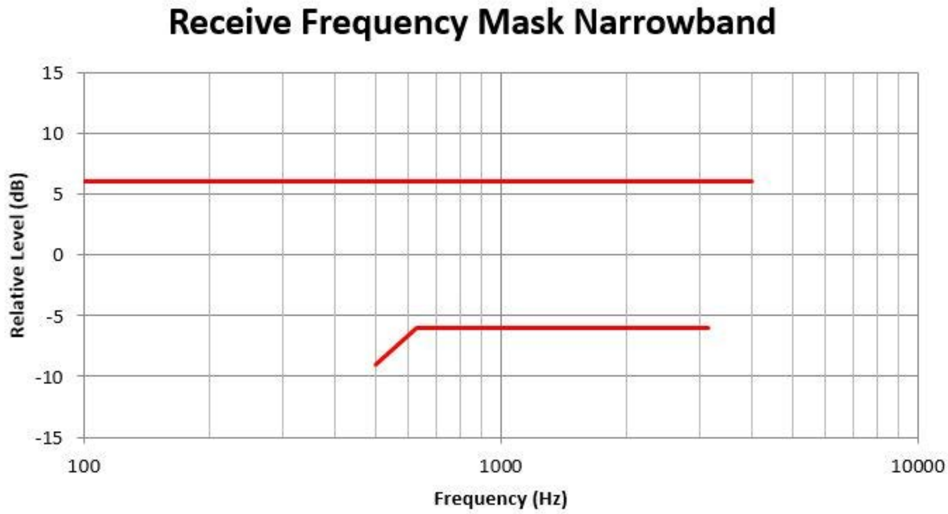


Figure 6.1.2.1-1: Handheld receive sensitivity/frequency response

Measurement method

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

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

S_{Jeff} Receive Sensitivity; Junction to HATS Ear with freefield 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 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 [14]. The test signal level shall be -16 dBm0, measured according to Recommendation ITU-T P.56 [9] at the digital reference point or the equivalent analogue point.

The terminal is set up as described in clause 5.2.

The measurement is conducted at nominal value of volume control.

HATS is freefield equalized as described in Recommendation ITU-T P.581 [16]. The equalized output signal is power-averaged on the total time of analysis. The 1/3rd octave band data are considered as the input signal to be used for calculations or measurements.

Measurements shall be made at 1/3rd octave intervals as in IEC 61260-1 [18] for frequencies from 100 Hz to 4 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)

Requirements

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

Type 2: The requirement as for handheld terminal shall apply.

6.1.2.4 Desktop Terminal

Requirement

The receive sensitivity frequency response from the measurement input (digital or analog input according measurement system used) to ear of HATS freefield 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 terminal receive sensitivity/frequency response

Frequency (Hz)	Upper limit	Lower limit
100	6	
315	6	-9
400	6	-6
3 150	6	-6
4 000	6	-∞
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (kHz) scale		

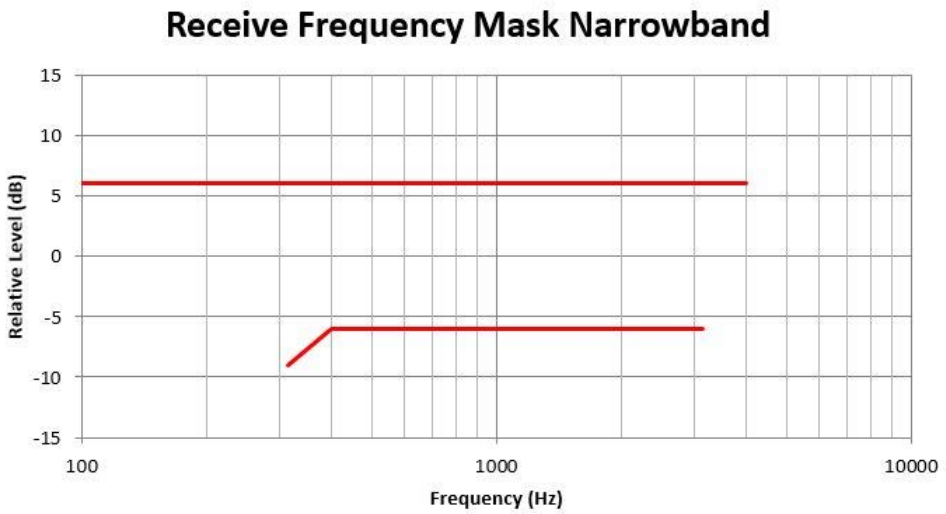


Figure 6.1.2.4-1: Desktop terminal receive sensitivity/frequency response

Measurement method

The terminal is positioned as described in clause 5.2.

Measurements methods defined 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 ± 3 dB

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 [14] The spectrum of acoustic signal produced by the artificial mouth is calibrated under freefield 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 terminal is set up as described in clause 5.2.

Calibration is realized as explained in clause 5.5.1.

The send sensitivity shall be calculated from each band of the 14 frequencies given in table 1 of Recommendation ITU-T P.79 [11], bands 4 to 17. For the calculation the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band at the MRP. SLR shall be calculated according to Recommendation ITU-T P.79 [11].

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 will be positioned 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 [14] The spectrum of acoustic signal produced by the artificial mouth is calibrated under freefield 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.

Calibration is realized as explained in clause 5.5.1.

The send sensitivity shall be calculated from each band of the 14 frequencies given in table 1 of Recommendation ITU-T P.79 [11], bands 4 to 17. For the calculation the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band at the MRP.

SLR shall be calculated according to Recommendation ITU-T P.79 [11].

6.2.3 Receive Loudness Ratings

Requirement

Handheld terminal

Nominal value of RLR shall 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 max ≤ +5 dB.

Range of volume control shall be equal or exceed 15 dB.

Desktop Terminal

Nominal value of RLR shall be +5 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 -2 dB: RLR max ≤ -2 dB.

Range of volume control shall be equal to or exceed 15 dB.

Softphone (computer-based terminals)

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

Type 2: requirement as for handheld terminal shall apply.

Group audio terminal

Nominal value of RLR shall be 5 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 -6 dB: RLR ≤ -6 dB.

Range of volume control shall be equal or exceed 19 dB.

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 [14]. The test signal level shall be -16 dBm0, measured at the electrical reference point and averaged over the complete test signal sequence.

The terminal is set up positioned as described in clause 5.2.

The RLR shall be calculated according to Recommendation ITU-T P.79 [11].

The receive sensitivity shall be calculated from each band of the 14 frequencies given in table 1 of Recommendation ITU-T P.79 [11], bands 4 to 17. 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 dB Pa/V and the RLR shall be calculated according to the formula 5-1 of Recommendation ITU-T P.79 [11], using the receive weighting factors from table 1 and according to clause 6. The RLR shall then be computed as RLR minus 14 dB according to Recommendation ITU-T P.340 [12] and without the LE factor.

The test shall be repeated for maximum volume control setting.

6.3 Send and receive noise

6.3.1 Send Noise

Requirement

The send noise level shall not exceed -64 dBm0p.

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 is set up 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 of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [14]. 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 psophometric noise level at the output of the test setup is measured. The psophometric filter is described in Recommendation ITU-T O.41 [17]. The send noise is measured at the POI in the frequency range from 100 Hz to 4 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.

Spectral peaks are measured in the frequency domain from 100 Hz to 3,4 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) 1/3rd 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

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

Octave band spectrum
The level in any 1/3rd 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: For softphone the test condition excludes fan noise.

Measurement method

The terminal is set up 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 of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [14]. 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 freefield equalization active. The level in any 1/3rd octave band, between 100 Hz and 10 kHz shall be measured at DRP of the artificial ear with the freefield equalization active.

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) 1/3rd 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 2: 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 coder-dependant requirements but to assess the electro acoustic performances of the terminal.

6.4.1 Send distortion

Requirement

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)
315	26
400	30
1 000	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 is set up as described in clause 5.2.

After the correct activation of the system, a sinewave signal at frequencies of 315 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz and 1 000 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 3,15 kHz.

The female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [14] shall be used for activation. 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	
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 (as for (desktop terminal) shall apply.

Type 2 requirement given in table 6.4.2-1 (as for handheld terminals) shall apply.

Measurement method

The terminal is set up as described in clause 5.2.

The signal used is an activation signal followed by a sine-wave signal at frequencies at 315 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz and 1 000 Hz is applied at the digital interface, 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 [14]. 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 [14] shall be used for activation. 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 weighted (TCLw)

Requirement

In order to meet talker echo objective requirements, the recommended weighted terminal coupling loss during single talk (TCL_W) should be greater than 55 dB when measured under freefield conditions at nominal setting of volume control.

The weighted Terminal Coupling loss TCL_W shall be ≥46 dB for all volume settings [if supplied].

NOTE 1: A $TCL_W \geq 50$ dB is recommended as a performance objective. Depending on the idle channel noise in the sending 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 test signal is the compressed real speech signal described in clause 7.3.3 of Recommendation ITU-T P.501 [14]. The signal level shall be -10 dBm0.

The TCL_W is calculated according to Recommendation ITU-T G.122 [2], clause B.4 (trapezoidal rule). 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).

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

NOTE 2: Care should be taken when measuring TCL_W so that the echo return is not to be masked by the residual noise or the comfort noise when implemented.

6.6 Stability Loss (or similar parameters)

Requirement

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

Measurement method

Test set-up is identical as for TCL_W. 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 [14] 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 [14] 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 200 Hz to 4 kHz.

6.7 Double talk performance

6.7.0 General

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 [12] and P.502 [15]):

- 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 to 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 $A_{H,S,dt}$

Requirement

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.

Table 6.7.1-1

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

In general, this table 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 done also for desktop terminals and softphones with variable echo path.

The category of the terminal according to table 6.7.1-1 shall be noted in the test report.

Measurement method

The terminal is set up as described in clause 5.2.

The long conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [14] 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 the double talk speech sequence as defined in clause 7.3.5 of Recommendation ITU-T P.501 [14] as shown in figure 6.7.1-1. The competing speaker is always inserted as the double talk sequence sdt(t) either in send or receive and is used for analysis.

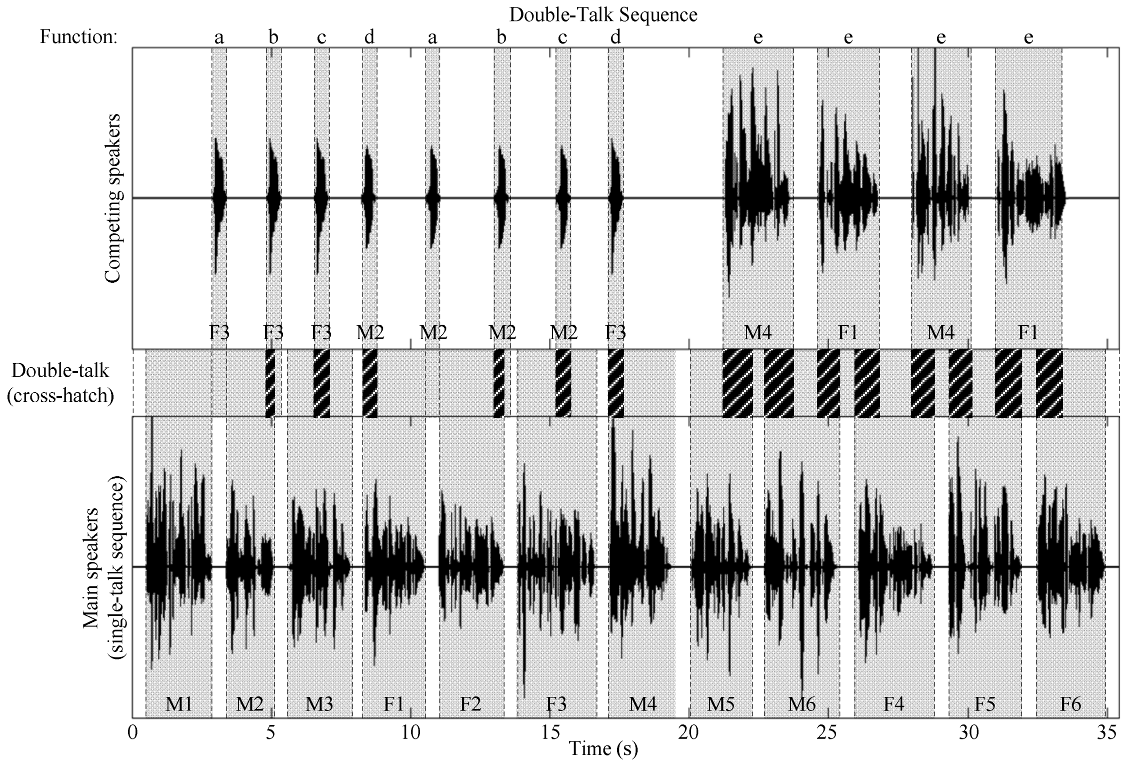


Figure 6.7.1-1: Double Talk Test Sequence with overlapping speech sequences in send and receive direction

The attenuation range during double talk is determined as described in Appendix III of Recommendation ITU-T P.502 [15]. The double talk performance is analysed for the sequence of words and the sequence of sentences produced by the competing speaker. The requirement has to be met for both sequences.

The settings for the test signals are according to table 6.7.1-2.

Table 6.7.1-2

	Receive direction	Send direction
Average signal level for speakerphone HFT	-16 dBm0	-1,7 dBPa
Average signal level for desktop HFT	-16 dBm0	-1,7 dBPa

When determining the attenuation range in send direction, the signal measured at the electrical reference point is referred to the test signal inserted.

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

Requirement

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.

Table 6.7.2-1

Category (according to Recommendation ITU-T P.340 [12])	1	2a	2b	2c	3
	Full Duplex Capability	*Partial Duplex Capability*			*No Duplex Capability*
$A_{H,R,dt}$ ***[dB]***	*≤ 3*	*≤ 5*	*≤ 8*	*≤ 10*	*> 10*

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.

The measurement shall be done also for desktop terminals and softphones with variable echo path.

The category of the terminal according to table 6.7.2-1 shall be noted in the test report.

Measurement method

Test setup is described in clause 5.2.

The long conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [14] 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. A sequence of speech signals shall be used which is inserted in parallel in send and receive direction. The test signals are synchronized in time at the acoustical interface. The delay of the test arrangement shall be constant during the measurement.

The settings for the test signals are according to table 6.7.2-2.

Table 6.7.2-2

	Receive direction	Send direction
Average signal level for speakerphone HFT	-16 dBm0	-1,7 dBPa
Average signal level for desktop HFT	-16 dBm0	-1,7 dBPa

When determining the attenuation range in receive direction, the signal measured at the loudspeaker of the hands-free terminal is referred to the test signal inserted.

The attenuation range during double talk is determined as described in Appendix III of Recommendation ITU-T P.502 [15]. The double talk performance is analysed for the sequence of words and the sequence of sentences produced by the competing speaker. The requirement has to be met for the sequence of words and the sequence of sentences produced by the competing speaker.

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. Under these conditions the requirements given in table 6.7.3-1 are applicable (more information can be found in annex A of the Recommendation ITU-T P.340 [12]).

Table 6.7.3-1

Category (according to Recommendation ITU-T P.340 [12])	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 measurement shall be done also for softphones and desktop terminals.

The category of the terminal according to table 6.7.3-1 shall be noted in the test report.

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 [14].

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

The signal settings are given in table 6.7.3-2.

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

Sending direction		Receiving direction	
$f_0^{(1)}$ (Hz)	$\pm\Delta f^{(1)}$ (Hz)	$f_0^{(2)}$(Hz)	$\pm\Delta f^{(2)}$ (Hz)
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
NOTE: Parameters of the Shaping Filter: f ≥ 250 Hz: Low Pass Filter, 5 dB/oct.			

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 [14]. 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 Introduction

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 [14] with increasing level for each single word.

The settings of the test signal are as follows:

Table 6.8.1-1: Settings for the signal

	Single word Duration/ Pause Duration	**Level of the Single word (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 note)	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 test signal according to Recommendation ITU-T P.501 [14]. NOTE 2: The signal level is determined for each utterance individually according to Recommendation ITU-T P.56 [9].			

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 psophometric 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
NOTE: All sensitivity values are expressed in dB on an arbitrary scale.		

Measurement method

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

The terminal is set up as described in clause 5.2.

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 [14] 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 narrowband systems shall be tested based on ETSI TS 103 106 [21]. The test method described leads to three MOS-LQO quality numbers:

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

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

- N-MOS-LQOn ≥ 3.0.
- S-MOS-LQOn ≥ 3.0.
- G-MOS-LQOn ≥ 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 terminal is set-up as described in clause 5.2. The background noise simulation as described in clause 5.5.3. is used.

The background noise shall be applied for at least 5 s in order to adapt noise reduction algorithms in advance of 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 [21]. 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 6 kHz.
3) The send signal is recorded at the electrical reference point.

N-MOS-LQOn, S-MOS LQOn and G-MOS LQOn are calculated as described in ETSI TS 103 106 [21].

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. The test is carried out by comparing the noise level transmitted in the sending direction under reference conditions with no far end speech, to the noise level transmitted in the sending direction under test conditions including far end speech. During and after the end of the speech signal the signal level in send direction should 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 (receiving 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 nonlinear processor working in conjunction with the echo canceller and erroneously switching or modulating the transmitted background noise.

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

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 should 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 [14] is applied in receive direction with duration of at least 10 s. The test signal level in the receiving 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

Requirement

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 terminal is set up as described in clause 5.2.

The test signal consists of periodically repeated Composite Source Signal according to Recommendation ITU-T P.501 [14], 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 [14] 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 by the integration time (35 ms) of the level analysis taking into account the exponential character of the integration time in any tolerance scheme.

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	-20 dB
200 Hz	-30 dB
300 Hz	-38 dB
800 Hz	-34 dB
1 500 Hz	-33 dB
2 600 Hz	-24 dB
4 000 Hz	-24 dB
NOTE 1: All sensitivity values are expressed in dB on an arbitrary scale. NOTE 2: 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 terminal is set up as 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 [14]. 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 [14]. 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 [14] 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.

No level peak shall be more than 10dB above the minimum noise level during the measurement.

6.11 Send and receive delay or round trip delay

Requirements

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 of the terminal T_{rtd} (sum of send and receive delay) shall be less than 100 ms (category B in Recommendation ITU-T P.1010 [i.7]). From the user perspective, a value less than 50 ms (Category A in Recommendation ITU-T P.1010 [i.7]) is preferred.

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 to POI. The delay measured in send direction is $Ts+ t_{system}$

NOTE 1: The delay should be minimized! This can, e.g. be accomplished by designing the speech decoder output, the additional radio link, and the hands-free system in a way, that sample-based processing and frame-based processing interoperate by using common buffers at their interfaces.

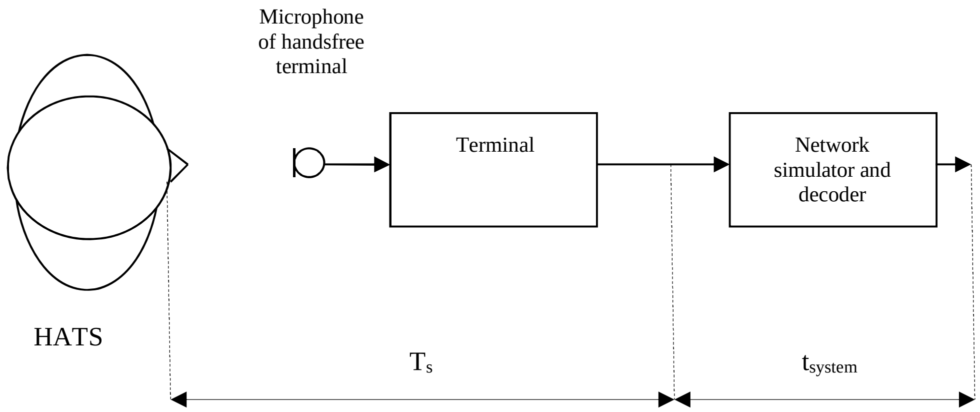


Figure 6.11-1: Different blocks contributing to the delay in send direction

The system delay t_{system} is depending on the transmission method used and the network simulator. The delay t_{system} shall be known:

1) For the measurements a Composite Source Signal (CSS) according to Recommendation ITU-T P.501 [14] 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 [16]). The equalization of the artificial mouth is made at the MRP. The reference signal is the original signal (test signal).

   - The setup of the hands-free terminal is in correspondence to 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 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 to the Drum Reference Point (DRP). The delay measured in receive direction is T_r+ t_{system}.

NOTE 2: The delay should be minimized! This can, e.g. be accomplished by designing the speech decoder output, the additional radio link, and the hands-free system in a way, that sample-based processing and frame-based processing interoperate by using common buffers at their interfaces. Careful matching of frame shift and DFT size for the signal processing in the hands-free system to the additional radio link and to the speech coder allows to (partially) embed the delay of one block into the preceding one.

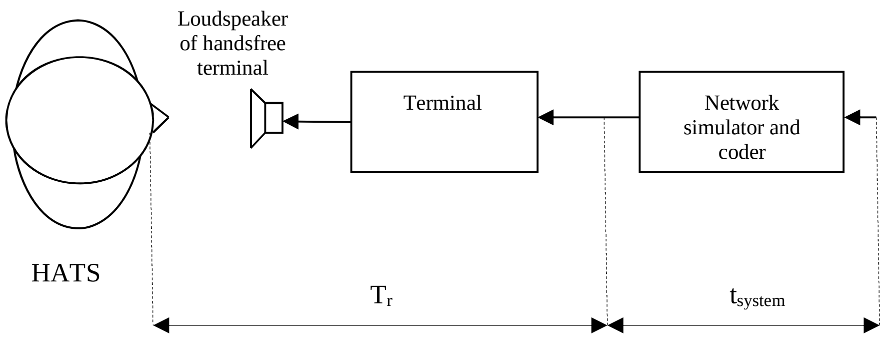


Figure 6.11-2: Different blocks contributing to the delay in receive direction

The system delay t_{system} is depending on the transmission system and on the network simulator used. The delay t_{system} shall be known:

1) For the measurements a Composite Source Signal (CSS) according to Recommendation ITU-T P.501 [14] 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 according to clause 5.2. Artificial head is free field adjusted according to Recommendation ITU-T P.581 [16]. 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 Coders

The present document is intended to be applicable for different speech coders implemented in access networks and additional links.

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

Table 7.1-1: List of speech coders

System	Codec
GSM 850, 900, 1 800, 1 900	AMR-NB ETSI TS 126 171 [1] GSM Full Rate Codec (ETSI TS 146 010 [19]), Enhanced FullRate (ETSI TS 146 060 [20])
UMTS (WCDMA), VoLTE	Recommendation ITU-T G.729 [6]
	EVS-NB (ETSI TS 126 441) [24]
Over The Top (OTT)	Recommendation ITU-T G.729.1 [7]
	Recommendation ITU-T G.711 [4], with PLC
	Recommendation ITU-T G.726 [5]

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 speech quality

7.3.0 Introduction

The terminal is set up as described in clause 5.2.

The measurements of this section are only applicable if the 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 are as follows.

Table 7.3.1-1

Speech coder	MOS-LQON (P.863 or TOSQA 2001)	MOS-LQOM (TOSQA 2001)
AMR-NB ETSI TS 126 171 @ 12,2 kbit/s [1]	**(ffs)**	**(ffs)**
EVS-NB ETSI TS 126 441 @ 13,2 kbit/s [24]	**(ffs)**	**(ffs)**
Recommendation ITU-T G.711 [4]	**(ffs)**	**(ffs)**
Recommendation ITU-T G.726 [5]	**(ffs)**	**(ffs)**
Recommendation ITU-T G.729 [6]	**(ffs)**	**(ffs)**
Recommendation ITU-T G.729.1 @ 8 kbit/s [7]	**(ffs)**	**(ffs)**

NOTE 1: Not sufficient experience is available so far with Recommendation ITU-T P.863 [22] and TOSQA 2001 (ETSI EG 201 377-1 [i.6] for measuring hands-free terminals. Therefore, the numbers for MOS-LQON and MOS-LQOM are for further study.

NOTE 2: The use of the codecs G.729 [6] and G.729.1 [7] is not recommended due to low quality.

Measurement method

For the assessment of objective listening speech quality, fullband mode of Recommendation ITU-T P.863 [22] shall be applied, taking Appendix III into account (see note 2).

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 [23]. The 4 sentence pairs are taken from Recommendation ITU-T P.501 [14]. It shall be stated, which sentence pairs were used. The test signal level shall be -1,7 dBPa, measured according to Recommendation ITU-T P.56 [9] at the MRP. 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 3: In narrowband acoustics, Recommendation ITU-T P.863 [22] is recommending using the fullband mode with a narrowband reference signal, resulting in a prediction on the narrowband scale (MOS-LQON).

NOTE 4: An alternative test method is TOSQA 2001 (ETSI EG 201 377-1) [i.6]. With TOSQA, terminals used in narrowband and wideband mode should be measured based on MOS-LQON. 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 requirements for the listening speech quality and the delay under clean network conditions are as follows.

Table 7.3.2.1-1

Speech coder	MOS-LQON (P.863)	MOS-LQOM (TOSQA 2001)
AMR-NB ETSI TS 126 171 @12,2 kbit/s [1]	(ffs)	(ffs)
EVS-NB ETSI TS 126 441 @13.2 kbit/s [24]	(ffs)	(ffs)
Recommendation ITU-T G.711 [4]	(ffs)	(ffs)
Recommendation ITU-T G.726 [5]	(ffs)	(ffs)
Recommendation ITU-T G.729 [6]	(ffs)	(ffs)
Recommendation ITU-T G.729.1@ 8 kbit/s [7]	(ffs)	(ffs)

NOTE 1: Not sufficient experience is available so far with Recommendation ITU-T P.863 [22] and TOSQA 2001 (ETSI EG 201 377-1 [i.6]) for measuring hands-free terminals. Therefore, the numbers for MOS-LQON and MOS-LQOM are for further study.

NOTE 2: The use of the codecs G.729 [6] and G.729.1 [7] is not recommended due to low quality.

Measurement method

For the assessment of objective listening speech quality, fullband mode of Recommendation ITU-T P.863 [22] shall be applied, taking Appendix III into account (see note 3).

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 [23]. The 4 sentence pairs are taken from Recommendation ITU-T P.501 [14]. It shall be stated, which sentence pairs were used. The test signal level is averaged over all sentence pairs (4 sentence pairs). The measurement is done 4 times, every time using another pair of the speech sentences. The result of the measurement is the averaged value of all 4 measurements.

NOTE 3: In narrowband acoustics, Recommendation ITU-T P.863 [22] is recommending using the fullband mode with a narrowband reference signal, resulting in a prediction on the narrowband scale (MOS-LQON).

NOTE 4: An alternative test method is TOSQA 2001 (ETSI EG 201 377-1 [i.6]). 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.

Requirements

The degradation between the error- and jitter-free condition (equals network condition 1) and impairment conditions shall not exceed the delta-values provided 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
AMR-NB ETSI TS 126 171 [1] @ 12,2 kbit/s	0	-0,1	-0,1	≤ Trtd + 5 ms
	1	0,0	0,0	≤ Trtd + 5 ms
	2	[ffs]	[ffs]	≤ Trtd + 25 ms
	3	[ffs]	[ffs]	≤ Trtd + 5 ms
	4	[ffs]	[ffs]	≤ Trtd + 25 ms
	5	[ffs]	[ffs]	≤ Trtd + 5 ms
EVS-NB @ 13,2 kbps	0	-0,1	-0,1	≤ Trtd + 5 ms
	1	0,0	0,0	≤ Trtd + 5 ms
	2	[ffs]	[ffs]	≤ Trtd + 25 ms
	3	[ffs]	[ffs]	≤ Trtd + 5 ms
	4	[ffs]	[ffs]	≤ Trtd + 25 ms
	5	[ffs]	[ffs]	≤ Trtd + 5 ms
Recommendation ITU-T G.711 [4]	0	-0,1	-0,1	≤ Trtd + 5 ms
	1	0,0	0,0	≤ Trtd + 5 ms
	2	-0,2	-0,2	≤ Trtd + 25 ms
	3	-0,2	-0,2	≤ Trtd + 5 ms
	4	-0,2	-0,2	≤ Trtd + 25 ms
	5	-0,2	-0,2	≤ Trtd + 5 ms
Recommendation ITU-T G.726 [5]	0	-0,1	-0,1	≤ Trtd + 5 ms
	1	0,0	0,0	≤ Trtd + 5 ms
	2	[ffs]	[ffs]	≤ Trtd + 25 ms
	3	[ffs]	[ffs]	≤ Trtd + 5 ms
	4	[ffs]	[ffs]	≤ Trtd + 25 ms
	5	[ffs]	[ffs]	≤ Trtd + 5 ms
Recommendation ITU-T G.729 [6]	0	-0,1	-0,1	≤ Trtd + 5 ms
	1	0,0	0,0	≤ Trtd + 5 ms
	2	-0,1	-0,1	≤ Trtd + 25 ms
	3	-0,1	-0,1	≤ Trtd + 5 ms
	4	-0,1	-0,1	≤ Trtd + 25 ms
	5	-0,5	-0,6	≤ Trtd + 5 ms
Recommendation ITU-T G.729.1 [7] @ 8 kbit/s	0	-0,1	-0,1	≤ Trtd + 5 ms
	1	0,0	0,0	≤ Trtd + 5 ms
	2	[ffs]	[ffs]	≤ Trtd + 25 ms
	3	[ffs]	[ffs]	≤ Trtd + 5 ms
	4	[ffs]	[ffs]	≤ Trtd + 25 ms
	5	[ffs]	[ffs]	≤ Trtd + 5 ms

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 (speech samples)

Condition	Packet Loss (Equal)	Delay Variation
0 (see note 2) (VAD)	0	No
1	0	No
2	0	20 ms (see note 1)
3	1 %	No
4	1 %	20 ms (see note 1)
5	3 %	No

NOTE 1: Delay variation produced with a Pareto-Distribution and r = 0,5.
NOTE 2: VAD on, all other conditions (1-5) 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: The settings are derived from the ones used in the ETSI Plugtest VoIP speech quality test events.
NOTE 5: The delay requirements for conditions with network impairments are based on the measured roundtrip delay of the terminal in the absence of network impairments Trtd (see clause 6.11). A small additional tolerance takes into account the variable behaviour of the delay.

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".
- ETSI TS 126 131: "Universal Mobile Telecommunications System (UMTS); LTE; Terminal acoustic characteristics for telephony; Requirements (3GPP TS 26.131 version 8.1.0 Release 8)".
- ETSI TS 126 132: "Universal Mobile Telecommunications System (UMTS); LTE; Speech and video telephony terminal acoustic test specification (3GPP TS 26.132 version 8.1.0 Release 8)".
- ETSI ES 202 737: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for narrowband VoIP loudspeaking and handsfree terminals from a QoS perspective as perceived by the user".
- ETSI ES 202 738: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for narrowband VoIP loudspeaking and handsfree terminals 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 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".
- ETSI EN 300 176-2: "Digital Enhanced Cordless Telecommunications (DECT); Test specification; Part 2: Speech".
- Recommendation ITU-T G.131: "Talker echo and its control".
- Recommendation ITU-T P.50: "Artificial voices".
- Recommendation ITU-T P.64: "Determination of sensitivity/frequency characteristics of local telephone systems".
- 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".
- 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".
- IETF RFC 3550: "RTP: A Transport Protocol for Real-Time Applications".

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