

ETSI TS 103 740 V1.1.1 (2009-11)

Technical Specification

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



Reference

DTS/STQ-00140-4

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

Important notice

Individual copies of the present document can be downloaded from:

<http://www.etsi.org>

The present document may be made available in more than one electronic version or in print. In any case of existing or perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF). In case of dispute, the reference shall be the printing on ETSI printers of the 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

<http://portal.etsi.org/tb/status/status.asp>

If you find errors in the present document, please send your comment to one of the following services:

http://portal.etsi.org/chaicor/ETSI_support.asp

Copyright Notification

No part may be reproduced except as authorized by written permission.
The copyright and the foregoing restriction extend to reproduction in all media.

© European Telecommunications Standards Institute 2009.
All rights reserved.

DECT™, **PLUGTESTS™**, **UMTS™**, **TIPHON™**, the TIPHON logo and the ETSI logo are Trade Marks of ETSI registered for the benefit of its Members.

3GPP™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners.

LTE™ is a Trade Mark of ETSI currently being registered

for the benefit of its Members and of the 3GPP Organizational Partners.

GSM® and the GSM logo are Trade Marks registered and owned by the GSM Association.

Contents

Intellectual Property Rights	5
Foreword.....	5
Introduction	5
1 Scope	6
2 References	6
2.1 Normative references	7
2.2 Informative references.....	7
3 Definitions and abbreviations.....	8
3.1 Definitions	8
3.2 Abbreviations	8
4 Configurations and interfaces.....	9
4.1 Access networks.....	9
4.2 Additional (radio) links between the terminal and external electroacoustical devices.....	9
5 Test Configurations	9
5.1 Set-up interface	9
5.2 Set-up for terminals	10
5.2.1 Hand-held terminal	10
5.2.2 Vehicle mounted hands-free	11
5.2.3 Desktop hands-free terminal	12
5.2.4 Additional test setup for handsfree function with softphone	13
5.2.4.1 Softphone including speakers and microphone.....	13
5.2.4.2 Softphone with separate speakers	14
5.3 Acoustical environment.....	15
5.4 Test signals.....	15
5.5 Calibration and test signal level	16
5.5.1 Send	16
5.5.2 Receive	16
5.5.3 Setup of background noise simulation.....	16
5.6 Environmental conditions for tests.....	17
5.7 Accuracy of test equipment	18
5.8 Power feeding conditions	18
6 Codec independent requirements and associated Measurement Methodologies.....	18
6.1 Send and receive frequency response	18
6.1.1 Send frequency response	18
6.1.2 Receive frequency response.....	20
6.1.2.1 Handheld terminal.....	20
6.1.2.2 Vehicle mounted hands-free.....	21
6.1.2.3 Softphone (computer-based terminals)	22
6.1.2.4 Desktop Terminal.....	22
6.2 Send and receive loudness ratings.....	22
6.2.1 Send Loudness Ratings	22
6.2.2 Receive Loudness Ratings.....	23
6.2.2.1 Handheld terminal.....	23
6.2.2.2 Vehicle mounted hands-free.....	23
6.2.2.3 Softphone (computer-based terminals)	24
6.2.2.4 Desktop Terminal.....	24
6.3 Send and receive noise	24
6.3.1 Send Noise	24
6.3.1 Receive Noise	25
6.4 Send and receive distortion	25
6.4.1 Send distortion	25
6.4.2 Receive distortion	26

6.5	TCLw (or similar parameters)	27
6.5.1	Handheld Terminal	27
6.5.2	Vehicle mounted hands-free	27
6.5.3	Softphone (computer-based terminals)	28
6.5.4	Desktop Terminal	28
6.6	Stability Loss (or similar parameters)	28
6.7	Double talk performance	29
6.7.1	Attenuation Range in Send Direction during Double Talk	29
6.7.2	Attenuation Range in Receive Direction during Double Talk	31
6.7.3	Detection of echo components during double Talk	31
6.7.4	Minimum activation level and sensitivity of double talk detection	33
6.8	Switching parameters	33
6.8.1	Activation in Send Direction	34
6.8.2	Minimum activation level and sensitivity in Receive direction	35
6.8.3	Automatic level control	35
6.8.4	Silence Suppression and Comfort Noise Generation	35
6.9	Background noise performance	35
6.9.1	Performance in send direction in the presence of background noise	35
6.9.2	Speech Quality in the Presence of Background Noise	36
6.9.3	Quality of Background Noise Transmission (with Far End Speech)	37
6.9.4	Quality of Background Noise Transmission (with Near End Speech)	37
6.10	Quality of echo cancellation	38
6.10.1	Temporal echo effects	38
6.10.2	Spectral Echo Attenuation	38
6.10.3	Occurrence of Artefacts	39
6.11	Send and receive delay or round trip delay	39
6.12	Objective listening Quality in send and receive direction	41
7	Codec dependent requirements and associated Measurement Methodologies	41
7.1	Speech Coders	41
7.2	Send and receive delay or round trip delay	41
7.3	Objective listening Quality in send and receive direction	41
8	Requirements and associated Measurement Methodologies (with an additional radio link between the terminal and external electroacoustical devices)	41
	Annex A (informative): Bibliography	42
	History	43

Intellectual Property Rights

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

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

Foreword

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

Introduction

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

1 Scope

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

References are either specific (identified by date of publication and/or edition number or version number) or non-specific.

- For a specific reference, subsequent revisions do not apply.
- Non-specific reference may be made only to a complete document or a part thereof and only in the following cases:
 - if it is accepted that it will be possible to use all future changes of the referenced document for the purposes of the referring document;
 - for informative references.

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

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

2.1 Normative references

The following referenced documents are indispensable for the application of the present document. For dated references, only the edition cited applies. For non-specific references, the latest edition of the referenced document (including any amendments) applies.

- [1] ETSI EG 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".
- [2] ITU-T Recommendation G.122: "Influence of national systems on stability and talker echo in international connections".
- [3] ITU-T Recommendation G.722: "7 kHz audio-coding within 64 kbit/s".
- [4] ITU-T Recommendation G.729.1: "G.729 based Embedded Variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable with G.729".
- [5] ITU-T Recommendation G.131: "Talker echo and its control".
- [6] ITU-T Recommendation P.50: "Artificial voices".
- [7] ITU-T Recommendation P.56: "Objective measurement of active speech level".
- [8] ITU-T Recommendation P.58: "Head and torso simulator for telephony".
- [9] ITU-T Recommendation P.79 and Corrigendum 2 (2001): "Calculation of loudness ratings for telephone sets".
- [10] ITU-T Recommendation P.340: "Transmission characteristics and speech quality parameters of hands-free terminals".
- [11] ITU-T Recommendation P.341: "Transmission characteristics for wideband (150-7 000 Hz) digital hands-free telephony terminals".
- [12] ITU-T Recommendation P.501: "Test signals for use in telephony".
- [13] ITU-T Recommendation P.581: "Use of head and torso simulator (HATS) for hands-free terminal testing".
- [14] ISO 3 (1973): "Preferred numbers - Series of preferred numbers".
- [15] ITU-T Recommendation G.711.1: "Wideband embedded extension for G.711 pulse code modulation".
- [16] ITU-T Recommendation G.722.2: "Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)".
- [17] IEC 61672: "Electroacoustics - Sound level meters".

2.2 Informative references

The following referenced documents are not essential to the use of the present document but they assist the user with regard to a particular subject area. For non-specific references, the latest version of the referenced document (including any amendments) applies.

- [i.1] ITU-T Recommendation G.107: "The E-model, a computational model for use in transmission planning".
- [i.2] ITU-T Recommendation G.108: "Application of the E-model: A planning guide".
- [i.3] ITU-T Recommendation G.109: "Definition of categories of speech transmission quality".
- [i.4] ITU-T Recommendation P.502: "Objective test methods for speech communication systems using complex test signals".

- [i.5] ITU-T Recommendation 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".
- [i.6] IEC 61260: "Electroacoustics - Octave-band and fractional-octave-band filters".
- [i.7] IUT-T Recommendation P.1100: "Narrowband hands-free communication in motor vehicles".
- [i.8] ETSI EG 202 396-3: "Speech Processing, Transmission and Quality Aspects (STQ); Speech Quality performance in the presence of background noise Part 3: Background noise transmission - Objective test methods".
- [i.9] ITU-T Recommendation P.800.1: "Mean Opinion Score (MOS) terminology".

3 Definitions and abbreviations

3.1 Definitions

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

freefield equalization: artificial head is equalized in such a way that for frontal sound incidence in anechoic conditions the frequency response of the artificial head is flat

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 center of the artificial head with no artificial head present)

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

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

Mouth Reference Point (MRP): is 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, the setting which is closest to the nominal RLR

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
CSS	Composite Source Signal
DECT	Digital Enhanced Cordless Telecommunications
DRP	ear Drum Reference Point
EL	Echo Loss
HATS	Head And Torso Simulator
HFRP	Hands Free Reference Point
MOS-LQOy	Mean Opinion Score - Listening Quality Objective

MRP	Mouth Reference Point
NOTE:	See ITU-T Recommendation P.800.1 [i.9].
NLP	Non Linear Processor
POI	Point Of Interconnect
QoS	Quality of Service
RLR	Receive Loudness Rating
SLR	Send Loudness Rating
TCL _w	Terminal Coupling Loss (weighted)
TELR	Talker Echo Loudness Rating

4 Configurations and interfaces

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, DECT, Bluetooth, WIFI, WIMAX, CDMA, ...

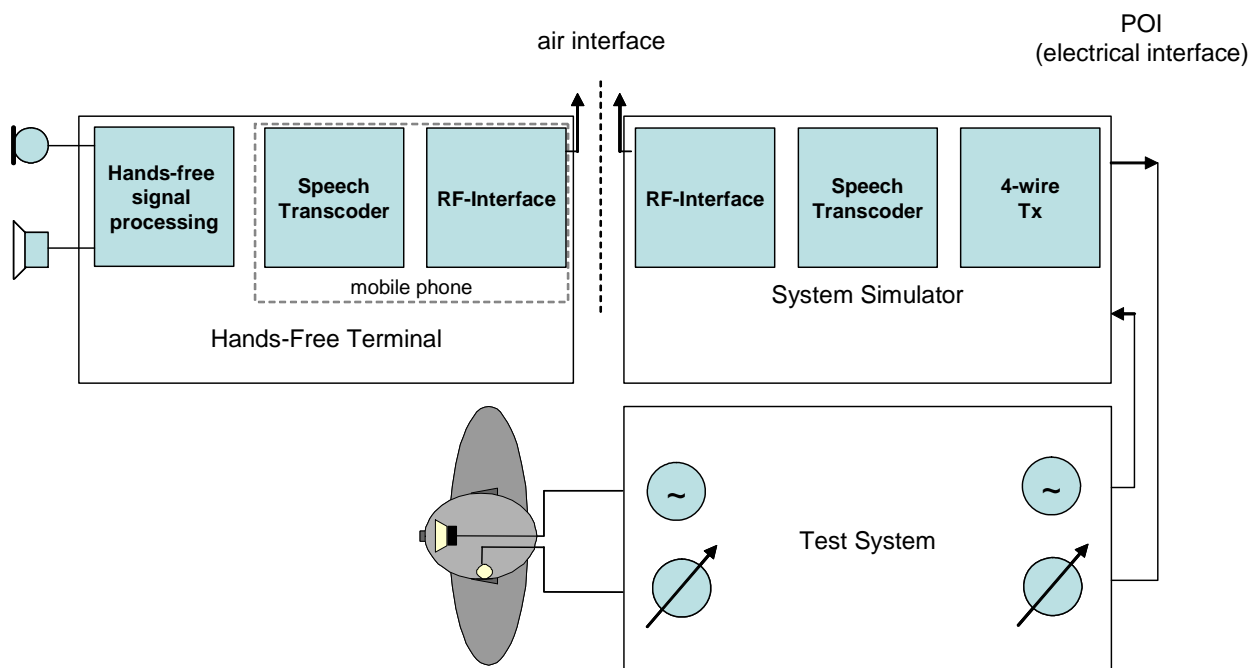
4.2 Additional (radio) links between the terminal and external electroacoustical devices

The present document also apply 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 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 the figure is not the additional radio link.

Figure 5.1: Set-up interface

5.2 Set-up for terminals

For electroacoustical testing, HATS as described in ITU-T Recommendation P.58 [8] 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.:

- ITU-T Recommendation G.729.1 [4]: 32 kbit/s.

5.2.1 Hand-held terminal

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

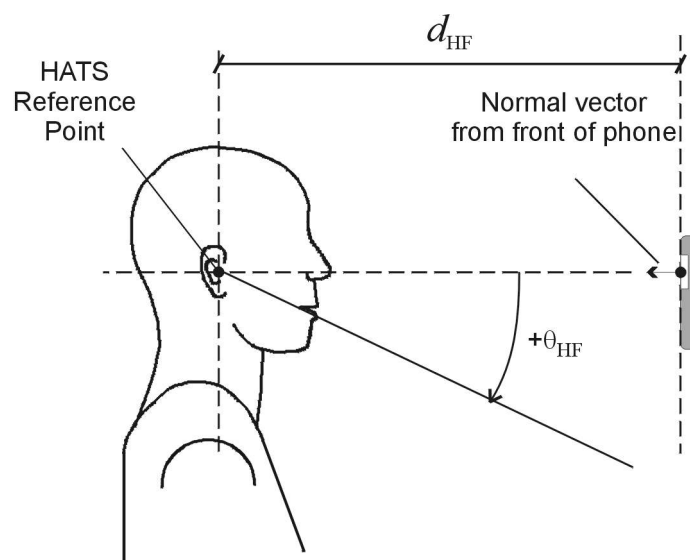


Figure 5.2: Configuration of Hand-held Hands-free UE relative to the HATS

5.2.2 Vehicle mounted hands-free

The hands-free terminal is installed according to the requirements of the manufacturers. The positioning of the microphone/microphone array and loudspeaker are given by the manufacturer. If no position requirements are given, the test lab will fix the arrangement.

Typically, the terminal's microphone is positioned close to the rear-view mirror, the terminal's loudspeaker is typically positioned in the footwell of the driver, respectively of the co-driver. In any case the exact positioning has to be noted. Hands-free terminals installed by the car manufacturer are measured in the original arrangement.

The artificial head (HATS Head and Torso Simulator according to ITU-T Recommendation P.58 [8]) is positioned at the driver's seat for the measurement. The position has to be in line with the average user's position, therefore all positions and sizes of users have to be taken into account. Typically the 95 % of the tallest people and 5 % of the smallest people have to be considered. The size of these persons can be derived e.g. from the 'anthropometric data set' for the corresponding year (e.g. based on data used by the car manufacturers). The position of the HATS (mouth/ears) within the positioning arrangement is given individually by each car manufacturer.

The position used has to be reported in detail in the test report. If no requirements for positioning are given, the distance from the microphone to the MRP is defined by the test lab.

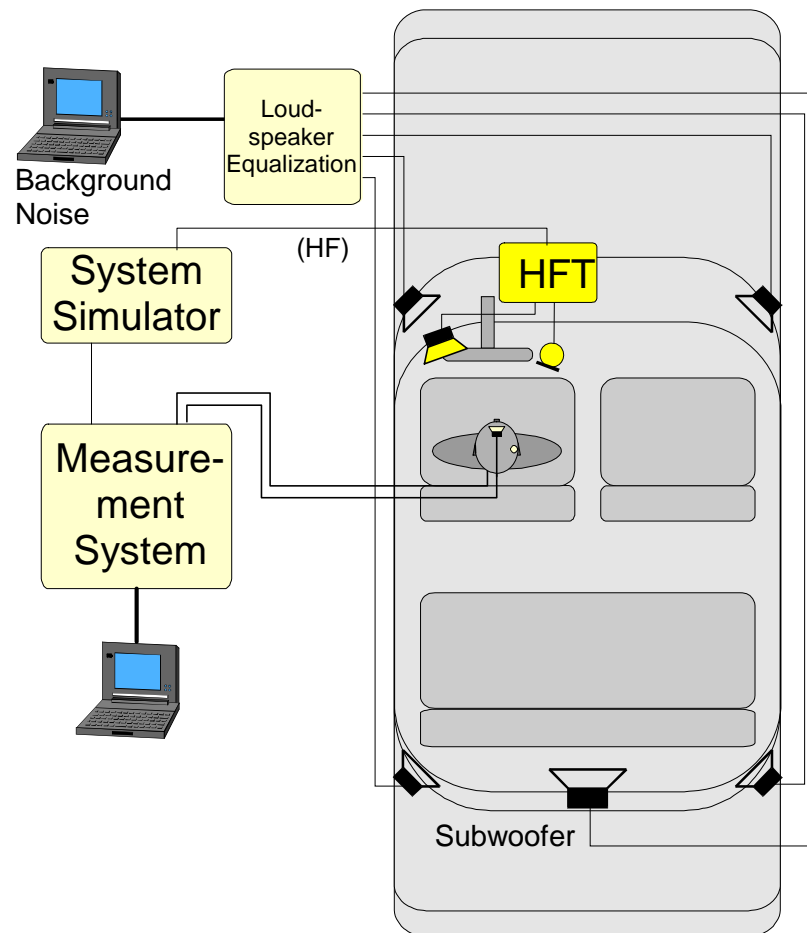


Figure 5.3: Test arrangement with background noise simulation

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 in ITU-T Recommendation P.581 [13].

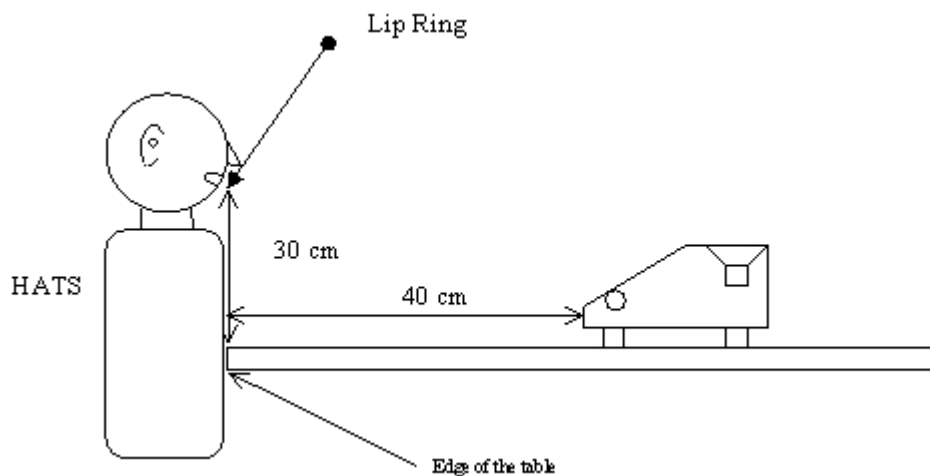


Figure 5.4: Position for test of desktop hands free terminal side view

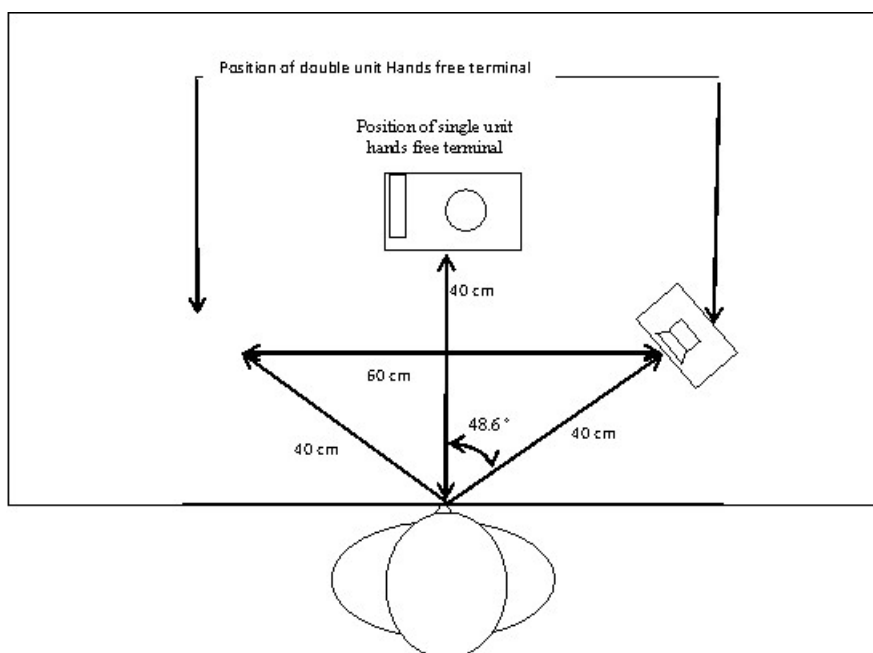


Figure 5.5: Position for test of desktop hands free terminal top sight

5.2.4 Additional test setup for handsfree function with softphone

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

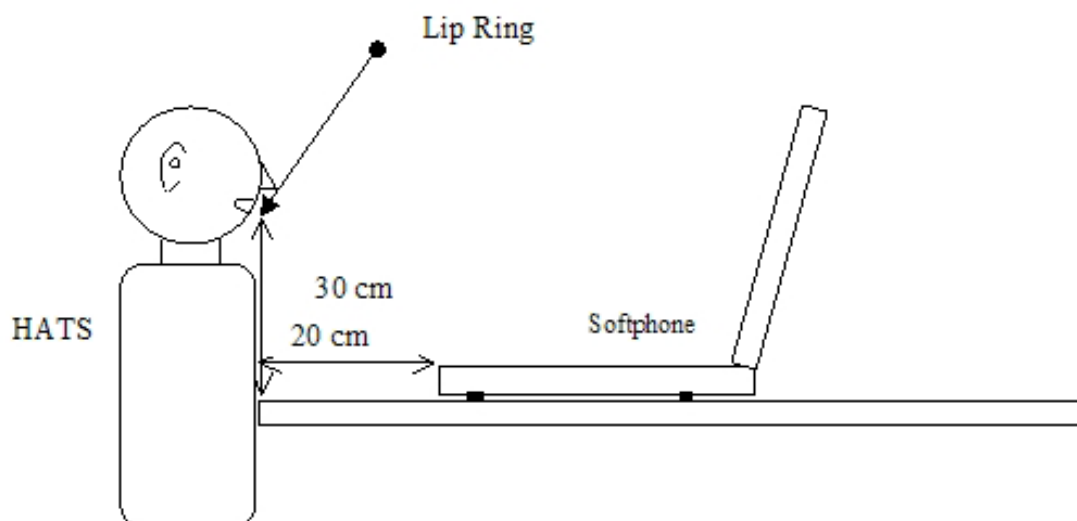


Figure 5.6: Configuration of softphone relative to the HATS side view

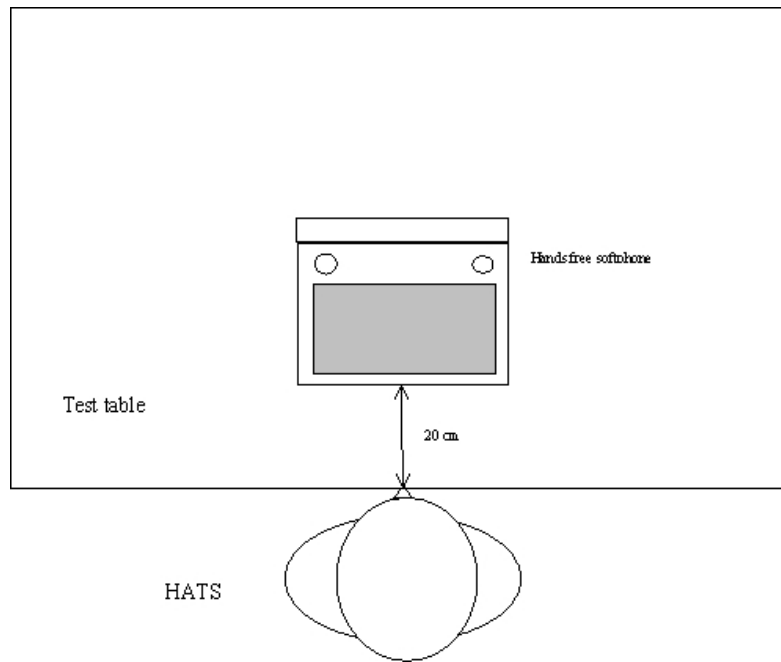


Figure 5.7: Configuration of softphone relative to the HATS-top sight

5.2.4.2 Softphone with separate speakers

When separate loudspeakers are used, system will be positioned as in figure 5.8.

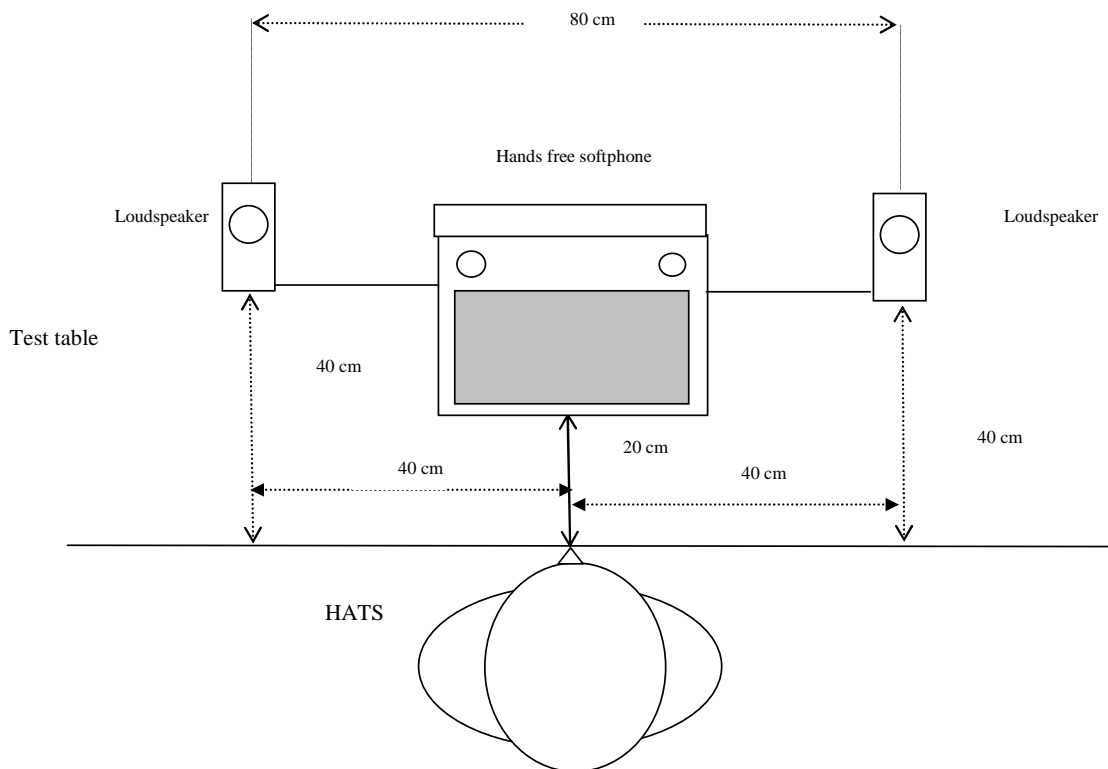


Figure 5.8: Configuration of softphone using external speakers relative to the HATS-top sight

When external microphone and speakers are used, system will be positioned as in figure 5.9.

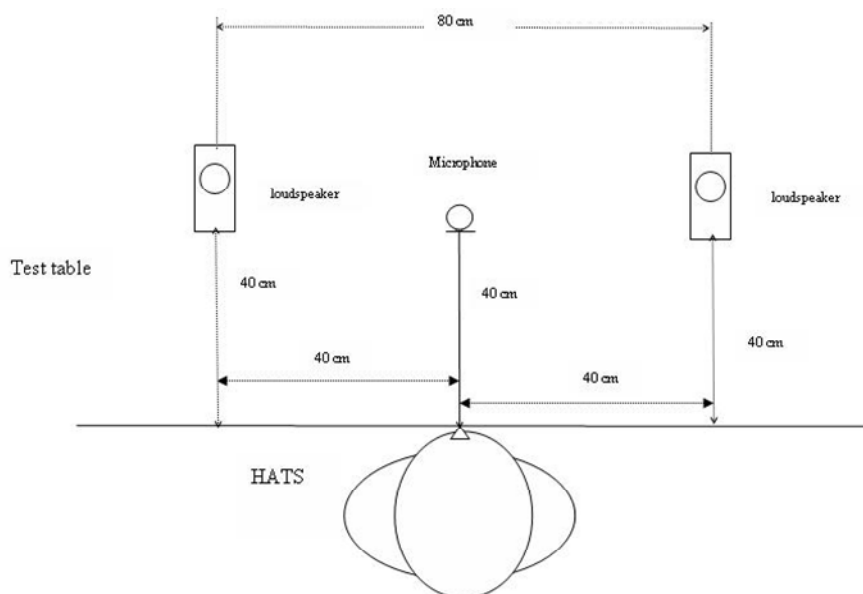


Figure 5.9: Configuration of softphone using external speakers and microphone relative to the HATS-top sight

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, test laboratory, in the case where its test room does not conform to anechoic conditions as given in ITU-T Recommendation 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 must 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 ITU-T Recommendation P.50 [6] and P.501 [12]. Normative requirements for the use of test signals from P.501 [12] are for further study. More information can be found in the test procedures described below.

For testing the wide-band 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.

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 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.

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 following procedure shall be used to perform the calibration of the artificial mouth of:

- The input signal from the artificial mouth of HATS is first calibrated under free-field conditions at the MRP. The total level on the frequency range is set to -4,7 dBPa.
- The spectrum at MRP is recorded.
- Then the level is adjusted to the level given further in this text (depending of type of terminal tested (for example -24,3 dBPa at 30 cm for handheld terminal)).
- The level at MRP (measured in third octave bands) adjusted at the first step (with total level of -4,7 dBPa) is used as the reference for send characteristics.

The test setup shall be in conformance with figure 5.10 but, depending on the type of terminal, the appropriate distance and level will be used. When using this calibration method, send sensitivity must be calculated as follows:

$$SmJ = 20 \log Vs - 20 \log PMRP$$

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 during the first step of calibration.

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.

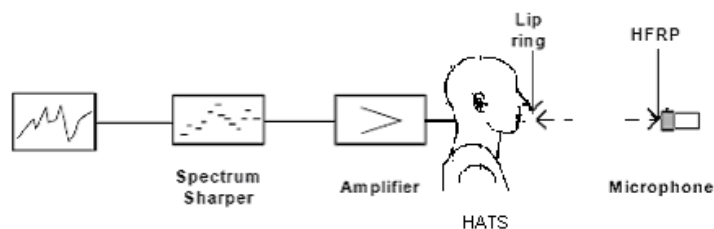


Figure 5.10: Calibration at HFRP for HATS

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

Desktop terminal: 50 cm and level to adjust - 28,7 dBPa.

Handheld terminal: 30 cm with - 24,3 dBPa.

Softphone: 36 cm with - 25,8 dBPa.

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 ITU-T Recommendation P.581 [13].

5.5.3 Setup of background noise simulation

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

EG 202 396-1 [1] 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 5.11.

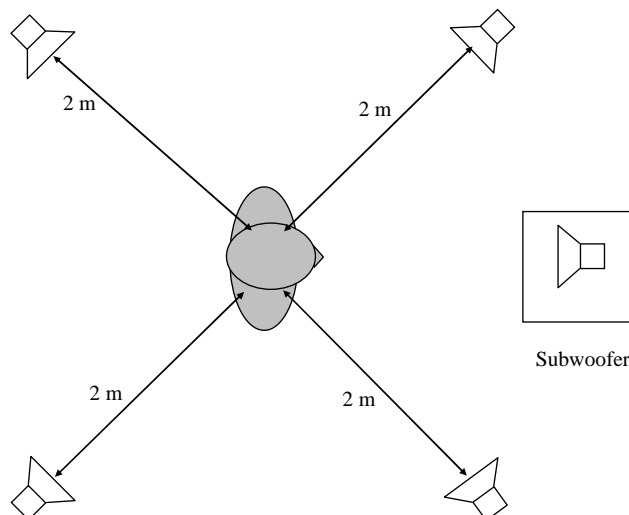


Figure 5.11: Loudspeaker arrangement for background noise simulation

The equalization and calibration procedure for the setup is described in detail in EG 202 396-1 [1].

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

The following noises of EG 202 396-1 [1] shall be used.

Table 5.1: Noises used for background noise simulation

Recording in pub	Pub_Noise_binaural	30 s	L: 77,8 dB(A) R: 78,9 dB(A)	binaural
Recording at sales counter	Cafeteria_Noise_binaural	30 s	L: 68,4 dB(A) R: 67,3 dB(A)	binaural
Recording in business office	Work_Noise_Office_Callcener_binaural	30 s	L: 56,6 dB(A) R: 57,8 dB(A)	binaural
Recording at the drivers position in a car	Midsized_Car1_130Kmh_binaural	30 s	L: 67,0 dB(A) R: 65,9 dB(A)	binaural

5.6 Environmental conditions for tests

The following conditions shall apply for the testing environment:

- Ambient temperature: 15 °C to 35 °C (inclusive);
- Relative humidity: 5 % to 85 %;
- 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 5.2: Accuracy of measurements

Item	Accuracy
Electrical Signal Power	$\pm 0,2$ dB for levels ≥ -50 dBm
Electrical Signal Power	$\pm 0,4$ dB for levels < -50 dBm
Sound pressure	$\pm 0,7$ dB
Time	$\pm 0,2$ %
Frequency	$\pm 0,2$ %

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

Table 5.3: 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	$\pm 0,4$ dB across the whole frequency range
Frequency generation	± 2 % (see note)
Time	$\pm 0,2$ %
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 [17] 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.

6 Codec independent requirements and associated Measurement Methodologies

6.1 Send and receive frequency response

6.1.1 Send frequency response

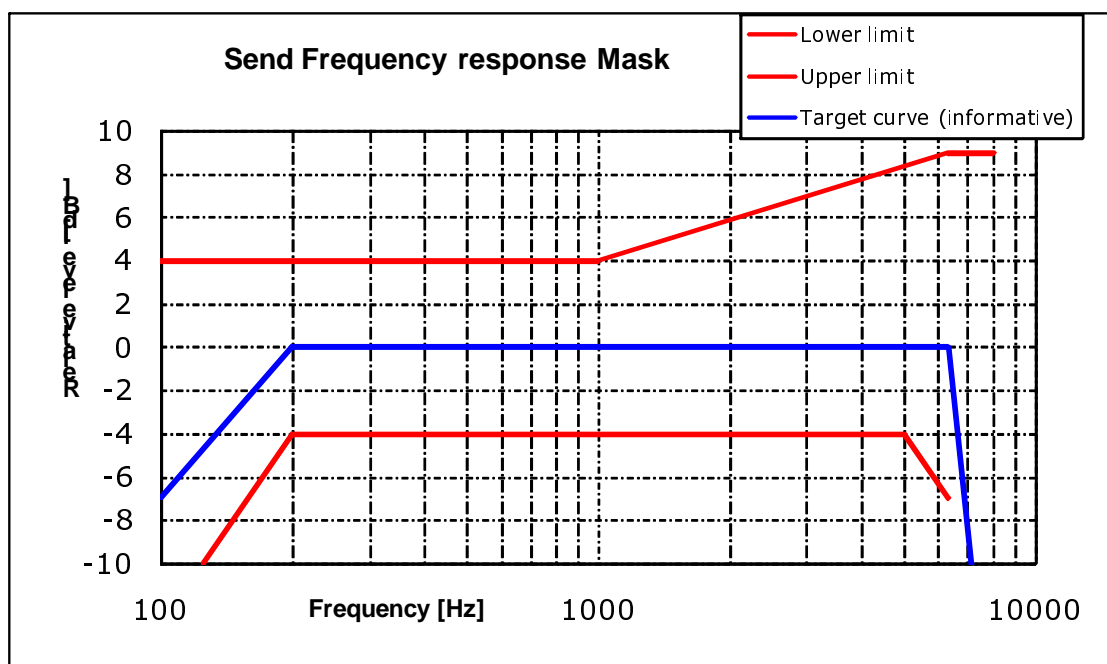
Requirement

The send sensitivity frequency response from the MRP to the measurement output (digital or analog output according measurement system used) shall within the mask which can be drawn with straight lines between the breaking points in table 6.1 on a logarithmic (frequency) - linear (dB sensitivity) scale.

Table 6.1: Hands-free send sensitivity/frequency response

Frequency	Upper limit	Lower limit
100 Hz	4 dB	$-\infty$ dB
125 Hz	4 dB	-10 dB
200 Hz	4 dB	-4 dB
1 000 Hz	4 dB	-4 dB
5 000 Hz	(see note)	-4 dB
6 300 Hz	9 dB	-7 dB
8 000 Hz	9 dB	$-\infty$ dB

NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (Hz) scale.

**Figure 6.1: Hands-free send sensitivity/frequency response****Measurement method**

The terminal will be positioned as described in clause 5.2.

An artificial voice according to ITU-T Recommendation P.50 [6] or a speech like test signal as described in ITU-T Recommendation P.501 [12] can be used for test. The type of test signal used shall be stated in the test report. 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.

The spectrum at the MRP and the actual level at the MRP (measured in third octaves) are used as reference to determine the send sensitivity S_{mJ} needed to compute SLR.

NOTE: For vehicle, to be reconsidered when the Focus Group Carcom has finalized the specification for wide band speech.

6.1.2 Receive frequency response

6.1.2.1 Handheld terminal

Requirement

The receive sensitivity frequency response from the measurement input (digital or analog 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 6.2 on a logarithmic (frequency) - linear (dB sensitivity) scale.

Table 6.2: Handheld terminal receive sensitivity/frequency response

Frequency	Upper limit	Lower limit
200 Hz	6 dB	$-\infty$ dB
250 Hz	6 dB	$-\infty$ dB
315 Hz	6 dB	$-\infty$ dB
400 Hz	6 dB	-12 dB
500 Hz	6 dB	-6 dB
4 000 Hz	6 dB	-6 dB
5 000 Hz	6 dB	-9 dB
6 300 Hz	6 dB	-12 dB
7 000 Hz	6 dB	$-\infty$ dB
8 000 Hz	6 dB	$-\infty$ dB

NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (Hz) scale.

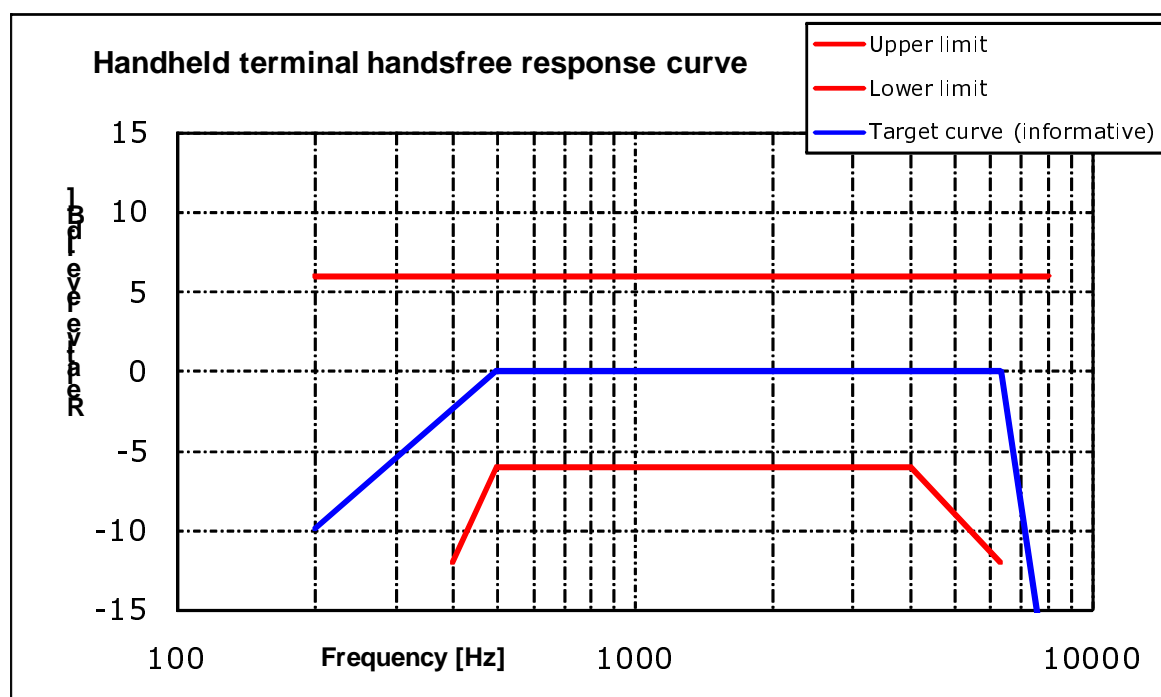


Figure 6.2: Handheld receive sensitivity/frequency response

Measurement method

The terminal will be positioned as described in clause 5.2.

The measurement is conducted at nominal volume control setting.

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

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

S_{Jeff}	Receive Sensitivity; Junction to HATS ear with free field correction.
p_{eff}	DRP Sound pressure measured by ear simulator Measurement data are converted from the Drum Reference Point to free field.
v_{RCV}	Equivalent RMS input voltage.

The test signal to be used for the measurements shall be the artificial voice according to ITU-T Recommendation P.50 [6]. If the signal to noise ratio in the high frequency domain is not sufficient CSS as defined in ITU-T Recommendation P.501 [12] shall be used. The test signal level shall be -16 dBm₀, measured according to ITU-T Recommendation P.56 [7] at the digital reference point or the equivalent analogue point.

HATS is free field equalized as described in ITU-T Recommendation P.581 [13]. The equalized output signal is power-averaged on the total time of analysis. The 1/12 octave band data are considered as the input signal to be used for calculations or measurements.

Measurements shall be made at one twelfth-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [14] for frequencies from 100 Hz to 10 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 Vehicle mounted hands-free

Requirement

Table 6.3: Vehicle mounted terminal receive sensitivity/frequency response

Frequency	Upper limit	Lower limit
125 Hz	8 dB	-∞ 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	-∞ dB

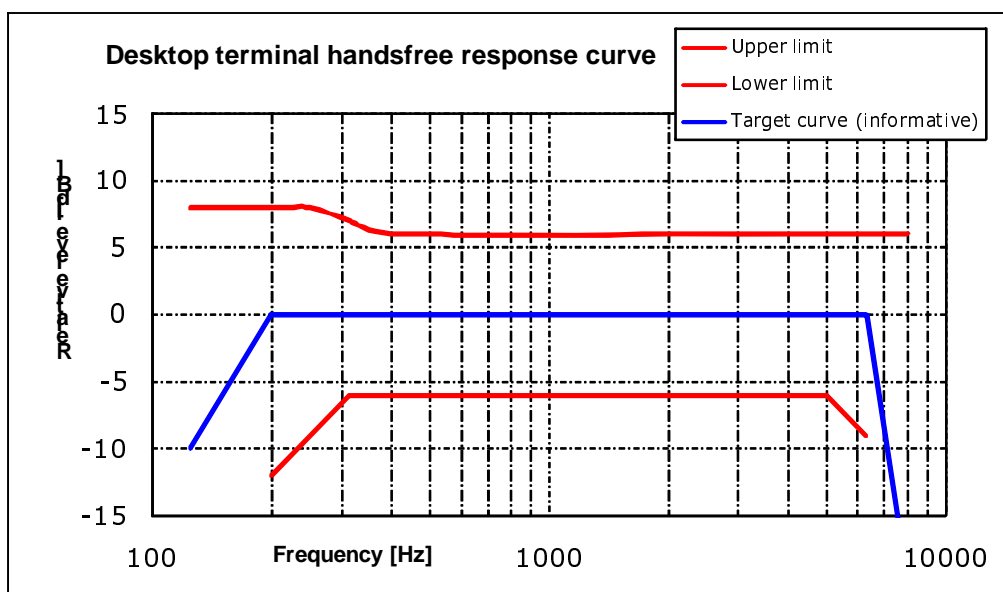


Figure 6.3: Vehicle mounted terminal receive sensitivity/frequency response

Measurement method

The terminal will be positioned as described in clause 5.2.

The test signal used for the measurements shall be artificial voice according to ITU-T Recommendation P.50 [6]. The test signal is -16 dBm₀, measured at the electrical reference point and averaged over the complete test signal sequence.

The test arrangement is according to clause 5.2. For the measurement of hands-free terminals the artificial head is free-field equalized according to ITU-T Recommendation P.581 [13]. The equalized output signal of the right ear is used for the measurement]. The receive sensitivity frequency response is determined in third octaves as given by the R.40 series of the preferred numbers in ISO 3 [14] for frequencies from 100 Hz to 10 kHz inclusive. In each third octave band the level of the measured signal is referred to the level of the reference signal, averaged over the complete test sequence length.

The sensitivity is determined in dBPa/V.

NOTE: Different listener position should be taken into account. Therefore the measurement should be repeated by moving the seat with the artificial head in different, typical positions.

6.1.2.3 Softphone (computer-based terminals)**Requirement**

Type 1 or softphone with external speakers: requirement defined in table 6.3 as for vehicle mounted or desktop terminal.

Type 2 requirement is defined in table 6.2 (as for hand-held terminal).

Measurement method

The terminal will be positioned as described in clause 5.2.

Measurement methods are defined in clause 6.1.2.1.

6.1.2.4 Desktop Terminal**Requirement**

Requirement according table 6.3 (as for vehicle mounted terminal)

Measurement method

The terminal will be positioned as described in clause 5.2.

Measurements methods are defined in clause 6.1.2.1.

6.2 Send and receive loudness ratings**6.2.1 Send Loudness Ratings****Requirement**

The nominal values of SLR shall be:

$$\text{SLR} = +13 \pm 3 \text{ dB}$$

Measurement method

The terminal will be positioned as described in clause 5.2.

An artificial voice according to ITU-Recommendation P. 50 [6] or a speech like test signal as described in ITU-T Recommendation P.501 [12] can be used to test. The type of test signal used shall be stated in the test report. The spectrum of acoustic signal produced by the artificial mouth is calibrated under free field conditions at the MRP. The test signal level shall be -4,7 dBPa, measured at the MRP. The test signal level is averaged over the complete test signal sequence.

Calibration is realized as explained in clause 5.2.1.

The send sensitivity shall be calculated from each band of the 20 frequencies given in table 1 of ITU-T Recommendation 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 ITU-T Recommendation P.79 [9], annex A.

6.2.2 Receive Loudness Ratings

6.2.2.1 Handheld terminal

Requirement

Nominal value of RLR will be $+9 \pm 3$ dB. This value has to be fulfilled for one position of volume range.

Value of RLR at upper part of volume range must be less than (louder) or equal to +2 dB: $RLR_{max} \leq +2$ dB.

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

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 artificial voice according to ITU-T Recommendation P.50 [6]. If the signal to noise ratio in the high frequency domain is not sufficient CSS as defined in ITU-T Recommendation P.501 [12] shall be used. The test signal level shall be -16 dBm0, measured according to ITU-T Recommendation P.56 [7] at the digital reference point or the equivalent analogue point.

The receive sensitivity shall be calculated from each band of the 20 frequencies given in table 1 of ITU-T Recommendation 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 ITU-T Recommendation P.79 [9], annex A. The RLR shall then be computed as $RLR(cal)$ minus 14 dB according to ITU-T Recommendation P.340 [10]), and without the LE factor.

6.2.2.2 Vehicle mounted hands-free

Requirement

$$RLR = +2 \text{ dB} \pm 4 \text{ dB}$$

If a user-specific volume control is provided, the requirement for RLR given above shall be measured at least for one setting of the volume control. It is recommended to provide a volume control which allows a loudness increase by at least 15 dB referred to the nominal value of RLR.

NOTE: To be reconsidered when the ITU-T Focus Group CARCOM has finalized the specification for wide band speech.

Measurement method

The terminal will be positioned as described in clause 5.2.

The test signal used for the measurements shall be artificial voice according to ITU-T Recommendation P.50 [6]. The test signal is -16 dBm0, measured at the electrical reference point and averaged over the complete test signal sequence.

For the measurement of hands-free terminals the artificial head is free-field equalized according to ITU-T Recommendation P.581 [13].

The sensitivity is expressed in terms of dBPa/V and the RLR shall be calculated according to ITU-T Recommendation P.79 [9], formula (A-23c), over bands 1 to 20, using $m = 0,175$ and the receive weighting factors from table A.2 of ITU-T Recommendation P.79 [9], Annex A.

For hands-free terminals the correction 14 dB according to ITU-T Recommendation P.340 [10] is used for the correction of the measurement results.

The test is repeated for maximum volume control setting.

6.2.2.3 Softphone (computer-based terminals)

Requirement

Type 1 or softphone with external speakers requirement defined in clause 6.2.2.2 as for vehicle mounted or desktop terminal

Type 2 requirement is defined in clause 6.2.2.1 as for handheld terminal

Measurement method

The terminal will be positioned as described in clause 5.2.

Measurement methods are defined in clause 6.2.2.1.

6.2.2.4 Desktop Terminal

Requirement

Nominal value of RLR will be $+5 \pm 3$ dB. This value has to be fulfilled for one position of volume range.

Value of RLR at upper part of volume range must be less than (louder) or equal to -2 dB: $RLR_{max} \leq -2$ dB.

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

Measurement method

The terminal will be positioned as described in clause 5.2.

Measurement methods are defined in clause 6.2.2.1.

6.3 Send and receive noise

6.3.1 Send Noise

Requirement

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 dependant of activity of system.

Measurement method

The terminal will be positioned as described in clause 5.2.

For a correct activation of the system, an artificial voice according to ITU-T Recommendation P. 50 [6] or a speech like test signal as described in ITU-T Recommendation P.501 [12] shall be used for activation. Level of this activation signal shall be -4,7 dBPa at the MRP.

The level at the output of the test setup is measured with a A filtering.

6.3.1 Receive Noise

Requirement

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

NOTE 1: No peaks in the frequency domain higher than 10 dB above the average noise spectrum should occur.

NOTE 2: For softphone fan noise must be avoided in order to fulfil this condition.

Measurement method

The terminal will be positioned as described in clause 5.2.

A signal is applied to input of test system in order to ensure correct activation of receive state. An artificial voice according to ITU-Recommendation P. 50 [6] or a speech like test signal as described in ITU-T Recommendation P.501 [12] can be used for activation. Level of this activation signal will be -16 dBm0.

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

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

It is not intended to provide coder-dependant requirements but to assess the electroacoustic 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: 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.

The signal used is an activation signal followed by a series sine-wave signal with a frequency at 200 Hz, 315 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz, 1 000 Hz and 2 kHz. The duration of the sine-wave shall be of less than 1 s. The sinusoidal signal level shall be calibrated to -4,7 dBPa at the MRP.

The signal to harmonic distortion ratio is measured selectively up to 6,3 kHz.

An artificial voice according to ITU-Recommendation P.50 [6] or a speech like test signal as described in ITU-T Recommendation P.501 [12] can be used for activation. Level of this activation signal will 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

- **vehicle mounted terminal**

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

Table 6.5: 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.

- **Handheld terminal**

The ratio of signal to harmonic distortion is given in table 6.5.

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

Type 1 or softphone with external speakers: requirement given in table 6.5 as for vehicle mounted or desktop terminal.

Type 2 requirement given in table 6.5 as for handheld terminal.

Measurement method

Test setup is described in clause 5.2.

The signal used is an activation signal followed by a sine-wave signal with a frequency at 315 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz, 1 000 Hz, 2 000 Hz, 3 000 Hz. The duration of the sine-wave shall be of less than 1 s. Appropriate signals for activation and signal combinations can be found in ITU-T Recommendation P.501 [12]. The sinusoidal signal level shall be calibrated to -16 dBm0.

An artificial voice according to ITU-Recommendation P.50 [6] or a speech like test signal as described in ITU-T Recommendation P.501 [12] can be used for activation. Level of this activation signal will be -16 dBm0.

The signal to harmonic distortion ratio is measured selectively up to 15 kHz.

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

6.5 TCLw (or similar parameters)

6.5.1 Handheld Terminal

Requirement

In order to meet the ITU-T Recommendation G.131 [5] talker echo objective requirements, the recommended weighted terminal coupling loss during single talk (TCL_{wst}) should be greater than 55 dB when measured under free field conditions at **nominal setting of the volume control**.

A TCL_w greater than 46 dB is considered as acceptable.

For terminals fitted with a volume control the TCL_w shall be not less than 40 dB for the higher gain settings above the nominal setting of the volume control.

Measurement method

The setup for terminal is described in clause 5.2.

For hands-free measurement, HATS is positioned but not used.

For loudspeaking measurement, handset is positioned on HATS (right ear).

Before the actual test a training sequence consisting of 10 s artificial voice male and 10 s artificial voice female according to ITU-T Recommendation P.50 [6] is altered. The training sequence level shall be -16 dBm₀ in order to not overload the codec.

The test signal following immediately the training sequence is a PN-sequence complying with ITU-T Recommendation P.501 [12] with a length of 4 096 points (for the 48 kHz sampling rate) and a crest factor of 6 dB. The length of the complete test signal composed of at least four sequences of CSS shall be at least one second (1,0 s). The test signal level is -3 dBm₀ (from 50 Hz to 7 kHz). The low crest factor is achieved by random alternation of the phase between -180° and 180°.

The TCL_w is calculated according to ITU-T Recommendation G.122 [2], clause B.4 (trapezoidal rule), but using the frequency range of 300 Hz to 6 700 Hz instead of 300 Hz to 3 400 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. For the measurement a time window has to be applied adapted to the duration of the actual pn-sequence of the test signal (200 ms) choosing the pn-sequence of the third CS-.Signal.

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

6.5.2 Vehicle mounted hands-free

Requirement

The TCL_w in quiet environments should be at least 50 dB for nominal setting of the volume control. For maximum setting of the volume control TCL_w should be higher than 50 dB. The implemented echo control mechanism should provide a sufficient echo loss for all typical environments and typical impulse responses.

When conducting the tests it should be checked whether the signal measured is an echo signal and not comfort noise inserted in send direction in order to mask an echo signal or noise emitted by the loudspeakers. This could be checked e.g. by conducting the idle channel noise measurement with maximum volume control setting.

NOTE: There may be implementations where echo problems may be observed although the TCL_w test gives a high number. In such cases it is recommended to verify the echo performance by subjective tests including different situations which are not addressed in this test.

Measurement method

All tests are conducted in the car cabin, the test arrangement is described in clause 5.2.2. The noise level measured at the electrical access point (idle channel noise) shall be less than -63 dBm₀. The attenuation between the input of the electrical reference point to the output of the electrical reference point is measured using a speech-like test signal.

Before the actual measurement a training sequence consisting of 10 s of artificial voice (male) and 10 s of artificial voice (female) according to ITU-T Recommendation P.50 [6] is inserted. The training sequence level shall be -16 dBm0.

The test signal is a pn sequence according to ITU-T Recommendation P.501 [12] with a length of 4 096 points (48 kHz sampling rate) and a crest factor of 6 dB. The duration of the test signal is 250 ms, the test signal level is -3 dBm0. The low crest factor is achieved by random alternation of the phase between -180° and +180°.

TCLW is calculated according to ITU-T Recommendation G.122 [2], annex B, clause B.4 (trapezoidal pseudo rule) but using the frequency range of 300 Hz to 6 700 Hz instead of 300 Hz to 3 400 Hz. For the calculation the average measured echo level at each frequency band is referred to the average level of the test signal measured in each frequency band. For the measurement a time window has to be applied which is adapted to the duration of the actual test signal (250 ms).

6.5.3 Softphone (computer-based terminals)

Requirement

Type 1 or softphone with external speakers: requirement defined in clause 6.6.2 as for vehicle mounted or desktop terminal.

Type 2 requirement as in clause 6.6.1 for hand-held terminal (as for hand-held terminal).

Measurement method

The terminal will be positioned as described in clause 5.2.

Measurement method are defined in clause 6.6.1 (as for hand-held terminal).

6.5.4 Desktop Terminal

Requirement

Same requirement as in clause 6.6.1 (as for handheld terminal).

Measurement method

The terminal will be positioned as described in clause 5.2.

Measurement method are defined in clause 6.6.1 (as for handheld terminal).

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

Measurement method

For handsfree mode test set-up is identical as for TCLw.

Before the actual test a training sequence consisting of 10 s artificial voice male and 10 s artificial voice female according to ITU-T Recommendation P.50 [6] is altered. The training sequence level shall be -16 dBm0 in order not to overload the codec.

The test signal is a PN sequence complying with ITU-T Recommendation 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

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 ITU-T Recommendations P.340 [10] and P.502 [i.4]):

- 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

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 6.6.

Table 6.6

Category (according to ITU-T Rec. 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

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 category of the terminal according to table 6.6 shall be noted in the test report.

Measurement method

The test setup is described in clause 5.2. The test signal to determine the attenuation range during double talk is shown in figure 6.4. A sequence of uncorrelated CS signals is used which is inserted in parallel in send and receive direction.

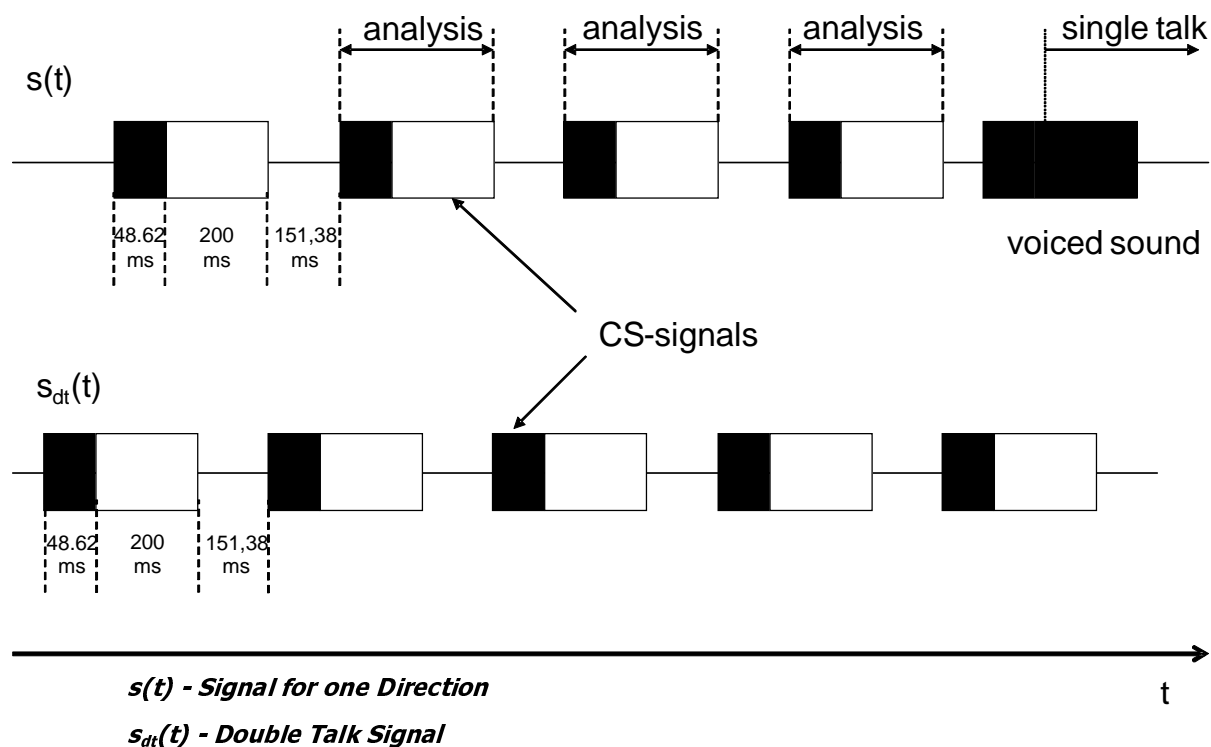


Figure 6.4: Double Talk Test Sequence with overlapping CS signals in send and receive direction

Figure 6.4 indicates that the sequences overlap partially. The beginning of the CS sequence (voiced sound, black) is overlapped by the end of the pn-sequence (white) of the opposite direction. During the active signal parts of one signal the analysis can be conducted in send and receive direction. The analysis times are shown in figure 6.4 as well. The test signals are synchronized in time at the acoustical interface. The delay of the test arrangement should be constant during the measurement.

NOTE: The length of voiced sound of the double talk signal is achieved by repeating one period of the voiced sound for double talk according to ITU-T Recommendation P.501 [12] 10 times and cutting off the initial 3,3 ms of the period of the first voiced sound.

The settings for the test signals are as follows.

Table 6.7

	Receive Direction (sdt(t))	Send Direction (s(t))
Pause Length between two Signal Bursts	151,38 ms	151,38 ms
Average Signal Level (Assuming an Original Pause length of 101,38 ms)	-16 dBm0	-4,7 dBPa
Active Signal Parts	-14,7 dBm0	-3 dBPa
NOTE: When the test laboratories implement different values (within the accuracy range defined in clause 5.7) it should be indicated in the test report.		

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

The level is determined as level vs. time from the time domain. The integration time of the level analysis is 5 ms. The attenuation is determined from the level difference measured at the beginning of the double talk always with the beginning of the CS-signal in send direction until its complete activation (during the pause in the receive channel). The analysis is performed over the complete signal starting with the second CS-signal. The first CS-signal is not used for the analysis.

6.7.2 Attenuation Range in Receive Direction during Double Talk

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 6.8.

Table 6.8

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

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 category of the terminal according to table 6.8 shall be noted in the test report.

Measurement method

The test setup is described in clause 5.2.

The test signal to determine the attenuation range during double talk is shown in figure 6.4. A sequence of uncorrelated CS signals is 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 should be constant during the measurement.

The settings for the test signals are as follows.

Table 6.9

	Receive Direction (s(t))	Send Direction (sdt(t))
Pause Length between two Signal Bursts	151,38 ms	151,38 ms
Average Signal Level (Assuming an Original pause Length of 101,38 ms)	-16 dBm0	-4,7 dBPa
Active Signal Parts	-14,7 dBm0	-3 dBPa
NOTE:	When the test laboratories implement different values (within the accuracy range defined in clause 5.7) it should be indicated in the test report.	

When determining the attenuation range in receive direction the signal measured at the artificial ear referred to the test signal inserted.

The level is determined as level vs. time from the time domain. The integration time of the level analysis is 5 ms. The attenuation is determined from the level difference measured at the beginning of the double talk always with the beginning of the CS-signal in receive direction until its complete activation (during the pause in the send channel). The analysis is performed over the complete signal starting with the second CS-signal. The first CS-signal is not used for the analysis.

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 the table below are applicable (more information can be found in annex A of the ITU-T Recommendation P.340 [10]).

Table 6.10

Category (according to ITU-T Rec. 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).

Measurement method

The test setup is 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. The measurement signals used are shown in the figure below. A detailed description can be found in ITU-T Recommendation 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).

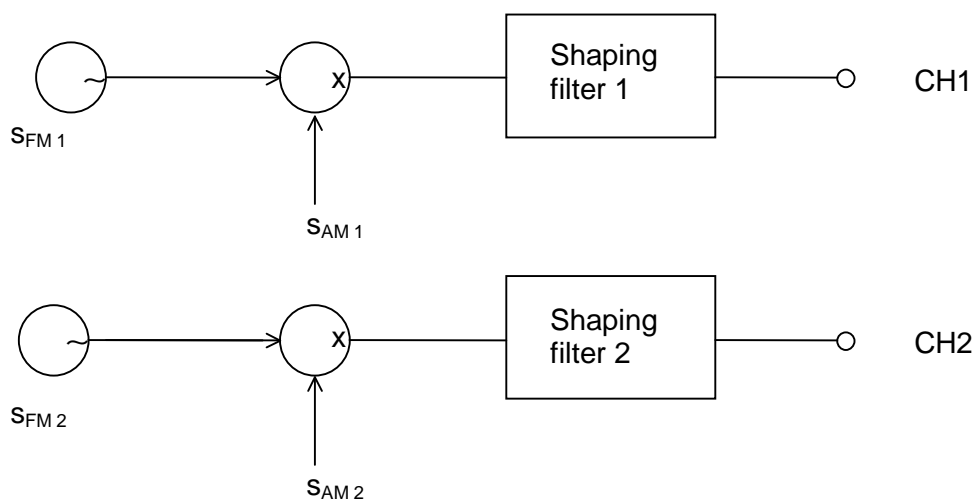


Figure 6.5: Measurement signals

$$s_{FM1,2}(t) = \sum A_{FM1,2} * \cos(2\pi n * F_{01,2}) ; n= 1,2,\dots \quad (2)$$

$$s_{AM1,2}(t) = A_{AM1,2} * \cos(2\pi t F_{AM1,2}); \quad (3)$$

The settings for the signals are as follows.

Table 6.11: Parameters of the two Test Signals for Double Talk Measurement based on AM-FM modulated sine waves

Receive Direction			Send Direction			
f_m [Hz]	$f_{mod(fm)}$ [Hz]	F_{am} [Hz]		f_m [Hz]	$f_{mod(fm)}$ [Hz]	F_{am} [Hz]
125	$\pm 2,5$	3		150	$\pm 2,5$	3
250	± 5	3		270	± 5	3
500	± 10	3		540	± 10	3
750	± 15	3		810	± 15	3
1 000	± 20	3		1 080	± 20	3
1 250	± 25	3		1 350	± 25	3
1 500	± 30	3		1 620	± 30	3
1 750	± 35	3		1 890	± 35	3
2 000	± 40	3		2 160	± 35	3
2 250	± 40	3		2 400	± 35	3
2 500	± 40	3		2 650	± 35	3
2 750	± 40	3		2 900	± 35	3
3 000	± 40	3		3 150	± 35	3
3 250	± 40	3		3 400	± 35	3
3 500	± 40	3		3 650	± 35	3
3 750	± 40	3		3 900	± 35	3
4 000	± 40	3		4 150	± 35	3
4 250	± 40	3		4 400	± 35	3
4 500	± 40	3		4 650	± 35	3
4 750	± 40	3		4 900	± 35	3
5 000	± 40	3		5 150	± 35	3
5 250	± 40	3		5 400	± 35	3
5 500	± 40	3		5 650	± 35	3
5 750	± 40	3		5 900	± 35	3
6 000	± 40	3		6 150	± 35	3
6 250	± 40	3		6 400	± 35	3
6 500	± 40	3		6 650	± 35	3
6 750	± 40	3		6 900	± 35	3
7 000	± 40	3				

NOTE: Parameters of the Shaping Filter:
 $f \geq 250$ Hz: Low Pass Filter, 5 dB/oct; $f < 250$ Hz,: High Pass Filter

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 ITU-T Recommendation 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 the table above. 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

NOTE: 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).

Requirement

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

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

Measurement method

The test setup is described in clause 6.1.

The structure of the test signal is shown in figure 18. The test signal consists of CSS components according to ITU-T Recommendation P.501 [12] with increasing level for each CSS burst.

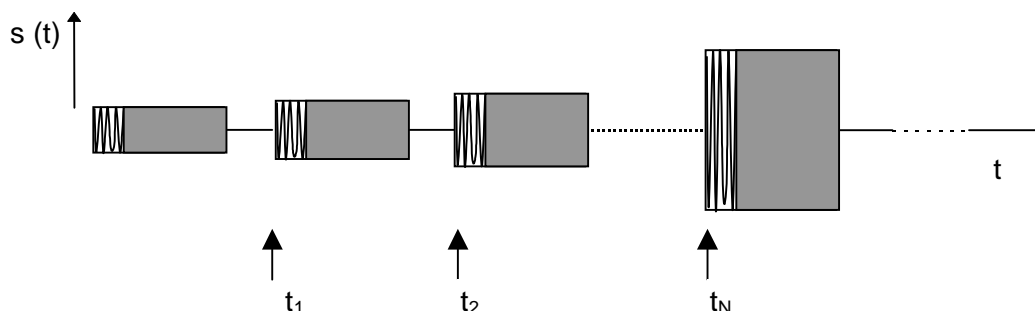


Figure 6.6: Test Signal to Determine the Minimum Activation Level and the Built-up Time

The settings of the test signal are as follows.

Table 6.12: 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
CSS to Determine Switching Characteristic in Send Direction	~250 ms / ~450 ms	-23 dBPa (see note)	1 dB
NOTE: The level of the active signal part corresponds to an average level of -24,7 dBPa at the MRP for the CSS according to ITU-T Recommendation P.501 [12] assuming a pause of about 100 ms.			

It is assumed that the pause length of about 450 ms is longer than the hang-over time so that the test object is back to idle mode after each CSS burst.

The level of the transmitted signal is measured at the electrical reference point. The measured signal level is referred to the 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 CSS burst which indicates the first activation of the test object. The time between the beginning of the CSS burst and the complete activation of the test object is measured.

NOTE: If the measurement using the CS-Signal does not allow to clearly identify the minimum activation level, the measurement may be repeated by using a one syllable word instead of the CS-Signal. The word used should be of similar duration, the average level of the word should be adapted to the CS-signal level of the according CS-burst.

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 6.13.

Table 6.13: 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 test setup is described in clause 5.2.

The background noise simulation as described in clause 5.5.3 (5.2.2 for vehicle mounted terminal) 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 Composite Source Signal 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. Alternatively other speech like test signals (e.g. artificial voice) with the same signal level can be used.

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 analyzed and compared to the requirements.

6.9.2 Speech Quality in the Presence of Background Noise

Requirement

Speech Quality for wideband systems can be tested based on EG 202 396-3 [i.8]. The test method described 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 $\geq 3,0$.
- S-MOS-LQOw $\geq 3,0$.
- G-MOS-LQOw $\geq 3,0$.

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 (clause 5.2.2 for vehicle mounted terminal) is used. The handset terminal is set-up as described in clause 5.2.

The background noise should 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 8 sentences of speech (2 male and 2 female talkers, 2 sentences each). Appropriate speech samples can be found in ITU-T Recommendation P.501 [12]. The preferred language is French since the objective method was validated with French language. The test signal level is -4,7 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 EG 202 396-3 [i.8].

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

Requirement

The test is carried out applying the Composite Source Signal in receive direction. During and after the end of Composite Source Signal bursts (representing the end of far end speech simulation) 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 6.1.

Measurement method

The test setup is described in clause 5.2.

The background noises are generated as described in clause 5.5.3 (clause 5.2.2 for vehicle mounted terminal).

First the 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 CS-signal at the far end. The exactly identical background noise signal is applied. The background noise signal must start at the same point in time which was used for the measurement without far end signal. The background noise should be applied for at least 10 s in order to allow adaptation of the noise reduction algorithms and should be mixed speech like signal e.g. CSS. After at least 10 s a Composite Source Signal according to ITU-T Recommendation P.501 [12] is applied in receive direction with a duration of ≥ 2 CSS periods. The test signal level is $-16 \text{ dB}_{\text{m0}}$ at the electrical reference point.

The send signal is recorded at the electrical reference point. 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 CS-signal is applied and after it stops. The level difference is determined from the difference of the recorded signal levels vs. time between reference signal and the signal measured with far end signal.

6.9.4 Quality of Background Noise Transmission (with Near End Speech)

Requirement

The test is carried out applying a simulated speech signal in send direction. During and after the end of the simulated speech signal (Composite Source Signal bursts) the signal level in send direction should not vary more than 10 dB.

Measurement method

The test setup is described in clause 5.2.

The background noises are generated as described in clause 5.5.3 (clause 5.2.2 for vehicle mounted terminal). The background noise should be applied for at least 5 s in order to allow adaptation of the noise reduction algorithms.

The near end speech is simulated using the Composite Source Signal according to ITU-T Recommendation P.501 [12] with a duration of ≥ 2 CSS periods. The test signal level is $-4,7 \text{ dBPa}$ at the MRP.

The send signal is recorded at the electrical reference point. The test signal level versus time is calculated using a time constant of 35 ms.

First the measurement is conducted without inserting the signal at the near end. The signal level is analyzed vs. time. In a second step the same measurement is conducted but with inserting the CS-signal at the near end. The level variation is determined by the difference between the background noise signal level without inserting the CS-signal and the maximum level of the noise signal during and after the CS-bursts in send direction.

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 should not decrease by more than 6 dB from the maximum measured during the TCLw test.

Measurement method

The test setup is described in clause 5.1.

The test signal consists of periodically repeated Composite Source Signal according to ITU-T Recommendation P.501 [12] with an average level of -5 dBm0 as well as an average level of -25 dBm0. The echo signal is analyzed 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.

NOTE 1: In addition tests with more speech like signals should be made, e.g. ITU-T Recommendation P.50 [6] 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 analyzed. The analysis time is reduced by the integration time of the level analysis (35 ms).

6.10.2 Spectral Echo Attenuation

Requirement

The echo attenuation vs. frequency shall be below the tolerance mask given in table 6.14.

Table 6.14: 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 10 s CS signal according to ITU-T Recommendation P.501 [12]. The level of the training sequence is -16 dBm0.

The test signal consists of a periodically repeated Composite Source Signal. The measurement is carried out under steady-state conditions. The average test signal level is -16 dBm0, averaged over the complete test signal. 4 CS signals including the pauses are used for the measurement which results in a test sequence length of 1,4 s. 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 analyzed in the frequency domain in dB.

6.10.3 Occurrence of Artefacts

For further study.

6.11 Send and receive delay or 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 recognised that the end to end delay should be as small as possible in order to ensure high quality of the communication.

The delay T_{rtt} in send direction T_s plus the delay in receive direction T_r shall be less than 70 ms if the hands-free system is implemented in conjunction with the speech coder and the RF-transmission. If the hands-free system is connected via additional radio link the delay in send direction T_s plus the delay in receive direction T_r shall be less than 70 ms plus the delay of the radio link and in case of Bluetooth link 120 ms.

NOTE 1: Those limits are based on the assumption that the mobile phone signal processing is deactivated and does not introduce any additional processing delay.

NOTE 2: Half of the round trip delay corresponds to the mean one-way delay.

As the actual delay depends on the codec implementations, complementary requirements and test methods are defined in clause 7.

Measurement method:

- 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:

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

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.

NOTE 2: The delay requirement assumes a delay of maximum 8 ms inserted by a potential additional radio link. Therefore tests should be made with a Bluetooth mobile phone which introduces a low delay.

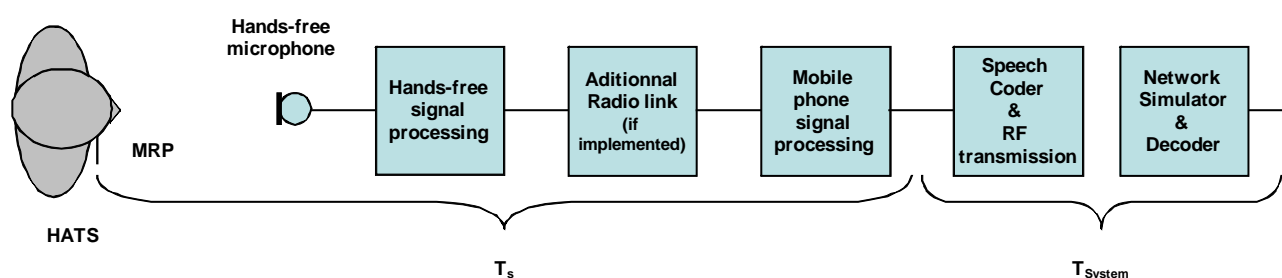


Figure 6.7: Different blocks contributing to the delay in send direction

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

- 1) For the measurements a Composite Source Signal (CSS) according to ITU-T Recommendation 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 ITU-T Recommendation P.581 [13]). 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 cross-correlation function is used for the determination.
- Receive direction

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

$$T_r + t_{\text{system}}$$

NOTE 3: 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.

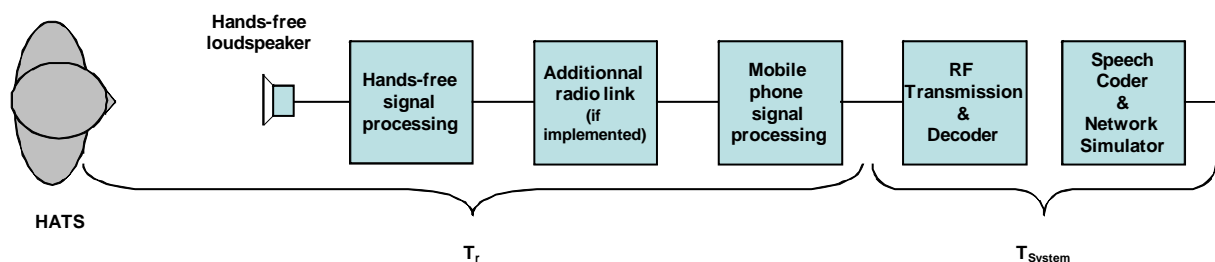


Figure 6.8: Different blocks contributing to the delay in receive direction

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

- 1) For the measurements a Composite Source Signal (CSS) according to ITU-T Recommendation 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).

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

- 2) The test arrangement is according to clause 5.2. Artificial head is free-field according to ITU-T Recommendation P.581 [13]. 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 cross-correlation function is used for the determination.

6.12 Objective listening Quality in send and receive direction

The aim is to provide the best listening quality whatever the implementation is.

Provisionnal target value: $MOS-LQO_M > 4,0$.

As the actual listening quality depends on the codec implementation, specific requirements and test methods are defined in clause 7.

This clause will be updated when the relevant quality model will be available.

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 defines a list of speech coders implemented (non exhaustive).

Table 7.1: List of speech coders

System	Codec
GSM 850, 900, 1 800, 1 900	AMR-WB (ITU-T G.722.2) @12,65 kbit/s [16]
UMTS (WCDMA)	AMR-WB (ITU-T G.722.2) @12,65 kbit/s [16]
Voice over Data Network (VoDN)	ITU-T G.722 [3] ITU-T G.722.2 [16] ITU-T G.729.1 [4] ITU-T G.711.1 [15]

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 Send and receive delay or round trip delay

To be completed in the next version of the present document.

7.3 Objective listening Quality in send and receive direction

For further study (this clause will be updated when the relevant quality model will be available).

8 Requirements and associated Measurement Methodologies (with an additional radio link between the terminal and external electroacoustical devices)

The intention is to provide requirements and test methods for the complete chain.

Annex A (informative): Bibliography

- 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 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 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".

History

Document history		
V1.1.1	November 2009	Publication