International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.1100

(10/2008)

SERIES P: TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS

Communications between cars

Narrow-band hands-free communication in motor vehicles

Recommendation ITU-T P.1100

ITU-T

ITU-T P-SERIES RECOMMENDATIONS

TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	Series	P.10
Subscribers' lines and sets	Series	P.30
		P.300
Transmission standards	Series	P.40
Objective measuring apparatus	Series	P.50
		P.500
Objective electro-acoustical measurements	Series	P.60
Measurements related to speech loudness	Series	P.70
Methods for objective and subjective assessment of quality	Series	P.80
		P.800
Audiovisual quality in multimedia services	Series	P.900
Transmission performance and QoS aspects of IP end-points	Series	P.1000
Communications between cars	**Series**	**P.1100**

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T P.1100

Narrow-band hands-free communication in motor vehicles

Summary

Recommendation ITU-T P.1100 describes performance requirements and test methods for narrow-band hands-free communication in vehicles. This Recommendation covers:

– build-in hands-free systems;

– after-market hands-free car kits;

– corded headsets; and

– wireless headsets;

to be used in vehicles for communication while driving.

This Recommendation addresses the test of complete systems as well as the subsystems of hands-free microphone and telephone with short-range wireless transmission link used to interconnect the hands-free system to the mobile network.

For testing, the test set-up and the recommended environmental conditions are described.

The methods, the analysis and the performance parameters described in this Recommendation are based on test signals and test procedures as defined in Recommendations ITU-T P.50, ITU-T P.501, ITU-T P.502, ITU-T P.340 and ITU-T P.380.

Source

Recommendation ITU-T P.1100 was approved on 22 October 2008 by ITU-T Study Group 12 (2005-2008) under Recommendation ITU-T A.8 procedure.

Keywords

Hands-free, headset, motor vehicle, quality of service, QoS.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

CONTENTS

Page

Recommendation ITU-T P.1100

Narrow-band hands-free communication in motor vehicles

1 Scope

The aim of this Recommendation is to define use cases and test methods for narrow-band hands-free communication in vehicles. This Recommendation covers:

– build-in hands-free systems;

– after-market hands-free car kits;

– corded headsets; and

– wireless headsets;

to be used in vehicles for communication while driving.

This Recommendation addresses the test of complete systems as well as the subsystems of hands-free microphone and telephone with short-range wireless transmission link used to interconnect the hands-free system to the mobile network.

For testing, the test setup and the recommended environmental conditions are described.

The methods, the analysis and the performance parameters described in this Recommendation are based on test signals and test procedures as defined in [ITU-T P.50], [ITU-T P.501], [ITU-T P.502], [ITU-T P.340] and [ITU-T P.380].

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T G.100.1] Recommendation ITU-T G.100.1 (2001), *The use of the decibel and of relative levels in speechband telecommunications*.
<http://www.itu.int/rec/T-REC-G.100.1>

[ITU-T G.111] Recommendation ITU-T G.111 (1993), *Loudness ratings (LRs) in an international connection*.
<http://www.itu.int/rec/T-REC-G.111>

[ITU-T G.122] Recommendation ITU-T G.122 (in force), *Influence of national systems on stability and talker echo in international connections*.
<http://www.itu.int/rec/T-REC-G.122>

[ITU-T G.711] Recommendation ITU-T G.711 (in force), *Pulse code modulation (PCM) of voice frequencies*.
<http://www.itu.int/rec/T-REC-G.711>

[ITU-T P.48] Recommendation ITU-T P.48 (in force), *Specification for an intermediate reference system*.
<http://www.itu.int/rec/T-REC-P.48>

[ITU-T P.50] Recommendation ITU-T P.50 (1993), *Artificial voices*.
<http://www.itu.int/rec/T-REC-P.50>

[ITU-T P.56] Recommendation ITU-T P.56 (in force), *Objective measurement of active speech level*.
<http://www.itu.int/rec/T-REC-P.56>

[ITU-T P.57] Recommendation ITU-T P.57 (in force), *Artificial ears*.
<http://www.itu.int/rec/T-REC-P.57>

[ITU-T P.58] Recommendation ITU-T P.58 (in force), *Head and torso simulator for telephonometry*.
<http://www.itu.int/rec/T-REC-P.58>

[ITU-T P.64] Recommendation ITU-T P.64 (in force), *Determination of sensitivity/frequency characteristics of local telephone systems*.
<http://www.itu.int/rec/T-REC-P.64>

[ITU-T P.79] Recommendation ITU-T P.79 (in force), *Calculation of loudness ratings for telephone sets*.
<http://www.itu.int/rec/T-REC-P.79>

[ITU-T P.340] Recommendation ITU-T P.340 (in force), *Transmission characteristics and speech quality parameters of hands-free terminals*.
<http://www.itu.int/rec/T-REC-P.340>

[ITU-T P.380] Recommendation ITU-T P.380 (in force), *Electro-acoustic measurements on headsets*.
<http://www.itu.int/rec/T-REC-P.380>

[ITU-T P.501] Recommendation ITU-T P.501 (in force), *Test signals for use in telephonometry*.
<http://www.itu.int/rec/T-REC-P.501>

[ITU-T P.502] Recommendation ITU-T P.502 (in force), *Objective test methods for speech communication systems using complex test signals*.
<http://www.itu.int/rec/T-REC-P.502>

[ITU-T P.581] Recommendation ITU-T P.581 (2000), *Use of head and torso simulator (HATS) for hands-free terminal testing*.
<http://www.itu.int/rec/T-REC-P.581>

[ITU-T P.800] Recommendation ITU-T P.800 (in force), *Methods for subjective determination of transmission quality*.
<http://www.itu.int/rec/T-REC-P.800>

[ITU-T P.800.1] Recommendation ITU-T P.800.1 (in force), *Mean Opinion Score (MOS) terminology*.
<http://www.itu.int/rec/T-REC-P.800.1>

[ITU-T P.862] Recommendation ITU-T P.862 (in force), *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*.
<http://www.itu.int/rec/T-REC-P.862>

[ITU-T P.862.1] Recommendation ITU-T P.862.1 (in force), *Mapping function for transforming P.862 raw result scores to MOS-LQO*.
<http://www.itu.int/rec/T-REC-P.862.1>

[IEC 60268-4] IEC 60268-4 (2004), *Sound system equipment – Part 4: Microphones*.
<http://webstore.iec.ch/webstore/webstore.nsf/artnum/031724>

[IEC 61260]	IEC 61260 (1995), *Electroacoustics – Octave-band and fractional-octave-band filters*. <http://webstore.iec.ch/webstore/webstore.nsf/artnum/019426>
[ISO 3]	ISO 3:1973, *Preferred numbers – Series of preferred numbers*. <http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=3564>

3 Definitions

This Recommendation defines the following terms:

3.1 artificial ear: Device 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.

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

3.3 composite source signal (CSS): Signal composed in time by various signal elements.

3.4 D-value: D-value is computed directly from measurements of the difference Δ_{Sm} between the send sensitivities for diffuse and direct sound, S_{si} (diff) and S_{si} (direct), respectively.

$$\Delta_{Sm} = S_{si}\ (\text{diff}) - S_{si}\ (\text{direct})$$

D is computed as a weighted average of Δ_{Sm}

3.5 ear-drum reference point (DRP): Point located at the end of the ear canal, corresponding to the ear-drum position.

3.6 free-field equalization: The transfer characteristics of the 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. This equalization is specific to the HATS used.

3.7 free-field reference point: Point located in the free sound field, at least in 1.5 m distance from a sound source radiating in free air (in case of a head and torso simulator (HATS) in the centre of the artificial head with no artificial head present).

3.8 hands-free reference point (HFRP): A point located on the axis of the artificial mouth, at 50 cm from the outer plane of the lip ring, where the level calibration is made, under free-field conditions. It corresponds to the measurement point 11, as defined in [ITU-T P.51].

3.9 hands-free terminal: Telephone set that does not require the use of hands during the communications session; examples are headset, speakerphone and group-audio terminal.

3.10 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.

3.11 headset: Device which includes a telephone receiver and transmitter which is typically secured to the head or the ear of the wearer.

3.12 maximum setting of the volume control: When a receive volume control is provided, the maximum setting of the volume control is chosen.

NOTE – The maximum volume should be carefully chosen in order to provide sufficient loudness for typical driving conditions but not to overload the audio system and introduce non-linearities in the echo path.

3.13 mean opinion score – listening-only quality objective (MOS-LQO): The score is calculated by means of an objective model which aims at predicting the quality for a listening-only test situation. Objective measurements made using the model given in [ITU-T P.862] give results in terms of MOS-LQO (for further information see Annex A).

3.14 mean opinion score – talking-only quality objective (MOS-TQO): The score is calculated by means of an objective model which aims at predicting the quality for a talking-only test situation. Methods generating a MOS-TQO are currently under development and are not yet standardized.

3.15 mouth reference point (MRP): The month reference point is located on the axis and 25 mm in front of the lip plane of a mouth simulator.

3.16 nominal setting of the volume control: When a receive volume control is provided, the setting which is closest to the nominal receiving loudness rating of 2 dB.

3.17 receiving loudness rating (RLR): The loudness loss between an electric interface in the network and the listening subscriber's ear (the loudness loss is here defined as the weighted (dB) average of driving electromotive force to measured sound pressure).

3.18 sending loudness rating (SLR): The loudness loss between the speaking subscriber's mouth and an electrical interface in the network (the loudness loss is here defined as the weighted (dB) average of driving sound pressure to measured voltage).

3.19 wideband speech: Voice service with enhanced quality compared to PCM (see [ITU-T G.711]) and allowing the transmission of a vocal frequency range of at least 150 Hz to 7 kHz.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

ACR	Absolute Category Rating
A/D	Analogue/Digital
AGC	Automatic Gain Control
$A_{H,R}$	Attenuation Range in receiving direction
$A_{H,R,dt}$	Attenuation Range in receiving direction during double talk
$A_{H,S}$	Attenuation Range in sending direction
$A_{H,S,dt}$	Attenuation Range in sending direction during double talk
BGN	BackGround Noise
CSS	Composite Source Signal
D/A	Digital/Analogue
DI	Digital Interface
DRP	ear-Drum Reference Point
DTX	Discontinuous Transmission
EC	Echo Cancellation
ERL	Echo Return Loss
ERP	Ear Reference Point
FFT	Fast Fourier Transform
HATS	Head And Torso Simulator
HATS-HFRP	Head And Torso Simulator – Hands-Free Reference Point

HF	Hands-Free
HFT	Hands-Free Terminal
JLR	Junction Loudness Rating
$L_{R,min}$	minimum activation level (Receiving direction)
$L_{S,min}$	minimum activation level (Sending direction)
MOS	Mean Opinion Score
MRP	Mouth Reference Point
NR	Noise Reduction
PCM	Pulse Code Modulation
POI	Point Of Interconnection
QoS	Quality of Service
RF	Radio Frequency
RLR	Receiving Loudness Rating
SLR	Sending Loudness Rating
S_{si}(diff)	Diffuse field Sensitivity
S_{si}(direct)	Direct sound Sensitivity
SRW	Short Range Wireless interface (details are under study; see: http://www.itu.int/ITU-T/studygroups/com12/fgfit/index.html)
SRWR	Short Range Wireless transmission Reference point
S/N	Signal-to-Noise ratio
TCLw	weighted Terminal Coupling Loss
TEMS	Terrestrial Ecosystem Monitoring Site
$T_{r,R}$	built-up time (Receiving direction)
$T_{r,S}$	built-up time (Sending direction)

5 Conventions

dBm:	Absolute power level relative to 1 milliwatt, expressed in dB.
dBm0:	Absolute power level in dBm referred to a point of zero relative level (0 dBr point).
dBm0p:	Weighted dBm0, according to [b-ITU-T O.41].
dBm0(C):	C-weighted dBm0, according to [b-ISO 1999].
dBPa:	Sound pressure level relative to 1 Pa, expressed in dB.
dBPa(A):	A-weighted sound pressure level relative to 1 Pa, expressed in dB.
dBSPL:	Sound pressure level relative to 20 µPa, expressed in dB; (94 dBSPL=0 dBPa)
dBV(P):	P-weighted voltage relative to 1 V, expressed in dB, according to [b-ITU-T O.41].
dBr:	Relative power level of a signal in a transmission path referred to the level at a reference point on the path (0 dBr point).
DELSM:	DELSM is sometimes used for Δ_{Sm} (see D-Value)
N:	Newton.

Vrms: Voltage – root mean square.

cPa: Compressed Pascal, sound pressure at the output of the hearing model in the "relative approach" after non-linear signal processing by the human ear.

6 How to use this Recommendation

This Recommendation addresses different parts and stages of development of hands-free terminals.

In case of headset hands-free terminals, clause 11 describes the appropriate tests and requirements. If not mentioned specifically, the set-up as described in clause 7.1 is applied and the requirements are identical for headset and speakerphone hands-free systems.

In case of speakerphone hands-free testing, different clauses may apply when focusing on different parts or components of the system:

– The test of hands-free microphones is described in clause 10.

– A digital interface concept for testing and debugging (not mandatory) is described in clause 8.

– The complete test of a speakerphone hands-free terminal is described in clause 11.

– The test of the performance of the short-range wireless transmission link when using a mobile phone with short-range wireless transmission interface to be connected to the hands-free system is described in clause 12.

– In case additional subjective testing is desired in order to validate the speakerphone hands-free performance under driving conditions, guidance can be found in clause 13.1.

The applicability of the different clauses of this Recommendation during a typical development process in the car industry is shown in Figure 6-1.

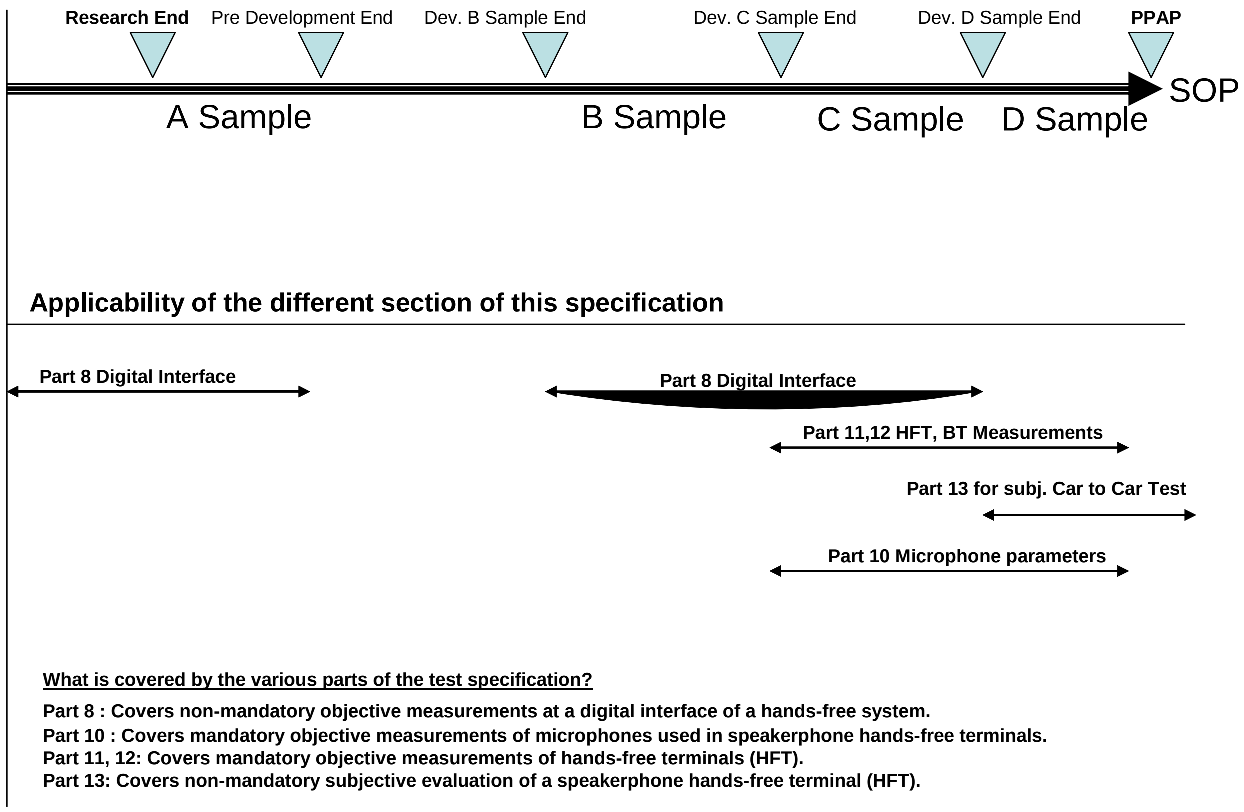


Figure 6-1 – Typical development cycle for a car speakerphone hands-free system and the applicability of the different clauses of this Recommendation during this process

7 Test arrangement

The acoustical interface for all hands-free terminals (speakerphones and headsets) is realized by using an artificial head (HATS – head and torso simulator) according to [ITU-T P.58]. The properties of the artificial head shall conform to [ITU-T P.58] for sending as well as for receiving acoustical signals.

All hands-free terminals are connected to a system simulator conforming to the required transmission standard with implemented, calibrated audio interface. For some requirements in this Recommendation, the performance limits depend on the transmission system and the speech codec used in this transmission system. The corresponding tables are found in each clause. Table 7-1 provides an overview of the narrow-band speech codecs used for the tests.

Table 7-1 – Overview of speech codecs used

System	Codec
GSM 850, 900, 1800, 1900	GSM full rate codec ([b-3GPP TS 46.010])
UMTS (WCDMA)	AMR-WB ([b-ITU-T G.722.2])@12.2 kbit/s
CDMA2000 (IS-2000)	EVRC ([b-TIA-127-A]) @ <= 8.55 kbit/s SMV ([b-TIA-893]) @ <= 8.55 kbit/s VMR-WB ([b-TIA-1016]) @ <= 13.3 kbit/s EVRC-B ([b-TIA-127-B]) @ <= 8.55 kbit/s EVRC-WB ([b-TIA-127-C]) @ <= 8.55 kbit/s
TIA/EIA IS-95-A/B CDMA	EVRC [b-TIA-127-A] @ <= 8.55 kbit/s
TIA/EIA IS-136 TDMA	[b-TIA/EIA-136-410]

The settings of the system simulator shall be chosen so that the audio signal is not influenced by any signal processing (e.g., DTX).

The test signals are fed electrically to the system simulator or acoustically to the artificial head. The test arrangement is shown in Figure 7-1.

NOTE 1 – Different codecs as well as the variation of the bit rate of codecs with variable bit rates will influence the speech quality. In order to take into account "real life" conditions, bit rates used in the real network should be used for testing and optimization.

NOTE 2 – For some mobile phones used in the hands-free set-up, the signal processing cannot be switched off completely. Therefore, care should be taken to use only such phones for tests which do not introduce additional AGC.

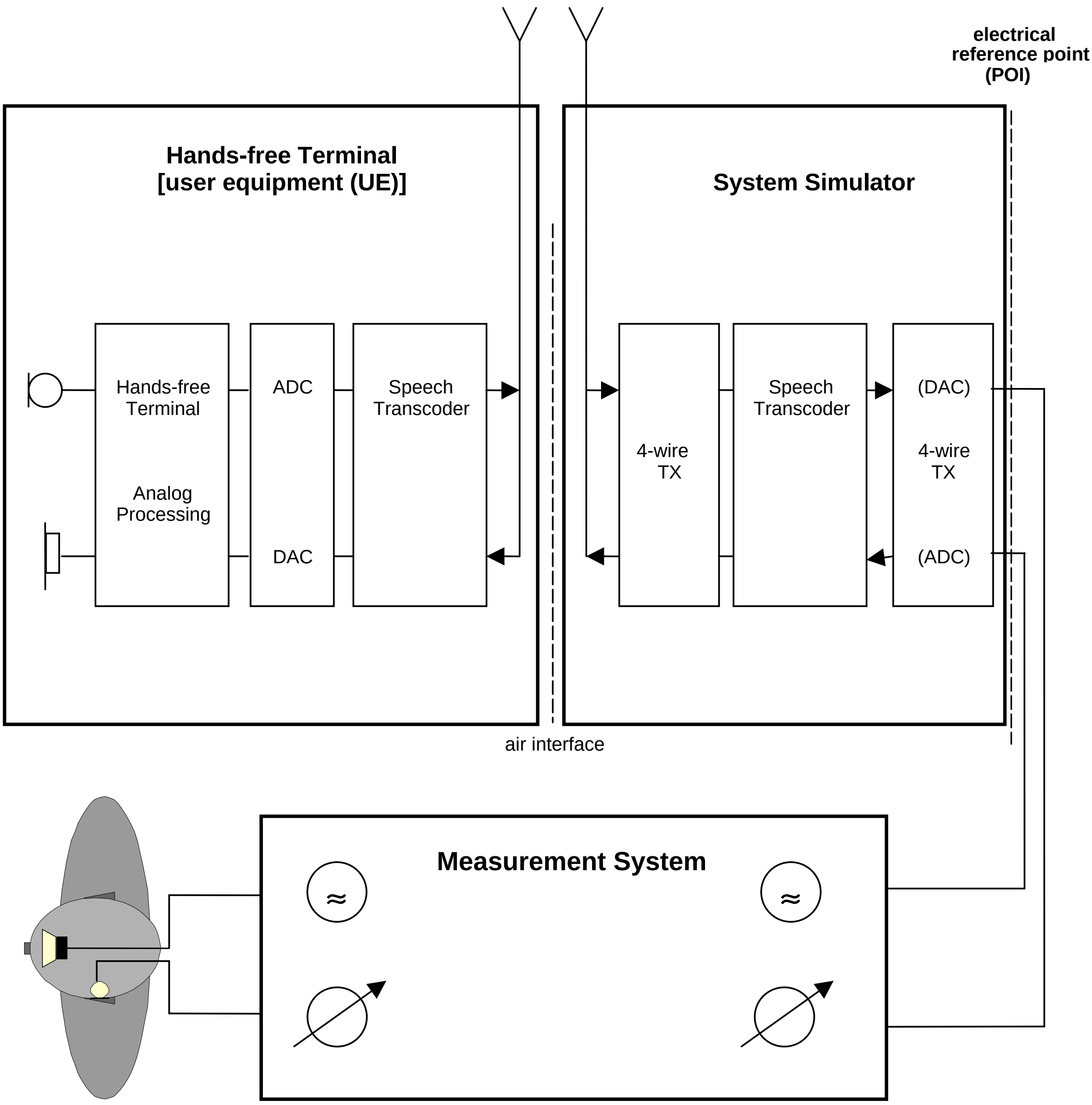


Figure 7-1 – Test arrangement for hands-free terminals

The test circuit for microphone measurements is shown in Figure 7-2.

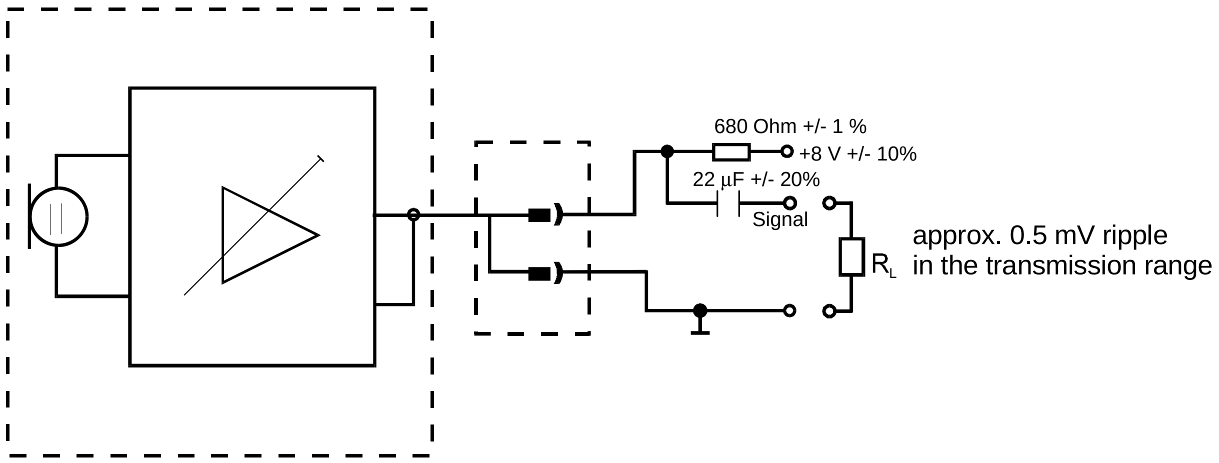


Figure 7-2 – Test arrangement for hands-free microphones and microphone arrangements

Care has to be taken that the ripple of the supply voltage does not exceed 0.5 mVrms. Furthermore, the ripple on the microphone output signal shall not exceed 0.5 mVrms measured in narrow-band. R_L shall be >10 kΩ.

7.1 Test arrangement in a car

7.1.1 Microphone-related simulation

The transmission performance of car hands-free terminals is measured in a car cabin. In order to simulate a realistic driving situation, background noise is inserted using a four-loudspeaker arrangement with subwoofer, while measurements with background noise are conducted. In Figure 7-3 the simulation arrangement is shown. More information on the test arrangement can be found in [b-ETSI EG 202 396-1]. The source signal used is recorded by a measurement microphone positioned close to the hands-free microphone. If possible, the output signal of the hands-free microphone can be used directly. The recordings are conducted in a real car. The loudspeaker arrangement is equalized and calibrated so that the power density spectrum measured at the microphone position is equal to the recorded one. For equalization, either the measurement microphone or the hands-free microphone used for recording is used. The maximum deviation of the A-weighted sound pressure level shall be ±1 dB. The third octave power density spectrum between 100 Hz and 10 kHz shall not deviate more than ±3 dB from the original spectrum. A detailed description of the equalization procedure as well as a database with background noises can be found in [b-ETSI EG 202 396-1].

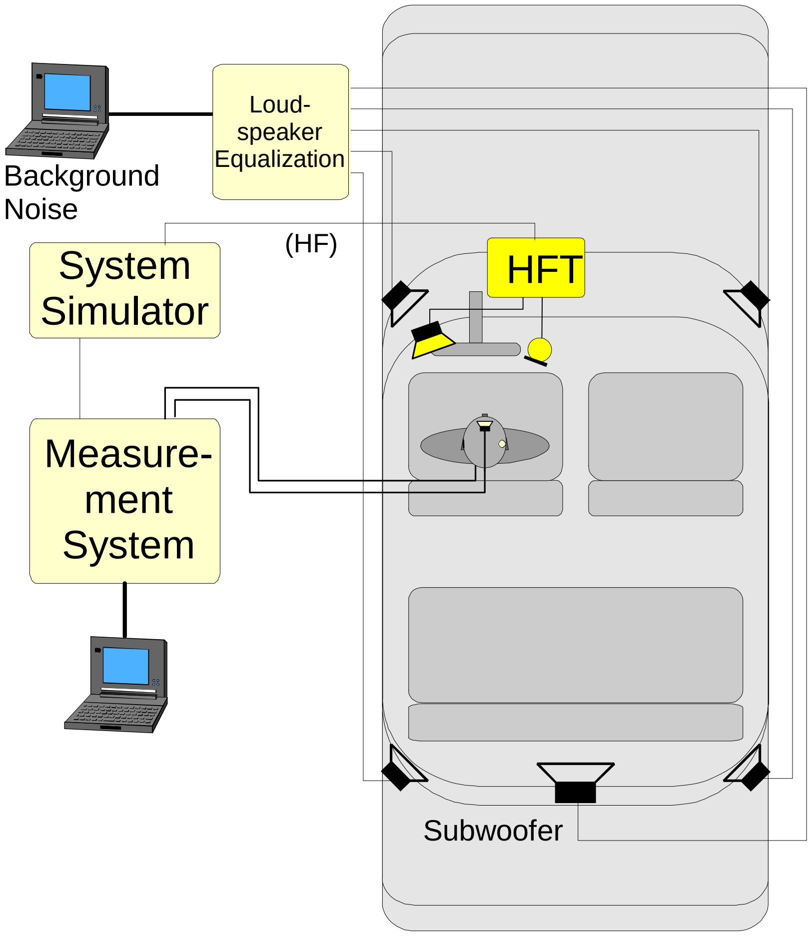


Figure 7-3 – Test arrangement with background noise simulation

7.1.2 Positioning of the hands-free terminals

The speakerphone hands-free terminal is installed according to the requirements of the manufacturer. The positioning of the microphone/microphone array and loudspeaker are given by the manufacturer. If no position requirements are given, the test lab will define the arrangements. Typically, the microphone is positioned close to the rear-view mirror, the 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 car manufacturers are measured in the original arrangement.

Headset hands-free terminals are positioned according to the requirements of the manufacturer. If no position requirements are given, the test lab will define the arrangements according to [ITU-T P.380].

If not stated otherwise, the artificial head (HATS – head and torso simulator, according to [ITU-T P.58]) is positioned in 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, all except the tallest 5% and the shortest 5% of the driving population 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.

By using suitable measures (marks in the car, relative position to A- or B-pillar, height from the floor, etc.) the exact reproduction of the artificial head position must be possible at any later time.

NOTE – Different positions of the artificial head may greatly influence the test results. Depending on the application, different positions of the artificial head may be chosen for the tests. It is recommended to check the worst-case position, e.g., those positions where the SNR and/or the speech quality in sending may be worst.

7.1.3 Artificial mouth

The artificial mouth of the artificial head shall conform to [ITU-T P.58]. The artificial mouth is equalized at the MRP according to [ITU-T P.340].

In the case of speakerphone hands-free terminals, the sound pressure level is calibrated at the HATS-HFRP so that the average level at HATS-HFRP is –28.7 dBPa. The sound pressure level at the MRP has to be corrected accordingly. A detailed description for equalization at the MRP and level correction at the HATS-HFRP can be found in [ITU-T P.581].

7.1.4 Artificial ear

For speakerphone hands-free terminals, the ear signal of the right ear of the artificial head is used (for the cars where the steering wheel is on the right hand side, the left ear is used). The artificial head is free-field equalized, more detailed information can be found in [ITU-T P.581].

For headset hands-free terminals, the type of ear to be used and the positioning is described in [ITU-T P.380].

7.1.5 Influence of the transmission system

Measurements may be influenced by signal processing (different speech codecs, DTX, comfort noise insertion, etc.) depending on the transmission system and the system simulator used in the test set-up. If requirements cannot be fulfilled due to impairments introduced by the transmission system or the system simulator, reference measurements of the hands-free unit or measurements without acoustical components should be made to document this behaviour.

7.1.6 Calibration and equalization

The following preparation has to be completed before running the tests:

Calibration:

– Acoustical calibration of the measurement microphones as well as of the HATS microphone.
– Calibration and equalization of the artificial mouth at the MRP.
– HATS-HFRP calibration (for speakerphone hands-free terminals only).

Equalization (for speakerphone hands-free terminals only):

– Free-field equalization of the artificial head.

Reference measurement:

– For the compensation of the different power density spectra of the measurement signals, it is required to refer the measured power density spectra to the power density spectra of the test signal. This is denoted as a reference measurement.
– In the sending direction, the reference spectrum is recorded at the MRP.
– In the receiving direction, the reference spectrum is recorded at the electrical interface.

7.1.7 System simulator settings

All settings of the system simulator have to ensure that the audio signal is not disturbed by any processing and the transmission of the HF signal is error-free. DTX shall be switched-off. For all networks, the RF-level shall be set to maximum. The settings shall be reported in the test report.

For measurements according to the GSM standard, the full rate codec shall be used. For measurements with an AMR codec, the highest bit rate of 12.2 kbit/s is used.

7.1.8 Environmental conditions

Unless specified otherwise, the background noise level shall be less than –54 dBPa(A) inside the car.

8 Digital interfaces for development, debugging and test

The interface concept and tests described in this clause is optional and may be used for the purpose of development, debugging and testing of hands-free implementations specifically during the development and optimization process. It can be applied if the digital interfaces are available, typically in the case of prototype or development boards, or in the case of factory-fitted HF devices.

8.1 Interfaces and access points

Digital interfaces allow the recording and investigating of signals at the specified access points. Some of the digital interfaces at access points *before* the HF system processing should also allow for writing/adding a digital signal to the signal path. This is true for the sending as well as for the receiving path.

Depending on the access point, any of the following three access means should be possible:

– READ: Writing the respective signal into a file.

– WRITE: Replacing a certain signal in the system by a digital signal from a file.

– ADD: Adding a digital signal from a file to a certain signal in the system.

Figure 8-1 gives an overview to the digital interfaces that are useful for development, debugging and test.

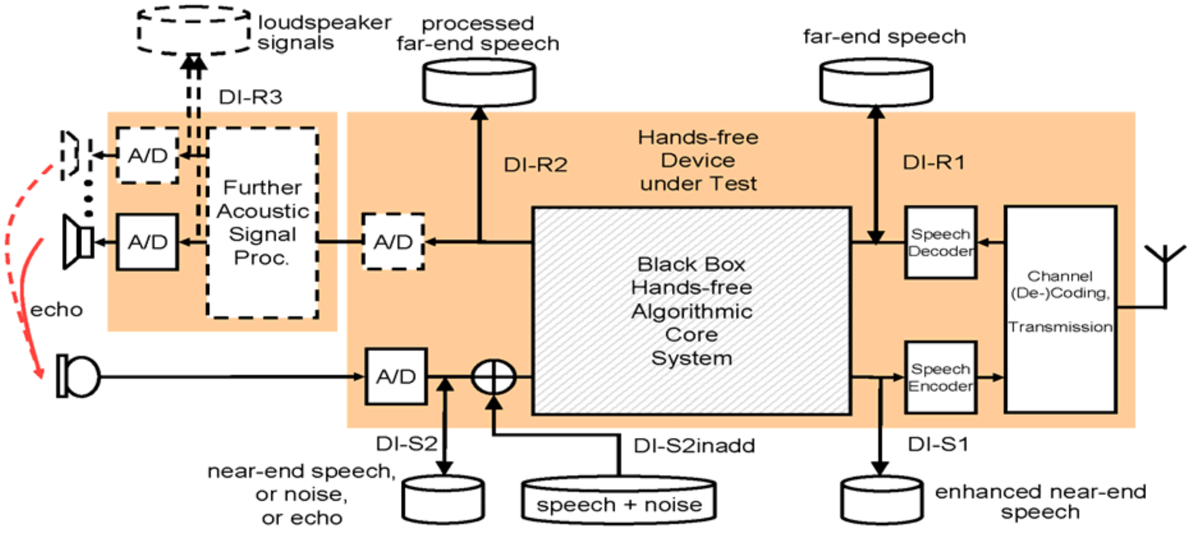


Figure 8-1 – Digital interfaces for the HF system

The digital interfaces (DIs) are called DI-{R | S}n with R standing for receiving path and S standing for sending path. The number n is used to distinguish between different digital interfaces in sending and receiving path, respectively.

DI-R1 can be used to record transmitted far-end speech (READ) or to test the hands-free device under test using recorded signals without actual involvement of a system tester (WRITE).

DI-R2, in comparison with DI-R1, can be used to evaluate the HF systems core algorithms in the receiving path. Here only READ access should be realized.

In some systems, further digital signal processing may be used, connected digitally or by analogue means to the HF algorithmic core system. In this case, DI-R3 yields useful signals to evaluate this system component. Such further acoustic signal processing may comprise an artificial bandwidth extension, or it may comprise typical audio processing functions related to a number of loudspeakers used (equalizers, room effects).

In the sending path, DI-S2 is the access point of greatest interest. If any of the digital access points is realized, this one shall be realized as well. It allows recording (READ) of any test case signals after the AD converter. Developers and testers may choose this access point to pre-record all near-end noises in their test scenario, stemming from real driving situations or from a background noise playback arrangement. Also, they may choose to pre-record all near-end speech or speech-like signals in their test scenario. DI-S2 should also allow WRITE access.

Given unchanged analogue processing and AD conversion in the sending path, the recorded noise and near-end signals can then be used to repeat test cases in an efficient way. This becomes possible by digital offline addition of near-end speech and noise, and by adding this signal to the sending input path to the HF system DI-S2inadd (ADD), while the HF system is in real-time operation in the sending and receiving path. In such cases, only the echo needs to be available in the car cabin. Therefore, no exact positioning of the HATS is required, or no HATS at all is necessary. A reduction in test effort is achieved by avoiding background noise simulation or even testing with real driving noise.

Finally, DI-S1 allows access to the HF system output signal in sending direction (READ). This signal gives important information about the core HF system's functionality; namely, acoustic echo cancellation and noise reduction. However, end-to-end system level performance should not be judged based on this access point because of complex interactions between the HF signal processing and the speech coder/network-side speech enhancement devices.

If digital interfaces are implemented for an HF system, at least one of the following formats shall be supported:

– 16 bit linear PCM.

– ITU-T G.711 A-law.

– ITU-T G.711 μ-law.

The sampling frequency of the digital interfaces should be 8 kHz, except where processing in the HF system is performed at different sampling rates. When using different sampling rates at the test system appropriate, up- and down-sampling should be used.

8.2 Test set-up and tests

In general, the digital interfaces can be used in virtually all test cases described in clause 10. If digital interfaces are available, the following recordings and tests should be done.

8.2.1 Recording and insert background noise

In many test cases, background noises are required. Recording of the background noises can be performed digitally via interface DI-S2, feedback into the system and addition to the microphone signal can be performed digitally with interface DI-S2inadd.

8.2.2 Recording and insert near-end speech recordings

In many test cases, near-end speech or artificial voice signals are required. Recording can be performed digitally via interface DI-S2, feedback into the system and addition to the microphone signal can be performed digitally with interface DI-S2inadd.

8.2.3 One-way speech quality in sending

In analogy to clause 11.5.1, the one-way speech quality in sending can be measured with stored near-end test signals (see Appendix I) via interface DI-S2. Feedback during the test shall be done via interface DI-S2inadd. Two measurement points shall be used: First, the electrical reference point (POI), in order to perform the test for the requirement described in clause 11.5.1, yielding MOS-LQO-N(POI). Second, the measurement can be done via the DI-S1 interface yielding MOS-LQO-N(S1) for diagnostic purposes. The requirement is:

$$\text{MOS-LQON(S1)} \geq \text{MOS-LQON(POI)} \geq 3.0$$

NOTE – It is known that there might be specific types of signal processing which lead to a degradation at the intermediate point S1 but which might improve the overall system performance. If it can be demonstrated that the optimized end-to-end system performance is achieved when this requirement is violated, this requirement does not apply.

The value of DELTA = MOS-LQON(S1) – MOS-LQON(POI) can be caused by:

– the codecs and the network;

– an interaction between the HF signal processing and the codecs/network;

– measurement error in the objective MOS prediction algorithm.

8.2.4 Speech distortion in double talk

The digital interface allows for a comfortable measurement of the distortion of the speech component in sending in double talk. The test is aimed to help optimizing the signal processing of the HFT algorithmic core system with respect to speech quality during double talk.

The test is based on the same stored near-end speech test signals as used in clause 8.2.3 (see clause I.1) recorded via interface DI-S2. These signals are used as reference signals for the determination of the speech distortion during double talk in sending.

The far-end speech test signals are the ones defined in clause I.2.

The processing steps for the test are the following:

– Before starting the double talk tests, the test lab should ensure that the echo canceller is fully converged. This can be done by an appropriate training sequence (see also clause 11.10).

– The HF system is to be processed in real-time with the speech input signals on both sides (interface DI-R1 in receiving, and DI-S2inadd in sending). It must always be ensured that different talkers are used for the receiving and sending directions. In 25% of the test cases, two female voices shall be applied; in 25% of the test cases, two male voices shall be applied; and in 50% of the test cases different genders in the receiving and sending directions shall be used. The echo as captured by the microphone is then added in real-time to the stored near-end speech signal accessed through interface DI-S2inadd.

– During processing, the echo signal is digitally stored via DI-S2. Also, the enhanced speech signal at the output of the HF system in sending is stored via DI-S1.

– Using the echo (DI-S2), the near-end speech (DI-S2inadd), the output of the HF system in sending (DI-S1), and the signal at the electrical reference point (POI) in sending, the following speech distortion measurements shall be applied.

Speech distortion shall be evaluated in terms of the quality of the speech component (1) at DI-S1 and (2) at the POI with the stored speech signal at DI-S2inadd as reference.

The speech component of the signal at DI-S1 or at the POI can be extracted using the signal separation methodology (for more information see [b-Fingscheidt]), using a Blackman window of 512 samples with a frame shift of ≤64 samples. In analogy to clause 8.2.3, the requirement is stated as:

$$\text{MOS-LQON(S1)} \geq \text{MOS- LQON(POI)} \geq 2.5.$$

NOTE – It is known that there might be specific types of signal processing which lead to a degradation at the intermediate point S1 but which might improve the overall system performance. If it can be demonstrated that the optimized end-to-end system performance is achieved when this requirement is violated, this requirement does not apply.

The MOS-LQON analysis is performed based on [ITU-T P.862] and [ITU-T P.862.1]. The value of DELTA = MOS-LQON(S1) – MOS-LQON(POI) can be caused by:

– codecs and the network;

– an interaction between the HF signal processing and the codecs/network;

– measurement error in the objective MOS prediction algorithm;

– echo control strategy of the hands-free terminal and its interaction with the mobile phone.

9 Test signals and test signal levels

9.1 Signals

Speech-like signals are used for the measurements which can be found in [ITU-T P.50] and [ITU-T P.501]. Detailed information about the test signal used is to be found in the corresponding clause of this Recommendation.

For narrow-band hands-free terminals, all test signals – which are used in the receiving direction – have to be band-limited. The band limitation is achieved by bandpass filtering in the frequency range between 200 Hz and 4 kHz using bandpass filtering providing 24 dB/octave. In the sending direction, the test signals are used without band limitation.

All test signal levels are referred to the average level of the test signals, averaged over the complete test sequence length, if not described otherwise. In the receiving direction, the band-limited test signal is measured; in the sending direction no band-limitation is applied.

The average signal levels for the measurements are as follows:

– –16 dBm0 in the receiving direction (typical signal level in networks).

– –4.7 dBPa in the sending direction at the MRP (typical average speech levels) (equivalent to –28.7 dBPa at the HATS-HFRP).

NOTE – If different networks' signal levels are to be used in tests, this is stated in the individual test. The "Lombard effect" (increased talker speech level due to high background noise) is considered in the background noise tests.

Some tests require exact synchronization of test signals in the time domain. Therefore, it is required to take into account the delays of the terminals. When analysing signals, any delay introduced by the test system codecs and terminals have to be taken into account accordingly.

9.2 Background noise signals

For some measurements, typical background noise is inserted. This is described in the corresponding clause. In general, such background noise should be car-specific and should be simulated for the car cabin tested. The test lab (together with the manufacturer) will decide which background noise is used for the test. Car-specific parameters, e.g., driving with an open roof in a cabriolet, have to be taken into account. Specific driving situations, e.g., driving with an open window, may be taken into account as well. In general, it is recommended to conduct all tests during constant driving conditions simulating fixed driving speed (e.g., 130 km/h). Under this condition, it is easier to conduct reproducible measurements.

If no requirements are made by the car manufacturers, a minimum background noise sound pressure level of –24 dBPa(A), measured at the inboard ear of the artificial head, has to be achieved. In any case, the recording of a real driving noise with constant speed shall be used.

Recording of driving noise

Background noise is recorded in the real car. The measurement microphone is positioned close to the hands-free microphone. Alternatively, the hands-free microphone can be used for the recording of the background noise if the microphone is easily assessable.

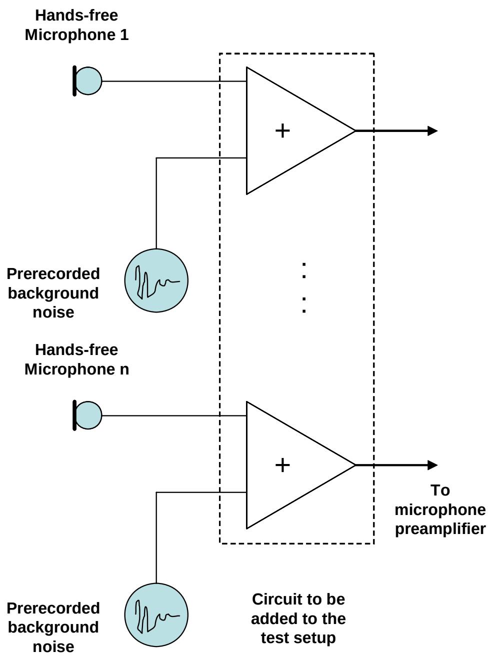
Playback of the recorded background noise

One possibility is that the test lab employs a four-loudspeaker arrangement for acoustic background noise reproduction in the car cabin. Typically, two loudspeakers are mounted in the front and in the rear (left and right side). The loudspeaker should be carefully positioned in order to minimize disturbance of the transmission paths between the loudspeakers and hands-free microphone and the artificial head at the driver's seat. Details can be found in [b-ETSI EG 202 396-1].

Alternatively, the background noise can be inserted electrically to the microphone signal. Therefore, the background noise signal recorded at the electrical output of the hands-free microphone(s) is inserted at the electrical access point which was used for the recording. Appropriate electronics allowing the mixing of the previously recorded background noise signal(s) with the microphone signal(s) at this access point has to be provided, see Figure 9-1. The test lab has to ensure the right calibration of the two signals.

As a third alternative, the background noise can be digitally recorded at the DI-S2 interface in Figure 8-1 and later digitally inserted (added) as described in clause 7 via interface DI-S2inadd in Figure 8-1.

NOTE – Both with analogue as well as digital electrical feedback of the noise signal (alternatives 2 and 3) structure-borne noise can be captured as well.



NOTE – Structure-borne noise is also covered with this arrangement, which is part of the microphone recording.

Figure 9-1 – Set-up for analogue electrical insertion of the pre-recorded background noise signal

10 Measurement parameters and requirements for microphones used in speakerphone hands-free systems

This clause is intended for the measurements of microphones without any additional electrical signal processing. Other types of microphones (e.g., beam-forming arrays) are measured in conjunction with the hands-free system as described in clause 11.

10.1 Microphone measurements in anechoic conditions

The scope of these measurements is the verification of microphone parameters in a defined acoustic environment without the influence of integration such as mounting, orientation and in-car acoustics.

10.1.1 Microphone sensitivity

10.1.1.1 Requirements

The microphone sensitivity has to be measured in the free sound field. The sensitivity refers to the sound pressure of the undisturbed free sound field (in the absence of the microphone). The sensitivity is measured at the output of the test circuit according to Figure 7-2.

The microphone sensitivity at 1 kHz shall be 300 mV/Pa ±3 dB when measured in the direction of its maximum sensitivity.

10.1.1.2 Test

1) The test signal is a sine wave of 1 kHz at a level of 0 dBPa at the microphone position in an undisturbed free sound field.
2) The microphone is positioned at a distance of 1 m on the acoustic centre-line of the loudspeaker.
3) The microphone is oriented to the loudspeaker in its direction of maximum sensitivity.
4) The sensitivity is determined in mV/Pa.

Further information can be found in [IEC 60268-4].

10.1.2 Microphone frequency response

10.1.2.1 Requirements

The microphone frequency response has to be measured in the free sound field. The frequency response refers to the sound pressure of the undisturbed free sound field (in the absence of the microphone). The frequency response is measured at the output of the test circuit according to Figure 7-2.

Table 10-1 – Tolerance mask for the sending sensitivity frequency response

Frequency [Hz]	Upper limit	Lower limit
200	0	
250	0	
315	0	–14
400	0	–13
500	0	–12
630	0	–11
800	0	–10
1 000	0	–8
1 300	2	–8
1 600	3	–8
2 000	4	–8
2 500	4	–8
3 100	4	–8
4 000	4	
NOTE – All sensitivity values are expressed in dB on an arbitrary scale.		

NOTE – Depending on customer demands, other tolerance schemes than described in Table 10-1 may be applied and have to be defined in an equivalent format.

10.1.2.2 Test

1) The test signals are sine waves at a level of 0 dBPa at the microphone position in an undisturbed free sound field covering at least the defined frequency range.
2) The microphone is positioned at a distance of 1 m on the acoustic centre-line of the loudspeaker.
3) The microphone is oriented to the loudspeaker in its direction of maximum sensitivity.
4) The sensitivity for each frequency is determined in mV/Pa.

Further information can be found in [IEC 60268-4].

10.1.3 Microphone directional characteristics

The directional characteristics of a microphone are described by different sensitivities at different angles of sound incidence.

10.1.3.1 Requirements

The front-to-back ratio is the ratio between the sensitivity in the direction of highest sensitivity to the sensitivity in the direction of lowest sensitivity, expressed in dB at 1 kHz. The front-to-back ratio is measured at the output of the test circuit according to Figure 7-2.

To achieve appropriate noise reduction, the front-to-back ratio shall be at least 10 dB.

NOTE – Depending on mounting and orientation, lower front-to-back ratios can also be an advantage.

10.1.3.2 Test

1) The test signal is a sine wave of 1 kHz at a level of 0 dBPa at the microphone position in an undisturbed free sound field.
2) The microphone is positioned at a distance of 1 m on the acoustic centre-line of the loudspeaker.
3) The first measurement is done with the microphone oriented to the loudspeaker in its direction of maximum sensitivity. The second measurement is done with the microphone oriented to the loudspeaker in its direction of minimum sensitivity. If the direction of minimum sensitivity is not known, it has to be determined by rotating the microphone until the minimum is found.
4) The front-to-back ratio is determined in dB.

Further information can be found in [IEC 60268-4].

10.1.4 Microphone distortion

10.1.4.1 Requirements

The microphone distortion refers to the sound pressure of the undisturbed free field. The distortion is measured at the output of the test circuit according to Figure 7-2.

The total harmonic distortion with a sound pressure level of 0 dBPa (94 dBSPL) at the position of the microphone shall be less than 1% in the narrow-band frequency range.

10.1.4.2 Test

1) The test signal is a sine wave with a frequency of 300 Hz, 500 Hz and 1 kHz at a level of 0 dBPa.
2) The microphone is positioned on the acoustic centre-line of the loudspeaker.
3) The microphone is oriented to the loudspeaker in its direction of maximum sensitivity.
4) The total harmonic distortion is expressed as a percentage.

Care has to be taken that the loudspeaker is able to produce the required sound pressure level with a lower distortion than the microphone under test.

10.1.5 Maximum sound pressure level

10.1.5.1 Requirements

The maximum sound pressure is defined by the sound pressure level where the total harmonic distortion of the microphone at 1 kHz is 3% in the narrow-band frequency range. The total harmonic distortion is measured at the output of the test circuit according to Figure 7-2.

The maximum sound pressure level should be higher than 106 dBSPL for a microphone with a typical sensitivity of 300 mV/Pa.

10.1.5.2 Test

1) The test signal is a sine wave with a frequency of 1 kHz and an increasing level to determine the level of 3% of total harmonic distortion.
2) The microphone is positioned on the acoustic centre-line of the loudspeaker.
3) The microphone is oriented to the loudspeaker in its direction of maximum sensitivity.
4) The maximum sound pressure level is expressed in dBSPL or dBPa.

Care has to be taken that the loudspeaker is able to produce the required sound pressure level with a lower distortion than the microphone under test.

NOTE – With a good microphone design, the maximum sound pressure level is electrically limited by the supply circuit according to Figure 7-2. A microphone with higher sensitivity will reach the electrical output limits at a lower sound pressure level compared to another microphone with lower sensitivity.

10.1.6 Self noise

10.1.6.1 Requirements

The maximum self noise measured at the output of the test circuit according to Figure 7-2 in quiet conditions shall be less than –72 dBV(P).

10.1.6.2 Test

1) For the measurement, no test signal is used.
2) The microphone has to be powered with a low noise voltage supply.
3) The self noise is measured at the output of the test circuit according to Figure 7-2 in the frequency range between 100 Hz and 4 kHz, psophometric weighting has to be applied.
4) The self noise is expressed in dBV(P).

Care has to be taken that the environmental noise is below the equivalent self noise of the microphone.

10.2 Microphone measurements in the car

Positioning of hands-free microphones

The speech quality in hands-free communication is significantly affected by the positioning of the hands-free microphone. As the optimal microphone position can vary strongly depending on vehicle design as well as on other specific requirements, there is no universally valid rule for the positioning of the microphone. However, there are some guidelines which should be considered. Nevertheless, in practice this often means finding the best compromise, as not all requirements can be equally fulfilled.

– The hands-free microphone should always be placed as close as possible to the speaker, since within the near field of a sound source (in a vehicle, this is up to 80-100 cm)[1] the speech level drops by $1/d^2$. In practical applications, this typically means an analogous loss in signal-to-noise ratio. For this reason, a single microphone placed nearby might give better performance than a microphone array, which is placed further away.
– There has to be a direct path between the speaker's mouth and the microphone. If this is not given, this might result in a significant decrease in signal-to-noise ratio as well as in speech quality since the speech signal becomes reverberant.
– The direction of the highest sensitivity of the microphone should point in the direction of the speaker's mouth. If different seating positions or several speakers are to be covered by one microphone, a compromise for the microphone position has to be found, since the direction of the highest sensitivity might not cover all. However, this often means a significantly reduced performance in comparison to an optimal alignment of the microphone for a single speaker. In this case, the application of additional microphones might be considered to achieve an optimal speech quality.
– A direct airstream towards the microphone, e.g., from the air conditioning, has to be avoided since the speech signal might be highly disturbed by wind buffeting.
– Saturation of the microphone by nearby loudspeakers, e.g., by a centre-speaker, has to be avoided. If necessary, the levels of the affected loudspeakers have to be reduced.

Coupling of structure-born sound to the microphone has to be avoided.

NOTE 1 – When the microphone performance is measured in the car, it is recommended to use the power supply provided by the car/car hands-free system.

NOTE 2 – If the microphone is integrated digitally into the car, it is recommended to measure the microphone performance at a digital access point, if available. Care should be taken in order to correctly calibrate the access point.

10.2.1 Microphone output level in the car

10.2.1.1 Requirements

The microphone sensitivity is determined from MRP to the output of the test circuit according to Figure 7-2.

For typical applications, the microphone output voltage should be in the range of

$$50\ \text{mV}_{\text{rms}} \pm 3\ \text{dB with 0 dBPa at MRP}$$

(equivalent to a microphone sensitivity of about 300 mV/Pa and a measurement in anechoic conditions at 50 cm distance from the microphone to the MRP).

However, depending on specific electrical/acoustical designs, arrangements inside the car, or other factors, the sensitivity requirement may be different. Therefore, this requirement has to be adapted to the individual arrangements inside a car.

10.2.1.2 Test

1) The test signal is a one-third octave noise signal with a mid-frequency of 1 kHz and a level of –10 dBPa measured at the MRP.
2) The microphone sensitivity is determined in a car with a microphone installed. The test arrangement is according to the arrangement described in clause 7.1.

[1] The near field is characterized by the distance (measured from the sound source) where the direct sound and the reflected sound are of equal intensity. In acoustics, this distance is often referred to as the critical distance.

3) The output voltage is determined in mV.

10.2.2 Overload point

10.2.2.1 Requirements

The overload resistance shall be >15 dB (referred to a nominal sound pressure level of –4.7 dBPa at the MRP and a distance of 50 cm).

10.2.2.2 Test

1) The test signal is a one-third octave noise signal with a mid-frequency of 1 kHz and a level of –10 dBPa and +5 dBPa measured at the MRP.
2) The overload point is determined in anechoic conditions. The distance between the MRP and the microphone is 30 cm (note that, since the artificial mouth is unable to produce a sound pressure of 10.3 dBPa, the distance between the artificial mouth and the microphone is reduced to 30 cm).
3) The output voltage is determined in mV. The deviation of the measured sensitivities shall be less than 0.1 dB.

10.2.3 Microphone frequency response in the car

10.2.3.1 Requirements

The microphone frequency response is measured from the MRP to the output of the test circuit according to Figure 7-2.

The tolerance mask for the sensitivity frequency response in the sending direction is given in Table 10-2. The limit at intermediate frequencies lies on a straight line drawn between the given values on a log (frequency) – linear (dB) scale.

Table 10-2 – Tolerance mask for the sending sensitivity frequency response

Frequency [Hz]	Upper limit	Lower limit
200	0	
250	0	
315	0	–14
400	0	–13
500	0	–12
630	0	–11
800	0	–10
1 000	0	–8
1 300	2	–8
1 600	3	–8
2 000	4	–8
2 500	4	–8
3 100	4	–8
4 000	4	
NOTE – All sensitivity values are expressed in dB on an arbitrary scale		

10.2.3.2 Test

1) The test arrangement is according to clause 7.1.

2) The test signal artificial voice according to [ITU-T P.50] is used. Alternatively, a periodic noise signal or CS signal according to [ITU-T P.501] can be used. The artificial mouth is equalized at the MRP, the test signal level shall be –4.7 dBPa at the MRP. The test signal level is the average level of the complete test signal. Finally, the level at the HATS-HFRP is adjusted to –28.7 dBPa.

3) The measured power density spectrum at the MRP is used as the reference power density spectrum for determining the frequency response in the sending direction.

4) The sensitivity frequency response is determined in one-third octave intervals as given by the R.40-series of preferred numbers in [ISO 3] for frequencies of 100 Hz and 4 kHz, inclusive. For calculation, the average measured level is referred to the level of the reference signal in each frequency band averaged over the complete test sequence length.

5) The sensitivity is determined in dBV/Pa.

10.2.4 Idle channel noise

10.2.4.1 Requirements

The maximum idle channel noise in the sending direction measured at the output of the test circuit according to Figure 7-2 in quiet conditions shall be less than –72 dBV(P). Spectral peaks in the frequency domain shall not exceed the averaged spectrum by more than 10 dB:

NOTE – It is recognized that fan noise, which is individually different for each car depending upon microphone and fan arrangement, may contribute significantly to the noise perceived by the far-end user. In order to determine the level and spectral content of this noise under different operating conditions, a noise test as described below may be used.

10.2.4.2 Test

1) For the measurement, no test signal is used. In order to ensure a reliable activation of active microphone arrangements, an activation signal is inserted before the actual measurement. The activation signal consists of a sequence of 4 composite source signals according to [ITU-T P.501]. The spectrum of the test at the MRP is calibrated under free-field conditions. The level of the activation signal is –28.7 dBPa measured at the HATS-HFRP.

2) The test arrangement is according to clause 7.1.

   The idle channel noise is measured at the output of the test circuit according to Figure 7-2 in the frequency range between 100 Hz and 4 kHz. The measurement requires a time window which starts exactly at the end of the activation signal. Any "ringing" of filters or receivers shall be taken into account, the time window must be shifted accordingly. The length of the time window is 1 second, which is the averaging time for the idle channel noise. The test lab has to ensure the correct activation of the microphone/microphone arrangement during the measurement. If the microphone arrangement is deactivated during measurement, the measurement window has to be cut to the duration when the microphone remains activated.

   The power density spectrum of the noise signal is determined using FFT (8 k samples/48 kHz sampling rate or equivalent). A Hanning window is used.

   If it is known that the microphone stays activated without any activation signal, no activation signal is required. In this case, a simple noise measurement is conducted.

3) The idle channel noise is determined by psophometric weighting. Spectral peaks are measured in the frequency domain. The average noise spectrum used for determining the spectral peak should be calculated as the arithmetic mean of the noise spectrum values when stated in dBV(P).

10.2.5 Ambient noise rejection (in the car)

10.2.5.1 Requirements

Microphones typically have to operate with high background noise without causing any impairment. Any background noise should be reduced as much as possible. It is known that low bit-rate speech coders which are optimized for speech are sometimes very sensitive to background noise.

The ambient noise rejection is determined as a D-value from DELSM according to [ITU-T G.111]. D should be ≥ –13 dB.

NOTE – Considering a microphone without additional signal processing, lower D-values than specified have to be expected. Table 10-3 gives measured D-values with an omnidirectional microphone in typical car installation positions. With an optimally integrated and oriented directional microphone, improvements of up to 3 dB, compared to the omnidirectional microphone, are possible. Vice versa, a wrong orientation or poor integration can lead to lower values. The D-value required depends on the distance between the microphone and the talker, the directivity of the microphone, placement and integration in the car, and other factors. A D-value of –13 dB can be achieved for excellent designs only.

Table 10-3 – D-value depending on distance and positioning in a car

Distance [mm]	D-value [dB]	Remark
445	–20.94	Microphone position = centre of OHC, small driver
455	–20.62	Microphone position = centre of OHC, large driver
450	–20.11	Microphone position = mirror base, small driver
530	–22.07	Microphone position = mirror base, large driver
220	–18.26	Microphone position = sun visor, small driver
245	–16.40	Microphone position = sun visor, medium driver
270	–18.56	Microphone position = sun visor, large driver

10.2.5.2 Test

1) The test arrangement is according to clause 7.1.
2) Depending on the manufacturer/test lab requirements, the background noise – which represents a typical driving condition – is inserted. The test shall be conducted under constant background noise conditions.
3) The sending signal is recorded at the output of the test circuit according to Figure 7-2. The measurements are performed in one-third octave bands according to [IEC 61260] in bands 4-17 (200 Hz-4 kHz). In each band, the diffuse field sensitivity S_{si}(diff) is measured. The result is scaled in dBV/Pa.
4) The speech signal is simulated by a composite source signal according to [ITU-T P.501] with a duration of ≥2 CS sequences. The level at the MRP is –4.7 dBPa. The direct sensitivity S_{si}(direct) is measured again in one-third octave bands according to [IEC 61260] in bands 4-17 (200 Hz-4 kHz) and scaled in dBV/Pa.
5) The D-value, according to equations E-2 and E-3 in Annex E of [ITU-T P.79], is calculated in bands 4-17. The coefficients K_i, as described in Table E.1 of [ITU-T P.79], are used.

11 Measurement parameters and requirements for hands-free terminals

11.1 Preparation measurements

Before conducting the tests, proper calibration and equalization of the test system has to be performed.

11.2 Delay

11.2.1 Delay in sending direction

11.2.1.1 Requirements

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

$$T_s + t_{System}$$

The delay T_s shall be less than 50 ms.

NOTE 1 – The delay should be minimized. This can, e.g., be accomplished by designing the speech decoder output, the short-range wireless transmission 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 short-range wireless transmission link and to the speech coder allows to (partially) embed the delay of one block into the preceding one.

NOTE 2 – The delay requirement assumes a delay of maximum 8 ms inserted by a potential short-range wireless transmission link. Therefore, tests should be made with a short-range wireless transmission mobile phone which introduces a low delay.

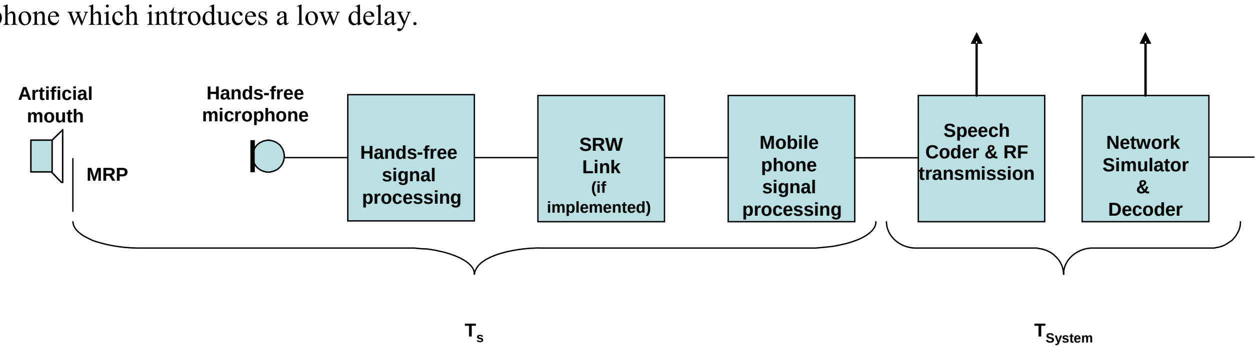


Figure 11-1 – Different blocks contributing to the delay in sending

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

11.2.1.2 Test

1) For the measurements, a composite source signal (CSS) according to [ITU-T P.501] 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. For speakerphone hands-free terminals, the test signal level is adjusted to –28.7 dBPa at the HATS-HFRP [ITU-T P.581]. The equalization of the artificial mouth is made at the MRP.

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

   The set up of the hands-free terminal is in correspondence to clause 7.1.

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.

11.2.2 Delay in receiving direction

11.2.2.1 Requirements

The delay in the receiving direction is measured from the POI (input of the reference speech coder of the system simulators) to the ear-drum reference point (DRP). The delay measured in the receiving direction is:

$$T_r + t_{System}$$

The delay T_r shall be less than 50 ms.

NOTE 1 – The delay should be minimized. This can, e.g., be accomplished by designing the speech decoder output, the short-range wireless transmission 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 short-range wireless transmission link and to the speech coder allows to (partially) embed the delay of one block into the preceding one.

NOTE 2 – The delay requirement assumes a delay of maximum 8 ms inserted by a potential short-range wireless transmission link. Therefore, tests should be made with a short-range wireless transmission mobile phone which introduces a low delay.

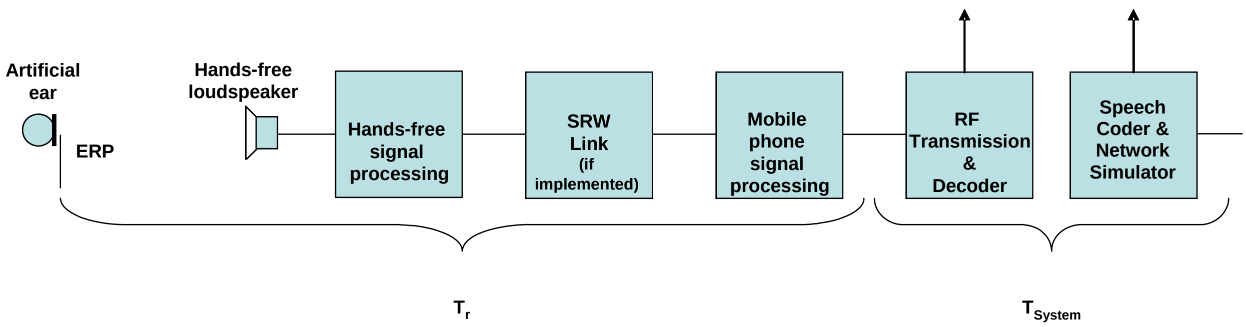


Figure 11-2 – Different blocks contributing to the delay in receiving

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

11.2.2.2 Test

1) For the measurements, a composite source signal (CSS) according to [ITU-T P.501] 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 7.1. For the measurement of speakerphone hands-free terminals, the artificial head is free-field equalized according to [ITU-T P.581]. The equalized output signal of the inboard ear is used for the measurement. For headset hands-free terminals, the sound pressure is measured at the DRP and corrected to the ERP according to [ITU-T P.57].

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.

11.3 Loudness ratings

11.3.1 Requirements

The nominal values of SLR/RLR from/to the electrical reference point (POI) shall be:

– for speakerphone hands-free terminals:

$$SLR = 13\ dB \pm 4\ dB;$$

$$RLR = 2\ dB \pm 4\ dB.$$

– for headset hands-free terminals:

$$SLR = 8\ dB \pm 4\ dB;$$

$$RLR = 2\ dB \pm 4\ dB.$$

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. The volume control range shall allow the setting of S/N ≥6 dB for all signal and noise conditions. This will allow sufficient loudness of the speech signal in the receiving direction in the presence of high background noise.

NOTE – It is recognized that the car may be a working place. Therefore, care has to be taken not to exceed the limits for daily noise exposure defined in the different regional standards and directives for working places.

11.3.2 Test loudness rating in sending direction

1) The test signal used for the measurements shall be artificial voice according to [ITU-T P.50]. The test signal is equalized at the MRP, the test signal level is –4.7 dBPa at the MRP. The test signal level is the average level of the complete test signal. For speakerphone hands-free terminals, the level at the HATS-HFRP is adjusted to –28.7 dBPa.

   The measured power density spectrum at the MRP is used as the reference power density spectrum for determining the sending sensitivity.

2) The test arrangement is according to clause 7.1. The sending sensitivity is calculated from each band of the 14 frequencies given in Table 1 of [ITU-T P.79], bands 4-17.

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

3) The sensitivity is expressed in dBV/Pa, the sending loudness rating (SLR) shall be calculated according to equation 5-1 of [ITU-T P.79], bands 4-17, m = 0.175 and the weighting factors in the sending direction according to Table 1 of [ITU-T P.79].

11.3.3 Test loudness rating in receiving direction

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

2) The test arrangement is according to clause 7.1. For the measurement of speakerphone hands-free terminals, the artificial head is free-field equalized according to [ITU-T P.581].

The equalized output signal of the inboard ear is used for the measurement. For headset hands-free terminals, the sound pressure is measured at the DRP and corrected to the ERP according to [ITU-T P.57]. The receiving sensitivity is determined by the bands 4-17 according to Table 1 of [ITU-T P.79].

For the calculation, the average signal level of each frequency band is referred to the signal level of the reference signal measured in each frequency band.

3) The sensitivity is expressed in dBPa/V, the receiving loudness rating (RLR) shall be calculated according to equation 5-1 of [ITU-T P.79], bands 4-17, m = 0.175 and the weighting factors in the receiving direction according to Table 1 of [ITU-T P.79].

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

5) The test is repeated for the maximum volume control setting.

11.3.4 Variation of loudness rating in sending direction

11.3.4.1 Requirements

For acoustical signal level variation in the range of –3 dB/+ 6 dB from the nominal signal level, the measured SLR shall not deviate more than ±0.5 dB from the SLR measured with the nominal signal level.

NOTE – It is recognized that, under certain conditions, the use of AGC not fulfilling the requirements stated above is useful. This, e.g., may be under certain network conditions. Under such conditions, the linearity requirement may not be appropriate.

11.3.4.2 Test

1) The test signal used for the measurements shall be artificial voice according to [ITU-T P.50]. The test signals are –7.7 dBPa and +1.3 dBPa, measured at the MRP. The test signal level is the average level of the complete test signal. For speakerphone hands-free terminals, the level at the HATS-HFRP is adjusted to –31.7 dBPa and –22.7 dBPa.

   The measured power density spectrum at the MRP is used as the reference power density spectrum for determining the sending sensitivity.

2) The test arrangement is according to clause 7.1. The sending sensitivity is calculated from each band of the 14 frequencies given in Table 1 of [ITU-T P.79], bands 4-17.

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

3) The sensitivity is expressed in dBV/Pa, the sending loudness rating (SLR) shall be calculated according to equation 5-1 of [ITU-T P.79], bands 4-17, m = 0.175 and the weighting factors in the sending direction according to Table 1 of [ITU-T P.79].

4) For both signal levels, the measured result is compared to the SLR measured nominal signal level.

11.3.5 Variation of loudness rating in receiving direction

11.3.5.1 Requirements

With nominal volume control setting for network signal level variations of ±5 dB from the nominal signal level, the measured RLR shall not deviate more than ±0.5 dB from the RLR measured with nominal signal level and nominal volume control setting.

11.3.5.2 Test

1) The test signal used for the measurements shall be artificial voice according to [ITU-T P.50]. The test signals are –11 dBm0 and –21 dBm0, measured at the electrical reference point and averaged over the complete test signal sequence.

2) The test arrangement is according to clause 7.1. For the measurement of speakerphone hands-free terminals, the artificial head is free-field equalized according to [ITU-T P.581]. The equalized output signal of the inboard ear is used for the measurement. For headset hands-free terminals, the sound pressure is measured at the DRP and corrected to the ERP according to [ITU-T P.57]. The receiving sensitivity is determined by the bands 4-17 according to Table 1 of [ITU-T P.79].

   For the calculation, the average signal level of each frequency band is referred to the signal level of the reference signal measured in each frequency band.

3) The sensitivity is expressed in dBPa/V, the receiving loudness rating (RLR) shall be calculated according to equation 5-1 of [ITU-T P.79], bands 4-17, m = 0.175 and the weighting factors in the receiving direction according to Table 1 of [ITU-T P.79].

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

5) For both signal levels, the measured result is compared to the RLR measured with nominal signal level.

11.4 Sensitivity frequency responses

11.4.1 Sending sensitivity frequency response

11.4.1.1 Requirements

The sending sensitivity frequency response is measured from the MRP to the output of the speech codec at the electrical point (output of the system simulators, POI).

The tolerance mask for the sending sensitivity frequency response is shown in Table 11-1, the mask is drawn by straight lines between the breaking points in Table 11-1 on a logarithmic (frequency) – linear (dB sensitivity) scale.

Table 11-1 – Tolerance mask for the sending sensitivity frequency response

Frequency [Hz]	Upper limit	Lower limit
200	0	
250	0	
315	0	–14
400	0	–13
500	0	–12
630	0	–11
800	0	–10
1 000	0	–8
1 300	2	–8

Table 11-1 – Tolerance mask for the sending sensitivity frequency response

Frequency [Hz]	Upper limit	Lower limit
1 600	3	–8
2 000	4	–8
2 500	4	–8
3 100	4	–8
4 000	0	
NOTE – All sensitivity values are expressed in dB on an arbitrary scale.		

11.4.1.2 Test

1) The test signal used for the measurements shall be artificial voice according to [ITU-T P.50]. The test signal is equalized at the MRP, the test signal level is –4.7 dBPa at the MRP. The test signal level is the average level of the complete test signal. For speakerphone hands-free terminals, the level at the HATS-HFRP is adjusted to –28.7 dBPa.

   The measured power density spectrum at the MRP is used as the reference power density spectrum for determining the sending sensitivity.

2) The test arrangement is according to clause 7.1. The sending sensitivity frequency response is determined in one-third octave bands as given by the R.40-series of preferred numbers in [ISO 3] for frequencies from 100 Hz to 4 kHz, inclusive. In each one-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.

3) The sensitivity is determined in dBV/Pa.

11.4.2 Receiving sensitivity frequency response

11.4.2.1 Requirements

The receiving sensitivity frequency response is measured from the electrical reference point (input of the system simulators, POI) to the ERP when headset hands-free terminals are measured. For speakerphone hands-free terminals, the sound pressure of the free-field equalized HATS is measured.

The tolerance mask for the receiving sensitivity frequency response is shown in Table 11-2, the mask is drawn by straight lines between the breaking points in Table 11-2 on a logarithmic (frequency) – linear (dB sensitivity) scale.

Table 11-2 – Tolerance mask for the receiving sensitivity frequency response

Frequency [Hz]	Upper limit	Lower limit
200	0	
250	0	
315	0	–15
400	0	–12
500	0	–12
630	0	–12
800	0	–12

Table 11-2 – Tolerance mask for the receiving sensitivity frequency response

Frequency [Hz]	Upper limit	Lower limit
1 000	0	–12
1 300	0	–12
1 600	0	–12
2 000	0	–12
2 500	0	–12
3 100	0	–12
4 000	0	
NOTE – All sensitivity values are expressed in dB on an arbitrary scale.		

11.4.2.2 Test

1) The test signal used for the measurements shall be artificial voice according to [ITU-T P.50]. The test signal is –16 dBm0, measured at the electrical reference point and averaged over the complete test signal sequence.
2) The test arrangement is according to clause 7.1. For the measurement of speakerphone hands-free terminals, the artificial head is free-field equalized according to [ITU-T P.581]. The equalized output signal of the inboard ear is used for the measurement. For headset hands-free terminals, the sound pressure is measured at the DRP and corrected to the ERP according to [ITU-T P.57]. The receiving sensitivity frequency response is determined in one-third octaves as given by the R.40-series of the preferred numbers in [ISO 3] for frequencies from 100 Hz to 4 kHz, inclusive. In each one-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.
3) The sensitivity is determined in dBPa/V.

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

11.5 Speech quality during single talk

11.5.1 One-way speech quality in sending

11.5.1.1 Requirement

The nominal values for the speech quality measured from/to the electrical reference point (POI) shall be:

$$\text{MOS-LQON} \geq 3.0$$

11.5.1.2 Test

A test method for measuring the one-way speech quality via the acoustic interface is currently under study. A possible, non-normative test procedure, is described in Appendix I.

11.5.2 One-way speech quality in receiving

11.5.2.1 Requirement

The nominal values for the speech quality measured from/to the electrical reference point (POI) shall be:

$$\text{MOS-LQON} \geq 3.0$$

11.5.2.2 Test

A test method for measuring the one-way speech quality via the acoustic interface is currently under study. A possible, non-normative test procedure, is described in Appendix I.

11.6 Idle channel noise

All tests are conducted with average RF-signal power settings. It is recommended to check the requirement, in addition with different RF-power settings. The requirement should be fulfilled for all RF-power settings.

For the measurements, it is desirable to have a background noise level of less than –74 dBPa(A) inside the car, a minimum background noise level of –64 dBPa(A) should not be exceeded.

11.6.1 Idle channel noise in sending direction

11.6.1.1 Requirements

The maximum idle channel noise in the sending direction, measured at the electrical reference point (POI) in quiet conditions shall be less than –64 dBm0(P).

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

11.6.1.2 Test

1) For the measurement, no test signal is used. In order to ensure a reliable activation of the terminal, an activation signal is inserted before the actual measurement. The activation signal consists of a sequence of 4 composite source signals according to [ITU-T P.501]. The spectrum of the test signal at the MRP is equalized under free-field conditions. The level of the activation signal is –28.7 dBPa, measured at the HATS-HFRP.

2) The test arrangement is described in clause 7.1.

   The idle channel noise is measured at the electrical reference point in the frequency range between 100 Hz and 4 kHz. The measurement requires a time window which starts exactly at the end of the activation signal. Any "ringing" of filters or receivers or reverberance influence shall be taken into account, the time window must be shifted accordingly. The length for the time window is 1 second, which is the averaging time for the idle channel noise. The test lab has to ensure that the terminal is activated during the measurement. If the terminal is deactivated during the measurement, the measurement window has to be cut to the duration while the terminal remains activated.

   The power density spectrum of the noise signal is determined using FFT (8 k samples/48 kHz sampling rate or equivalent). A Hanning window is used.

3) The idle channel noise is determined by psophometric weighting. Spectral peaks are measured in the frequency domain. The average noise spectrum used for determining the spectral peak should be calculated as the arithmetic mean of the noise spectrum values when stated in dBm0(P).

11.6.2 Idle channel noise in receiving direction

11.6.2.1 Requirements

The requirements for the maximum noise produced by the hands-free terminal in case no signal is applied to the receiving direction are as follows:

– If a user-specific volume control is provided, it is adjusted to the RLR value close to the nominal value. Hands-free terminals without user-specific volume controls are measured in normal operating conditions. The idle channel noise level measured at the DRP shall be less than –53 dBPa(A).

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

11.6.2.2 Test

1) For the measurements, no test signal is used. In order to ensure a reliable activation of the terminal, an activation signal is inserted before the actual measurement. The activation signal consists of a sequence of 4 composite source signals according to [ITU-T P.501]. The level of the activation level is adjusted to –16 dBm0, measured at the electrical reference point. The level of the activation signal is averaged over the complete duration of the activation signal.

2) The test arrangement is according to clause 7.1. For the measurement of speakerphone hands-free terminals, the artificial head is free-field equalized according to [ITU-T P.581]. The equalized output signal of the inboard ear is used for the measurement. For headset hands-free terminals, the sound pressure is measured at the DRP and corrected to the ERP according to [ITU-T P.57].

   The idle channel noise is measured at the DRP in the frequency range between 50 Hz and 8 kHz. The measurement requires a time window which starts exactly at the end of the activation signal. Any ringing of filters or receivers or reverberance influence shall be taken into account. The time window must be shifted accordingly. The length of the time window is 1 second, which is the averaging time for the idle channel noise.

   The power density spectrum of the noise signal is determined using FFT (8 k samples/48 kHz sampling rate or equivalent). A Hanning window is used.

3) The idle channel noise is A-weighted. Spectral peaks are measured in the frequency domain. The average noise spectrum used for determining the spectral peak should be calculated as the arithmetic mean of the noise spectrum values when stated in dBPa(A).

11.7 Out-of-band signals

For the measurements, it is desirable to have a background noise level of less than –74 dBPa(A) inside the car, a minimum background noise level of –64 dBPa(A) should not be exceeded.

11.7.1 Discrimination against out-of-band signals in sending direction

11.7.1.1 Requirements

If out-of-band signals are generated at the MRP, they are transmitted to the terminal and possibly to the input of the speech decoder. For signals that are measured at the output of the speech decoder at the electrical reference point (POI), the following requirements shall apply.

For the measurement, a white Gaussian noise band-limited in the frequency range between 4.6 kHz and 8 kHz with a level of –4.7 dBPa at the MRP is used. The total level, measured in a frequency range from 300 Hz to 3.4 kHz, is measured at the electrical reference point (POI) and shall be less than 35 dB referred to the reference level. The reference level is determined using artificial voice according to [ITU-T P.50], band-limited in the frequency range between 300 Hz and 3.4 kHz with a level of –4.7 dBPa at the MRP. For this signal, the in-band level averaged over the complete reference signal length is determined at the electrical reference point.

11.7.1.2 Test

NOTE 1 – Frequency shifting technology should be disabled during this test.

1) The test arrangement is according to clause 7.1.

2) In order to ensure a reliable activation of the hands-free terminal, an activation signal is generated before the actual measurement starts. The activation signal consists of a sequence of 4 composite source signals according to [ITU-T P.501]. The activation level shall be

–4.7 dBPa, measured at the MRP. The level of the activation signal is averaged over the complete activation sequence signal.

3) Directly after the activation signal, the actual test signal is inserted. A test signal is inserted exactly after the voiced sound of the last CSS burst (instead of the pn sequence). The duration of the test signal amounts to 200 ms.

4) The test signal is a white Gaussian noise, band-limited from 4.6 kHz to 8 kHz with a level of –4.7 dBPa at the MRP. The level of the test signal is averaged over the complete test signal sequence.

5) For the analysis, a rectangular window is used which is adapted to the test signal duration (200 ms). Any "ringing" of filters or receivers or reverberance influence shall be taken into account, the time window must be shifted accordingly. The signal level is determined in the frequency range from 300 Hz to 3.4 kHz at the electrical reference point (POI). The level of the reference signal (artificial voice according to [ITU-T P.50], band-limited from 300 Hz to 3.4 kHz, –4.7 dBPa at the MRP) is determined at the electrical reference point (POI) as well.

NOTE 2 – With low sensitivity in the sending direction, the measured noise level may already exceed the required minimum out-of-band level.

11.7.2 Spurious out-of-band signal in receiving direction

11.7.2.1 Requirements

The test signal used is artificial voice according to [ITU-T P.50], band-limited in a frequency range between 300 Hz and 3.4 kHz with a level of –12 dBm0 in the receiving direction. The level of the out-of-band signal is measured in a frequency range between 4.6 and 8 kHz at the hands-free loudspeaker and shall be at least 45 dB below the level of the reference signal. The level of the reference signal is determined by measuring the acoustical level of the in-band signal at the hands-free loudspeaker.

11.7.2.2 Test

NOTE 1 – Bandwidth extension technology should be disabled during this test.

1) The test signal shall be artificial voice according to [ITU-T P.50]. The level of the test signal is averaged over the complete test signal sequence.

2) The output signal of a measurement microphone positioned close to the hands-free loudspeaker is used for the measurement. By this, the S/N of the tests can be improved as compared to measurements conducted with the artificial head. For headset hands-free terminals, the output signal is measured at the DRP and corrected to the ERP according to [ITU-T P.57].

3) The level of the out-of-band signal is determined between 4.6 and 8 kHz. The reference level is determined by measuring the in-band signal at the hands-free loudspeaker.

NOTE 2 – With low sensitivity in the receiving direction, the measured noise level may already exceed the required minimum out-of-band level.

NOTE 3 – This measurement method does not apply to systems including artificial bandwidth extension.

11.8 Distortion in sending

The distortion in sending up to 4 kHz is measured from the MRP to the electrical reference point (input of the system simulator, POI).

NOTE – This test cannot be performed over some networks, such as CDMA, because the speech coder does not pass pure tones.

For the measurements, it is desirable to have a background noise level of less than –74 dBPa(A) inside the car, a minimum background noise level of –64 dBPa(A) should not be exceeded.

11.8.1 Requirement

The harmonic distortion in sending should be not higher than 3%.

11.8.2 Test

NOTE – Signal processing that could invalidate the test should be disabled during this test.

1) The test signal is a sinusoidal signal with a frequency of 300 Hz, 500 Hz and 1 kHz. The test signal level is –4.7 dBPa. In order to guarantee a reliable activation of the hands-free terminal, a sequence of 4 composite source signals according to [ITU-T P.501] is sent to the terminal before the actual test signal. The activation signal level is –4.7 dBPa, measured at the MRP. The activation signal level is averaged over the total length of the activation signal.
2) The test signal is inserted immediately after the activation sequence, after the voiced sound of the last CSS burst (instead of the pn sequence). The test signal duration is 200 ms.
3) For the analysis, a Hanning window is used which is adapted to the duration of the test signal (200 ms).
4) The harmonic distortion produced by the hands-free terminal is measured at the electrical reference point.

11.9 Distortion in receiving

The distortion in receiving is measured from the POI to the artificial ear up to 8 kHz.

NOTE – This test cannot be performed over some networks, such as CDMA, because the speech coder does not pass pure tones.

For the measurements, it is desirable to have a background noise level of less than –74 dBPa(A) inside the car, a minimum background noise level of –64 dBPa(A) should not be exceeded.

11.9.1 Requirements

The distortion in receiving is measured from the POI to the DRP.

NOTE – If available, a linear access point at the hands-free system where no non-linear and time-variant signal processing except speech coding is present can be used alternatively for the test.

The harmonic distortion should be less than 3% when producing a sound pressure level needed to achieve an S/N ≥6 dB (see clause 11.3.1) and for the maximum volume control setting.

This test is applicable if a linear access point without any non-linear signal processing to the loudspeaker amplifier is available. If this access point is not available, the measurement may be conducted with some care since non-linear processing may influence the test result.

11.9.2 Test

NOTE – Signal processing (e.g., bandwidth extension techniques, etc.) that could invalidate the test should be disabled during this test.

1) The test signal is a sinusoidal signal with a frequency of 300 Hz, 500 Hz and 1 kHz. The test signal level is the level measured at the linear access point when inserting a test signal with –16 dBm0 at the POI. In order to guarantee a reliable activation of the hands-free terminal, a sequence of 4 composite source signals according to [ITU-T P.501] is sent to the terminal before the actual test signal. The activation signal level is the level equivalent to the level when inserting a test signal at the POI with –16 dBm0, measured at the linear access point. The activation signal level is averaged over the total length of the activation signal.
2) The test signal is inserted immediately after the activation sequence, after the voiced sound of the last CSS burst (instead of the pn sequence). The test signal duration is 200 ms.

3) For the analysis, a Hanning window is used which is adapted to the duration of the test signal (200 ms).

4) The harmonic distortion is measured for each test signal frequency.

11.10 Echo performance without background noise

Due to the expected delay in networks, the echo loss presented at the electrical reference point (POI) should be at least 50 dB during single talk. This echo loss (TCL_W) should be achieved for a wide range of acoustical environments and delays.

NOTE – When realizing echo loss by speech-activated attenuation/gain control, "comfort noise" should be inserted in case the signal is completely suppressed.

For the measurements, it is desirable to have a background noise level of less than –74 dBPa(A) inside the car, a minimum background noise level of –64 dBPa(A) should not be exceeded.

11.10.1 Terminal coupling loss (TCLw)

11.10.1.1 Requirements

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 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 the sending 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 are observed, although the TCLw 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.

11.10.1.2 Test

1) All tests are conducted in the car cabin, the test arrangement is described in clause 7.1. 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.

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

3) The test signal is a pn sequence according to [ITU-T P.501] with a length of 4096 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°.

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

11.10.2 Echo level versus time

11.10.2.1 Requirements

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

NOTE 1 – The echo path is kept constant during this test, and the test should begin 5 seconds after the initial application of a reference signal such that a steady state converged condition is achieved.

NOTE 2 – The analysis is conducted only during the active signal part.

11.10.2.2 Test

1) The test arrangement is according to clause 7.1.
2) The test signal consists of a periodically repeated composite source signal according to [ITU-T P.501] 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. In addition, the test is repeated with artificial voice according to [ITU-T P.50]. One sequence of male and one sequence of female voice are used. The average test signal level is –16 dBm0. The echo signal is analysed during the complete test signal. The integration time for the level analysis shall be 35 ms, the analysis is referred to the level analysis of the reference signal.
3) The measurement result is displayed as attenuation versus time. The exact synchronization between input and output signal has to be guaranteed.

NOTE – When testing using CSS, 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 of the level analysis (35 ms).

11.10.3 Spectral echo attenuation

11.10.3.1 Requirements

The echo attenuation versus frequency shall be below the tolerance mask given in Table 11-3.

Table 11-3 – Spectral echo attenuation mask

Frequency [Hz]	Upper limit
100	–20
200	–30
300	–38
800	–34
1500	–33
2600	–24
4000	–24
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 may be inserted in the sending direction in order to mask the echo signal.

NOTE – This requirement should be fulfilled at any point in time. Therefore, it should be verified at different time intervals of the test sequence.

11.10.3.2 Test

1) The test arrangement is according to clause 7.1.
2) Before the actual measurement, a training sequence is fed in consisting of 10 seconds CS signal according to [ITU-T P.501]. The level of the training sequence is –16 dBm0.
3) 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. Four 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 with 8 k points (48 kHz sampling rate or equivalent, rectangular window).
4) The spectral echo attenuation is analysed in the frequency domain in dB.

11.10.4 Initial convergence without background noise

11.10.4.1 Requirements

The initial convergence (echo attenuation versus time) during single talk immediately after activating the hands-free terminal with maximum volume control setting should conform to the requirement shown in Figure 11-3.

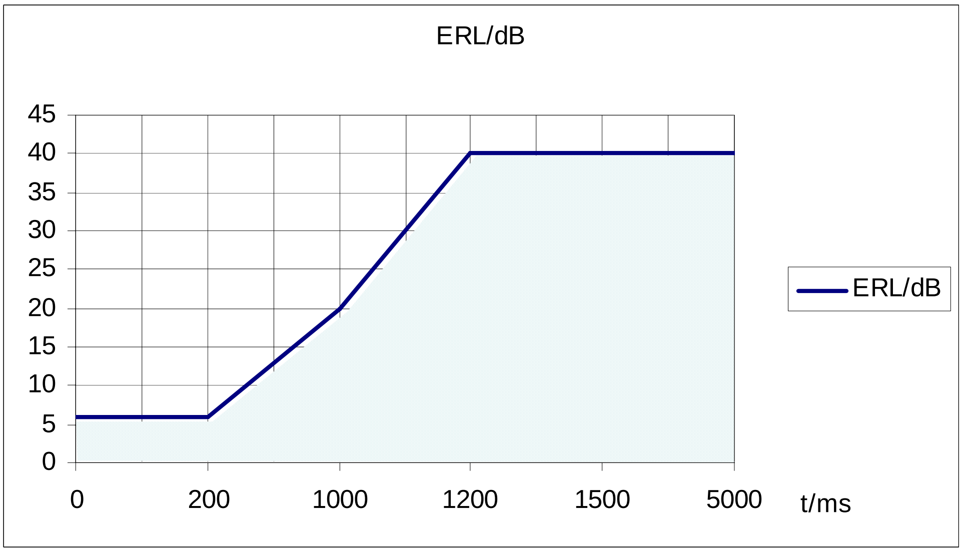


Figure 11-3 – Initial convergence, ERL versus time

11.10.4.2 Test

1) The test arrangement is described in clause 7.1. The noise level measured at the electrical access point (idle channel noise) shall be less than –63 dBm0.
2) The test signal is applied immediately after setting up the call and setting the volume control to its maximum.
3) The test signal is a composite source signal according to [ITU-T P.501], repeated periodically. The average signal level is –16 dBm0. The echo signal is analysed over a period of at least 5 s. The analysis integration time is 35 ms, the analysis is referred to the level analysis of the reference signal. In addition, the test is repeated with artificial voice according to [ITU-T P.50]. One sequence of male and a second one consisting of a female

voice are used. The starting point of each signal is as defined by the start of the sequence provided by [ITU-T P.50]. The average test signal level is –16 dBm0. The echo signal is analysed over a period of at least 5 s. The analysis integration time is 35 ms, the analysis is referred to the level analysis of the reference signal.

4) The measurement is displayed as echo attenuation versus time, measured signal and reference signal have to be synchronized in time.

NOTE 1 – The analysis of the CSS is performed only on the active signal parts, the pauses between the bursts of the composite source signal are not analysed. The analysis time is reduced by the time constant of the level analysis due to the integration time of 35 ms.

NOTE 2 – The required performance for artificial voice signals should be achieved for different starting points of the artificial voice signal.

11.10.5 Initial convergence with background noise

11.10.5.1 Requirements

The initial convergence (echo attenuation versus time) during single talk immediately after activating the hands-free terminal with background noise and with maximum volume control setting should conform to the requirement in Figure 11-4.

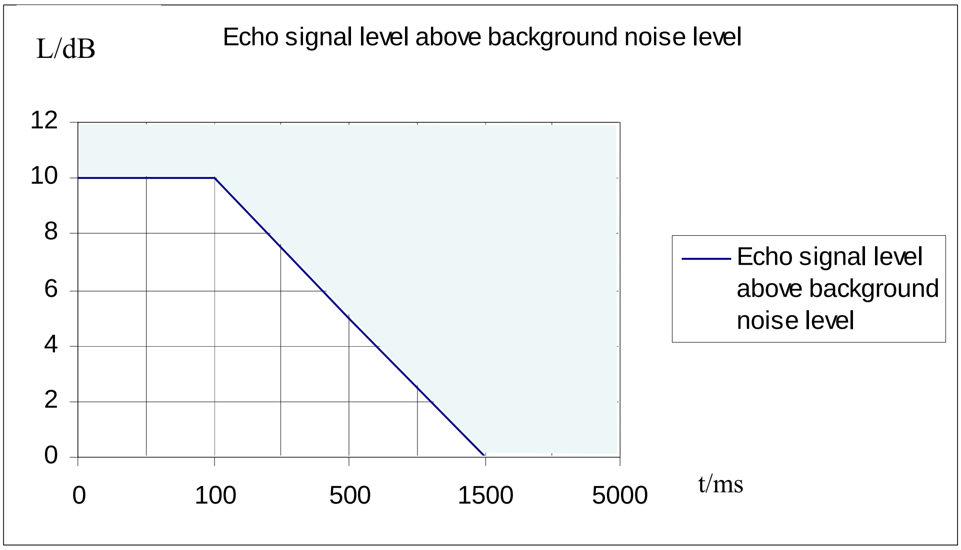


Figure 11-4 – Initial convergence with background noise, requirement on echo signal level versus time

11.10.5.2 Test

1) The test arrangement is described in clause 7.1.

2) The background noise defined by the manufacturer/test lab is played back at least 5 s before the start of the actual measurement. This allows time for some adaptive algorithms in the hands-free unit, which are constantly monitoring the microphone signal, to stabilize (e.g., AGC, NR). The test is conducted under simulated constant driving conditions.

3) The test signal is applied immediately after setting up the call and setting the volume control to its maximum.

4) The test signal is a composite source signal according to [ITU-T P.501] repeated periodically. The average signal level is –16 dBm0. The echo signal is analysed over a period of at least 5 s. The analysis integration time is 35 ms. In addition, the test is repeated with artificial voice according to [ITU-T P.50]. One sequence of male and a second one consisting of a female voice are used. The starting point of each signal is as defined by the

start of the sequence provided by [ITU-T P.50]. The average test signal level is –16 dBm0. The echo signal is analysed over a period of at least 5 s. The analysis integration time is 35 ms.

5) The measurement is displayed as echo attenuation versus time.

NOTE 1 – The analysis of the CSS is performed only on the active signal parts, the pauses between the bursts of the composite source signal are not analysed. The analysis time is reduced by the time constant of the level analysis due to the integration time of 35 ms.

NOTE 2 – The required performance for artificial voice signals should be achieved for different starting points of the artificial voice signal.

11.10.6 Echo performance with time variant echo path

11.10.6.1 Requirements

This test is intended to verify that the system will maintain sufficient echo attenuation with a time variant echo path. The measured echo attenuation measured with a time variant echo path should not decrease by more than 6 dB from the maximum measured during the test.

The time variant echo path is realized by a rotating 30 cm × 40 cm reflective surface (e.g., a piece of cardboard, wood or plastic) positioned on the co-drivers seat. The initial state of the reflecting surface (i.e., 0° position) is such that it is in the median plane (perpendicular to the front of the vehicle) with a bottom-to-top height of 40 cm, a front-to-back length of 30 cm, and the centre of the reflecting surface is at a point in the vehicle that is symmetric with the centre of the HATS in the driver's seat. The reflecting surface then pivots 90° such that the most forward edge of the reflecting surface rotates out towards the co-drivers side window; the centre of the reflecting surface serves as the axis point and stays in the same location during this rotation. At the 90° position, the reflecting surface is in the frontal plane (parallel with the front of the vehicle). The reflecting surface continuously rotates between the 0° and 90° positions during the measurements at a rate of 90°/second. The rotation of the reflecting plane is time-synchronized with the test signals by means of a control channel.

11.10.6.2 Test

1) Before conducting the test, the echo canceller should be fully converged.

2) The test arrangement is according to clause 7.1.

3) Rotation of the reflecting surface is time-synchronized with the playback of the test signals. The reflecting surface starts rotating from an initial position of 0° as soon as the test signals start playback.

4) The test signal consists of a periodically repeated composite source signal according to [ITU-T P.501] 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. In addition, the test is repeated with artificial voice according to [ITU-T P.50]. A sequence of male and female voices is used. The average test signal level is –16 dBm0. The echo signal is analysed during the complete test signal. The integration time for the level analysis shall be 35 ms, the analysis is referred to the level analysis of the reference signal.

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

NOTE – When using the CSS, 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 of the level analysis (35 ms).

11.10.7 Switching characteristics

11.10.7.1 Activation in sending direction

The activation in the sending 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 the sending 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 below is always referred to the test signal level at the mouth reference point (MRP).

11.10.7.1.1 Requirements

The minimum activation level $L_{S,min}$ should be ≤–20 dBPa.

The built-up time $T_{r,S,min}$ (measured with minimum activation level) should be ≤50 ms.

11.10.7.1.2 Test

The structure of the test signal is shown in Figure 11-5. The test signal consists of CSS components according to [ITU-T P.501] with increasing level for each CSS burst.

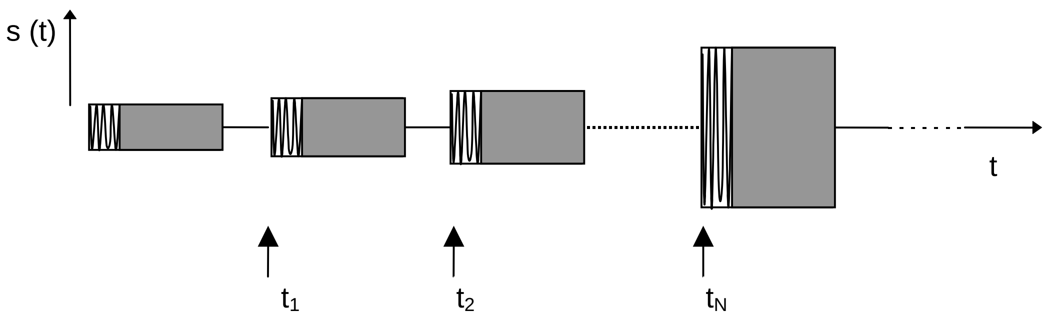


Figure 11-5 – Test signal to determine the minimum activation level and the built-up time

The settings of the test signal are as follows.

Table 11-4 – Settings of the CSS in Sending

	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 sending direction	248.62 ms/451.38 ms	–23 dBPa (Note 1)	1 dB
NOTE 1 – The level of the active signal part corresponds to an average level of –24.7 dBPa at the MRP for the CSS according to [ITU-T P.501] assuming a pause of 101.38 ms. NOTE 2 – When testing a speakerphone hands-free system, the signal level is corrected at the HATS-HFRP.			

It is assumed that the pause length of 451.38 ms is longer than the hang-over time so that the test object is back to idle mode after each CSS burst.

1) The test arrangement is described in clause 7.1.

2) 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 versus time. The levels are calculated from the time domain using an integration time of 5 ms.

3) 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 the one syllable word "test" instead of the CS signal. The word used should be of similar duration, the average level of the word must be adapted to the CS signal level of the according CSS burst.

11.10.7.2 Activation in receiving direction

The activation in the receiving direction is mainly determined by the built-up time $T_{r,R,min}$ and the minimum activation level ($L_{R,min}$). The minimum activation level is the level required to remove completely any attenuation inserted during the idle mode. The built-up time is determined from the level variation of the transmitted test signal which is applied with a minimum activation level.

The activation level described below is always referred to the test signal level at the electrical reference point (POI).

In order to guarantee a higher accuracy when recording the transmitted signal in the receiving direction, a measurement microphone is used for this test and positioned close to the loudspeaker of the hands-free terminal.

11.10.7.2.1 Requirements

The minimum activation level $L_{R,min}$ should be ≤–35.7 dBm0 (measured during the active signal part).

The built-up time $T_{r,R,min}$ (measured with minimum activation level) should be ≤50 ms.

11.10.7.2.2 Test

The signal construction is shown in Figure 11-5. The test signal settings are as follows.

Table 11-5 – Settings of the CSS in receiving

	CSS duration/pause duration	Level of the first CS signal (active signal part at the POI)	Level difference between two periods of the test signals
CSS to determine switching characteristics in receiving direction	248.62 ms/451.38 ms	–38.7 dBm0 (Note)	1 dB
NOTE – The level of the active signal part corresponds to an average level of –40 dBm0 at the POI for the CSS according to [ITU-T P.501] assuming a pause of 101.38 ms.			

1) The test arrangement is according to clause 7.1.

2) The transmitted signal is recorded by a microphone positioned close to the loudspeaker. For headset hands-free terminals, the sound pressure is measured at the DRP and corrected to the ERP according to [ITU-T P.57]. The measured signal level is referred to the test signal level and displayed versus time. The integration time of the level analysis used should be 5 ms.

3) The minimum activation level is determined from the CSS burst indicating the first activation of the test object. The duration between the beginning of this CSS burst and the complete activation of the terminal 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 must be adapted to the CS signal level of the according CSS burst.

11.10.7.3 Attenuation range in sending direction

The attenuation range in the sending direction is determined by applying the test signal in the sending direction after the terminal was activated in the receiving direction. During the measurement, the attenuation range in the sending direction ($A_{H,S}$) and the built-up time in the sending direction ($T_{r,S}$) is determined.

11.10.7.3.1 Requirements

The attenuation range $A_{H,S}$ should be less than 20 dB.

The built-up time $T_{r,S}$ should be less than 50 ms. It is recommended to reduce the attenuation within 15 ms to at least 13 dB below the final value.

11.10.7.3.2 Test

The structure of the test signals is shown in Figure 11-6. It consists of periodically repeated composite source signal bursts used for activating the receiving direction and the voiced sound used to measure the sending direction.

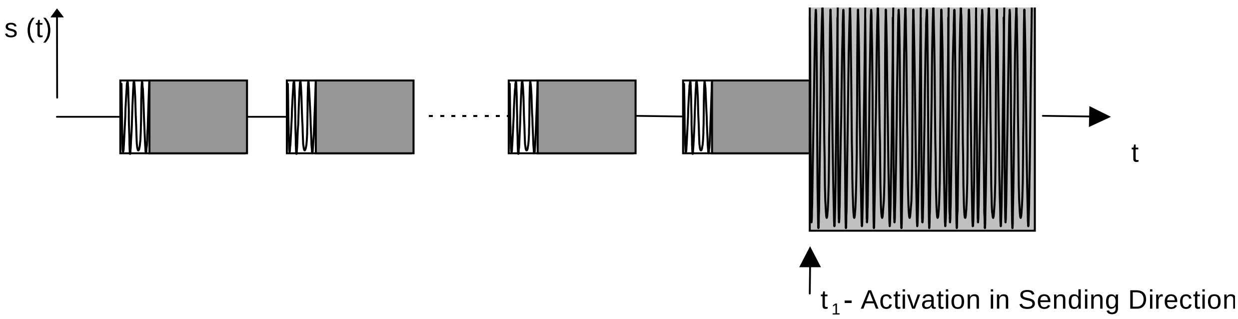


Figure 11-6 – Structure of the test signal for measuring the attenuation range

The test signals are synchronized in time at the acoustical interface. The delay of the test arrangement should be constant during the measurement.

11.10.7.3.3 Test

1) The test arrangement is according to clause 7.1.

2) The test signal used is according to Figure 11-6, the receiving direction is activated first. The measurement parameters are as follows.

Table 11-6 – Signal levels for double talk tests in sending and receiving

	Receiving direction	**Sending direction (at the MRP)**
Average signal level (including 101.38-ms pauses)	–16 dBm0	---
Active signal part	–14.7 dBm0	–3 dBPa

The level in the receiving direction is determined at the electrical reference point.

3) The level is determined as level versus time calculated from the time domain. The integration time of the levels analysis is 5 ms. The attenuation range is determined by calculating the difference between the measured level between the beginning of the test signal in the sending direction (t_1 in Figure 11-6) until complete activation in the sending direction.

11.10.7.4 Attenuation range in receiving direction

The attenuation range in the receiving direction is determined after the terminal was activated in the sending direction before. During the measurement, the attenuation range in the receiving direction ($A_{H,R}$), as well as the built-up time in receiving direction ($T_{r,R}$), are determined.

11.10.7.4.1 Requirements

The attenuation $A_{H,R}$ should be less than 15 dB.

The built-up time $T_{r,R}$ should be less than 50 ms. It is recommended to reduce the attenuation within 15 ms to less than 9 dB.

11.10.7.4.2 Test

The structure of the test signal is shown in Figure 11-6. Again, CSS bursts are used for activating the opposite direction (now the sending direction) and the voiced sound is used to measure the receiving direction. The test signals are synchronized in time at the acoustical interface. The delay of the test arrangement should be constant during the measurement.

1) The test arrangement is according to clause 7.1.

2) The test signal shown in Figure 11-6 is used, the sending direction is activated first.

The measurement parameters are as follows.

Table 11-7 – Signal levels for double talk tests in sending and receiving

	Receiving direction	**Sending direction (at the MRP)**
Average level (including 101.38-ms pauses)	---	–4.7 dBPa
Active signal part	–14.7 dBm0	–3 dBPa

The level in the receiving direction is determined at the electrical reference point.

3) The level is determined as level versus time calculated from the time domain. The integration time of the level analysis is 5 ms. The attenuation is determined by calculating the difference between the beginning of the measured test signal in the receiving direction (t_1 in Figure 11-6) and the complete activation in the receiving direction.

11.11 Double talk performance

NOTE – Before starting the double talk tests, the test lab should ensure that the echo canceller is fully converged. This can be done by an appropriate training sequence.

During double talk, the speech is mainly determined by two 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 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 P.340] and [ITU-T P.502]):

– Attenuation range in sending direction during double talk $A_{H,S,dt}$.
– Attenuation range in receiving direction during double talk $A_{H,R,dt.}$
– Echo attenuation during double talk.

11.11.1 Attenuation range in sending direction during double talk: $A_{H,S,dt}$

11.11.1.1 Requirements

Based on the level variation in the sending direction during double talk, $A_{H,S,dt}$, the behaviour of hands-free terminals can be classified according to Table 11-8.

Table 11-8 – Categorization of double talk capability according to [ITU-T P.340]

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

The requirements apply for nominal and maximum setting of the receive volume control.

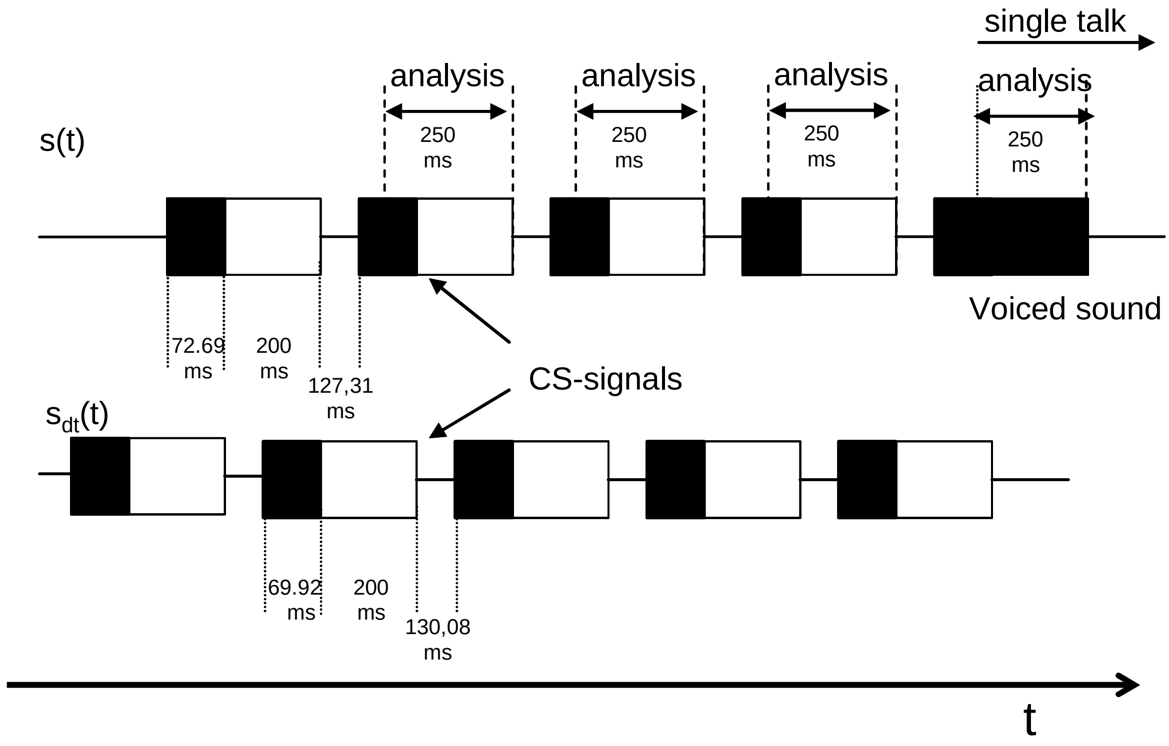
The requirements apply for nominal signal levels in the sending and receiving directions as well as for the level combinations +6 dB (re. nominal level) in the sending/–6 dB (re. nominal level) in receiving; and +6 dB (re. nominal level) in receiving/–6 dB (re. nominal level) in sending. Furthermore, the test is conducted with nominal levels but with maximum setting of the volume control.

NOTE – If the maximum setting of the volume control is chosen such that non-linearities occur in the echo path, the double talk performance will decrease.

In general, Table 11-8 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.

11.11.1.2 Test

The test signal to determine the attenuation range during double talk is shown in Figure 11-7. A sequence of uncorrelated CS signals is used, which is inserted in parallel in the sending and receiving directions.



s(t) – Signal for one direction

$s_{dt}(t)$ – Double talk signal

Signal timing for attenuation range in sending

Figure 11-7 – Double talk test sequence with overlapping CS signals in sending and receiving directions

Figure 11-7 indicates that the sequences overlap partially. The beginning of the CSS sequence (voiced sound, black) is overlapped by the end of the pn sequence (white) in the opposite direction. During the active signal parts of one signal, the analysis can be conducted in the sending and receiving directions. The analysis times are also shown in Figure 11-7. 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 11-9 – Timing of the double talk sequences

	Receiving direction	Sending direction
Voiced sound	69.92 ms (Note 2)	72.69 ms (Note 1)
Pseudo-random noise sequence/ noise sequence	200 ms	200 ms
Pause length between two signal bursts	130.08 ms	127.31 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 1 – 14 repetitions of the voiced sound for double talk according to [ITU-T P.501]. NOTE 2 – 23 repetitions of the voiced sound according to [ITU-T P.501].		

The tests are repeated with maximum volume control setting in receiving direction.

1) The test arrangement is according to clause 7.1. Before the actual test, a training sequence for the echo canceller consisting of 10 s male and 10 s female voice according to [ITU-T P.50] with a level of –16 dBm0 is applied to the electrical reference point.
2) When determining the attenuation range in the sending direction, the signal measured at the electrical reference point is referred to the test signal inserted.
3) The level is determined as level versus 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 the sending direction until its complete activation (during the pause in the receiving 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.
4) The test is repeated for all level combinations as defined in the requirements.

11.11.2 Attenuation range in receiving direction during double talk: $A_{H,R,dt}$

To ensure higher accuracy measuring the transmitted signal in the receiving direction, a measurement microphone is used which is positioned as close as possible to the loudspeaker of the hands-free terminal.

11.11.2.1 Requirements

Based on the level variation in the receiving direction during double talk, $A_{H,R,dt,}$ the behaviour of the hands-free terminal can be classified according to Table 11-10.

Table 11-10 – Categorization of double talk capability according to [ITU-T P.340]

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

The requirements apply for nominal and maximun setting of the receive volume control.

In general, Table 11-10 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.

11.11.2.2 Test

The test signal to determine the attenuation range during double talk is shown in Figure 11-7. A sequence of uncorrelated CS signals is used which is inserted in parallel in the sending and receiving directions. 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 11-11 – Timing of the double talk sequences

	Receiving direction	Sending direction
Voiced sound	69.92 ms	72.69 ms
Pseudo-random noise sequence	200 ms	200 ms
Pause length between two signal bursts	130.08 ms	127.31 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

The tests are repeated with maximum volume control setting in the receiving direction.

1) The test arrangement is according to clause 7.1.
2) When determining the attenuation range in the receiving direction, the signal measured at the loudspeaker of the hands-free terminal is referred to the test signal inserted. For headset hands-free terminals, the sound pressure is measured at the DRP and corrected to the ERP according to [ITU-T P.57].
3) The level is determined as level versus 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 the receiving direction until its complete activation (during the pause in the sending 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.
4) The test is repeated for all level combinations as defined in the requirements.

11.11.3 Detection of echo components during double talk

11.11.3.1 Requirements

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). "Echo loss" is the echo suppression provided by the hands-free terminal measured at the electrical reference point. Under these conditions, the requirements given in Table 11-12 are applicable (more information can be found in Annex A of [ITU-T P.340]).

Table 11-12 – Categorization of double talk capability according to [ITU-T P.340]

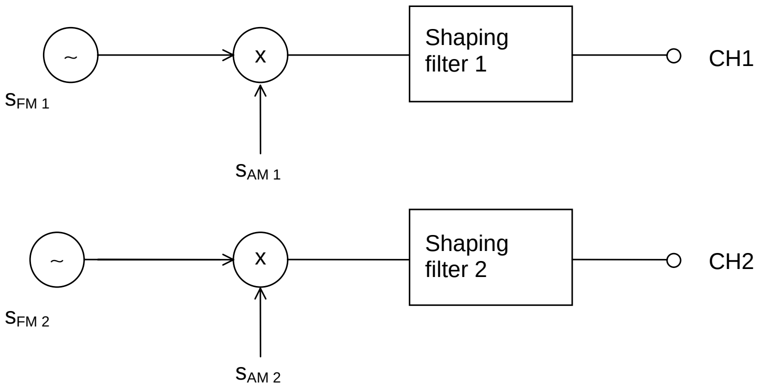
Category (according to [ITU-T P.340])	1	2a	2b	2c	3
	Full duplex capability	Partial duplex capability			No duplex capability
Echo loss [dB]	≥27	≥23	≥17	≥11	<11

11.11.3.2 Test

1) The test arrangement is according to clause 7.1.

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 Figure 11-8. A detailed description can be found in [ITU-T P.501].

The signals are fed simultaneously in the sending and receiving directions. The level in the sending direction is –4.7 dBPa at the MRP (nominal level), the level in the receiving direction is –16 dBm0 at the electrical reference point (nominal level).



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

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

The settings for the signals are as follows:

Receiving direction			**Sending direction**		
f_m[Hz]	$f_{mod(fm)}$[Hz]	F_{am}[Hz]	f_m[Hz]	$f_{mod(fm)}$[Hz]	F_{am}[Hz]
250	±5	3	270	±5	3
500	±10	3	540	±10	3
750	±15	3	810	±15	3
1000	±20	3	1080	±20	3
1250	±25	3	1350	±25	3
1500	±30	3	1620	±30	3
1750	±35	3	1890	±35	3
2000	±40	3	2160	±35	3
2250	±40	3	2400	±35	3
2500	±40	3	2900	±35	3
2750	±40	3	3150	±35	3
3000	±40	3	3400	±35	3
3250	±40	3	3650	±35	3
3500	±40	3	3900	±35	3
3750	±40	3			

Parameters of the shaping filter: Low pass filter, 5 dB/oct.

Figure 11-8 – Parameters of the two test signals for double talk measurement based on AM-FM modulated sine waves

3) The test signal is measured at the electrical reference point (sending 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 a comb filter using mid-frequencies and

bandwidth according to the signal components of the signal in the receiving direction (see [ITU-T P.501]). The filter will suppress frequency components of the double talk signal.

4) In each frequency band which is used in the receiving 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 11-12. The echo attenuation is to be achieved for each individual frequency band from 200 Hz to 3450 Hz according to the different categories.

11.11.4 Sent speech attenuation during double talk

11.11.4.1 Requirements

The sent speech attenuation during double talk is based on the parameter $A_{H,S,dt}$.

Based on the level variation in the sending direction during double talk, $A_{H,S,dt}$, the behaviour of hands-free terminals can be classified according to the following table.

Table 11-13 – Categorization of double talk capability according to [ITU-T P.340]

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

The requirements apply for nominal and maximum setting of the receive volume control.

The requirements apply for nominal signal levels in the sending and receiving directions as well as for the level combinations +6 dB (re. nominal level) in sending/–6 dB (re. nominal level) in receiving; and +6 dB (re. nominal level) in receiving/–6 dB (re. nominal level) in sending. Furthermore, the test is conducted with nominal levels but with maximum setting of the volume control.

In general, Table 11-13 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 test is conducted in addition to the test described in clause 11.11.1 in order to guarantee that no switching device with a short reaction time is classified as a duplex or partially duplex system.

11.11.4.2 Test

1) The test arrangement is according to clause 7.1.

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 Figure 11-8. A detailed description can be found in [ITU-T P.501].

   The signals are fed simultaneously in the sending and receiving directions. The level in the sending direction is –4.7 dBPa at the MRP (nominal level), the level in the receiving direction is –16 dBm0 at the electrical reference point (nominal level).

3) The test signal is measured at the electrical reference point (sending direction). The measured signal consists of the double talk signal which was fed in by the artificial mouth and the echo signal. The double talk signal (send signal) is filtered by a comb filter using mid-frequencies and bandwidth according to the signal components of the signal in the

sending direction (see [ITU-T P.501]). The filter will suppress frequency components of the echo signal.

4) In each frequency band which is used in the sending direction, the sent speech attenuation, $A_{H,S,dt}$, can be measured separately. The requirement for category 1 is fulfilled if in each frequency band the attenuation of the signal in the sending direction is below the required limit. If attenuation is detectable, the classification is based on Table 11-13 above. The sent speech attenuation $A_{H,S,dt}$ is to be achieved for each individual frequency band from 200 Hz to 3550 Hz according to the different categories.

5) The test is repeated for all level combinations as defined in the requirements.

11.12 Background noise transmission

11.12.1 Ambient noise rejection

11.12.1.1 Requirements

The efficiency of noise reduction is calculated as a D-value from DELSM according to [ITU-T G.111]. The result shall be higher than –13 dB. Note that a D-value higher than 0 dB is recommended.

11.12.1.2 Test

1) The test arrangement is according to clause 7.1.

2) According to the specifications of the manufacturer/test lab, a realistic background noise is played back. The noise scenario should represent a constant driving situation. Before the actual test, a background noise and a speech training sequence shall be simultaneously applied in order to allow the noise reduction and automatic gain control algorithms to adapt. The speech training sequence shall consist of 10 s male and 10 s female voice according to [ITU-T P.50] with a level of +1.3 dBPa at the MRP. Background noise shall begin a minimum of 5 seconds before the start of the speech training sequence.

3) The signal in the sending direction is recorded at the electrical reference point (POI). The measurements are carried out in one-third octaves according to [IEC 61260] in the frequency bands 4-17 (200 Hz–4 kHz). In each band, the diffuse field sensitivity S_{si}(diff) is measured. The result is expressed in dBV/Pa.

4) The near-end speech is simulated by applying a composite source signal according to [ITU-T P.501] with a duration of ≥2 CSS periods. The level is adjusted to +1.3 dBPa at the MRP in order to take into account the Lombard effect. The direct sound sensitivity S_{si}(direct) is also calculated in one-third octave bands according to [IEC 61260] in the frequency bands 4-17 (200 Hz-4 kHz). The result is expressed in dBV/Pa.

5) The D-value is calculated according to Annex E, equations E-2 and E-3, of [ITU-T P.79], in the frequency bands 4-17. The coefficients K_i are used according to Table E.1 of [ITU-T P.79].

11.12.2 Background noise transmission after call setup

11.12.2.1 Requirements

The analysis based on the relative approach [b-Sottek] (see Annex B) should not indicate remarkable characteristics exceeding 6 cp/cPa. The first transmitted signal peak in the sending direction should not cause higher excitation than 15 cp/cPa between 300 Hz and 3.4 kHz.

11.12.2.2 Test

1) The test arrangement is given in clause 7.1.
2) According to the specification of the manufacturer/test lab, the background noise is played back. The test should be carried out during a constant driving situation.
3) The terminal is switched off and on again (to provide a reset) and a call is established by the system simulation. The incoming call is answered at the terminal. Special care should be taken not to produce any disturbances or unwanted noise by touching the terminal's housing while answering the incoming call.
4) The transmitted signal in the sending direction is recorded at the POI starting at least 1 s before the call is answered and for at least 7 s after the call is established. The analysis range is chosen at 8 s including an initial pause of 1 s before the call was established.
5) The recorded signal is analysed using the relative approach [b-Sottek].

11.12.3 Quality of background noise transmission (with far-end speech)

11.12.3.1 Requirements

The test is carried out applying the composite source signal in the receiving direction. During and after the end of composite source signal bursts (representing the end of far-end speech simulation), the signal level in the sending direction should not vary by more than 10 dB (during transition to transmission of background noise without far-end speech).

11.12.3.2 Test

1) The test arrangement is according to clause 7.1.
2) According to the specification of the manufacturer/test lab, the background noise is played back. The test should be carried out during a constant driving situation. Before the actual test, background noise and a speech training sequence shall be simultaneously applied in order to allow the noise reduction and automatic gain control algorithms to adapt. The speech training sequence shall consist of 10 s male and 10 s female voice according to [ITU-T P.50] with a level of +1.3 dBPa at the MRP. Background noise shall begin a minimum of 5 seconds before the start of the speech training sequence.
3) First, the measurement is conducted without inserting the signal at the far end. At least 10 s of noise is recorded. The background signal level versus time is calculated using a time constant of 35 ms. This is the reference signal.
4) In a second step, the same measurement is conducted but with insertion of the CS signal at the far end. The 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 the far-end signal. The background noise should be applied for at least 5 seconds in order to allow adaptation of the noise reduction algorithms. After at least 5 seconds, a composite source signal according to [ITU-T P.501] is applied in the receiving direction with a duration of ≥2 CSS periods. The test signal level is –16 dBm0 at the electrical reference point.
5) The sending signal is recorded at the electrical reference point. The test signal level versus time is calculated using a time constant of 35 ms.
6) The level variation in the sending 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 versus time between the reference signal and the signal measured with the far-end signal.

11.12.4 Quality of background noise transmission (with near-end speech)

11.12.4.1 Requirements

The test is carried out applying a simulated speech signal in the sending direction. During and after the simulated speech signal (composite source signal bursts), the signal level in the sending direction should not vary more than 10 dB.

11.12.4.2 Test

1) The test arrangement is according to clause 7.1.
2) According to the specification of the manufacturer/test lab, the background noise is played back. The test should be carried out during a constant driving situation. Before the actual test, a background noise and a speech training sequence shall be simultaneously applied in order to allow the noise reduction and automatic gain control algorithms to adapt. The speech training sequence shall consist of 10 s male and 10 s female voice according to [ITU-T P.50] with a level of +1.3 dBPa at the MRP. Background noise shall begin a minimum of 5 seconds before the start of the speech training sequence.
3) The near-end speech is simulated using the composite source signal according to [ITU-T P.501] with a duration of ≥2 CSS periods. The test signal level is 1.3 dBPa at the MRP. For speakerphone hands-free systems, the HATS-HFRP correction has to be applied.
4) The sending signal is recorded at the electrical reference point. The test signal level versus time is calculated using a time constant of 35 ms.
5) First, the measurement is conducted without inserting the signal at the near end. The signal level is analysed versus time. In a second step, the same measurement is conducted but with the insertion of 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 CSS bursts in the sending direction.

11.12.5 "Comfort noise" injection

This clause is applicable only if comfort noise is inserted by the hands-free unit.

11.12.5.1 Requirements

1) The level of comfort noise is adjusted in a range of +2 and –5 dB to the original (transmitted) background noise. The noise level is calculated with psophometric weighting.
2) The spectral difference between comfort noise and original (transmitted) background noise shall be in-between the mask given through straight lines between the breaking points on a logarithmic (frequency) – linear (dB sensitivity) scale as given in Table 11-14.

Table 11-14 – Requirements for spectral adjustment of comfort noise (mask)

Frequency [Hz]	Upper limit	Lower limit
200	12	–12
800	12	–12
800	10	–10
2000	10	–10
2000	6	–6
4000	6	–6
NOTE – All sensitivity values are expressed in dB on an arbitrary scale.		

11.12.5.2 Test

1) The test arrangement is according to clause 7.1. Background noise is played back.
2) The test signal is applied in the receiving direction consisting of an initial pause of 10 s and a periodical repetition of the composite source signal in the receiving direction (duration ≥10 s) with nominal level to enable comfort noise injection.
3) The transmitted signal is recorded in the sending direction at the electrical reference point.
4) The power density spectrum measured in the sending direction during the initial pause of the test signal (8 k FFT/48 kHz sampling rate or equivalent, averaged over ≥5 s) is referred to the power density spectrum determined during the period with the periodical repetition of the composite source signals in the receiving direction (8 k FFT/48 kHz sampling rate or equivalent, averaged over ≥5 s). Spectral differences between both power density spectra are analysed and compared to the requirements given in Table 11-14.
5) The level of the transmitted signal in the sending direction is determined during the initial pause of the test signal in the receiving direction and referred to the level of the transmitted signal in the sending direction determined during the application of the test signal in the receiving direction. Both levels are calculated using psophometric weighting.

12 Verification of the transmission performance of short-range wireless transmission enabled phones

12.1 Interface definition and calibration

The principle of short-range wireless transmission testing is shown in Figure 12-1.

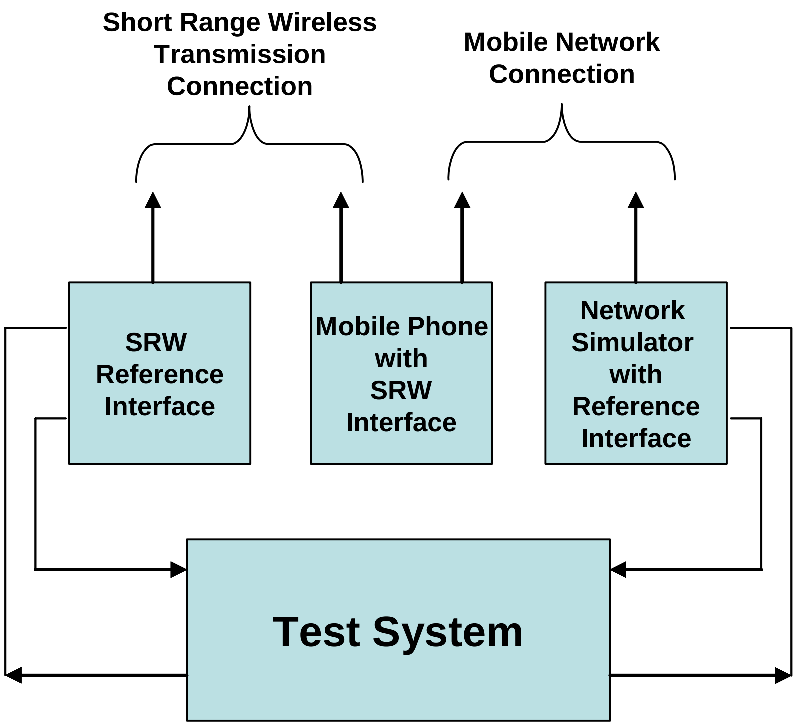


Figure 12-1 – Short range wireless transmission test setup

The mobile phone under test is connected to the short-range wireless transmission reference interface at the short-range wireless transmission reference point (SRWR) and to the network (system) simulator. The short-range wireless transmission reference interface as well as the network simulator reference interface are calibrated to the electrical inputs and outputs of the test system. While the 0 dBr point is clearly defined for the network reference interface, the calibration of the short-range wireless transmission reference interface, in principle, is free. However, in order to get a calibrated set up, the following calibration procedure is recommended.

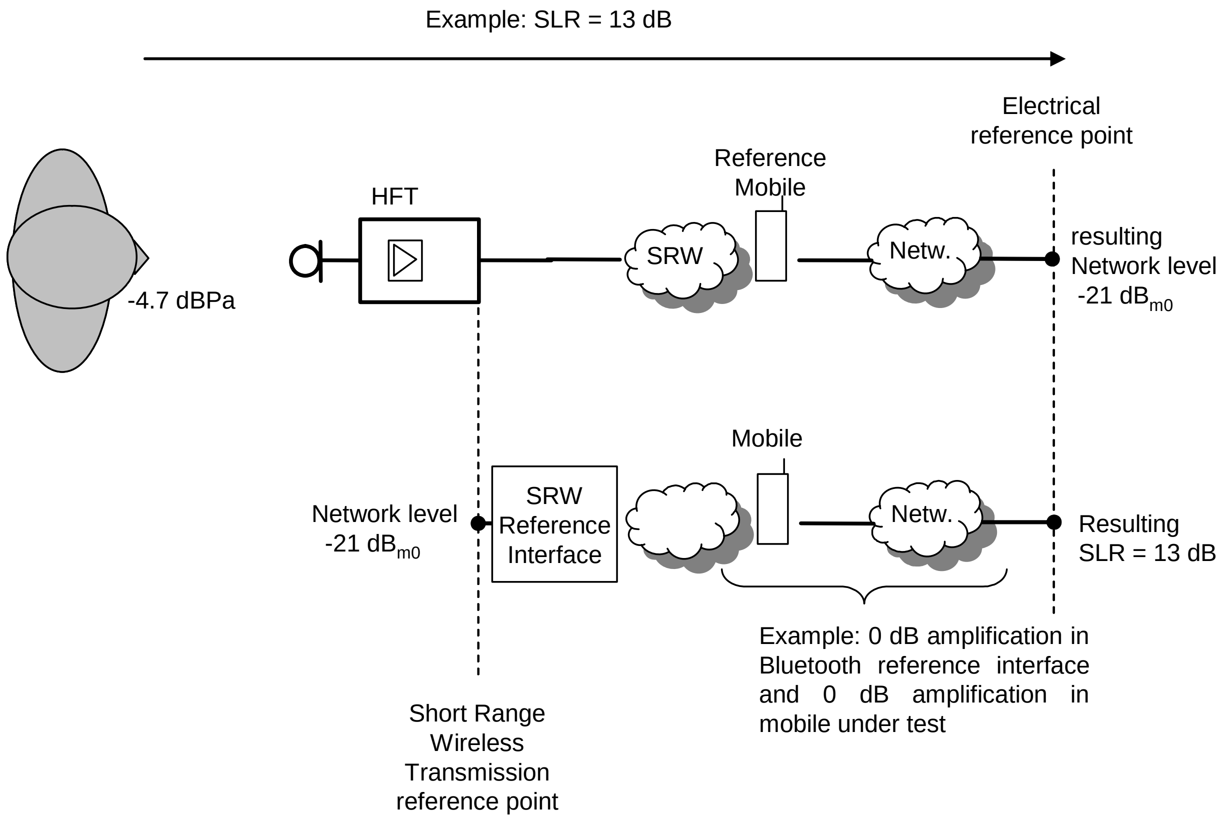


Figure 12-2 – Calibration of the short-range wireless transmission reference interface

The short-range wireless transmission reference interface is calibrated using a mobile phone with known characteristics which does not introduce any amplification or attenuation in the short-range wireless transmission link. No additional signal processing except the short-range wireless transmission coding shall be active. Typically, a "reference phone" is selected by evaluating a variety of different phones of different brands and selecting the one which most closely matches the requirements with respect to sensitivity and deactivated signal processing.

With the reference mobile phone linked to the system simulator and at the same time connected to the hands-free terminal via the SRW link, an SLR test is performed according to the specifications outlined in clause 11.3. The signal level measured at the electrical reference point during this test is the reference signal level, S_{SRWref}, which is to be used when calibrating an arbitrary mobile phone connected to the short-range wireless transmission reference interface.

With an arbitrary mobile phone linked to the system simulator and at the same time connected to the short-range wireless transmission reference interface via the SRW link, the level at the short-range wireless transmission reference point is adjusted until the level measured at the electrical reference point reaches the reference signal level. The corresponding signal level at the short-range wireless transmission reference point is S_{SRWsnd}. The test method is as outlined in clause 11.3. However, the acoustically calibrated excitation signal – properly corrected for attenuation due to the distance between MRP-HFRP – is applied at the short-range wireless transmission reference point rather than the HFRP.

NOTE 1 – The influence of the frequency response characteristic of the hands-free phone is not considered; however, the method ensures a similar overall signal level inserted in the short-range wireless transmission link and the phone.

NOTE 2 – Evaluation boards from short-range wireless transmission chipset vendors may be used for implementation of the short-range wireless transmission reference interface.

12.1.1 SRW delay in sending direction

12.1.1.1 Requirements

The delay in the sending direction is measured from the SRWR to the POI (reference speech codec of the system simulator, output). The delay measured in the sending direction is:

$$T_{SRWs} + t_{System}$$

T_{SRWs} shall be less than 10 ms.

NOTE – The delay T_{SRWs} should be minimized.

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

12.1.1.2 Test

1) For the measurements, a composite source signal (CSS) according to [ITU-T P.501] 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 S_{SRWsnd}.

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

   The test arrangement is according to clause 12.1.

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.

12.1.2 SRW delay in receiving direction

12.1.2.1 Requirements

The delay in the receiving direction is measured from the POI (input of the reference speech coder of the system simulators) to the SRW reference interface. The delay measured in the receiving direction is:

$$T_{SRWr} + t_{System}$$

T_{SRWr} shall be less than 10 ms.

NOTE – The delay T_{SRWr} should be minimized.

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

12.1.2.2 Test

1) For the measurements, a composite source signal (CSS) according to [ITU-T P.501] 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 12.1.

3) The delay is determined by cross-correlation analysis between the measured signal at the SRWR 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.

12.2 SRW loudness ratings

12.2.1 Requirements

The nominal values of JLR from and to the electrical reference point (POI) should be:

$$JLR_{SRWsnd} = 0 \pm 0.5 \text{ dB}$$

$$JLR_{SRWrcv} = 0 \pm 0.5 \text{ dB}$$

However, it is recognized that different short-range wireless transmission implementations deviate significantly (±6 dB) from this level. In order to maintain a good system performance, appropriate corrections have to be integrated in the hands-free implementation to ensure a good system performance. Based on these tests, appropriate level corrections can be inserted.

12.2.2 Test SRW junction loudness rating in sending direction

1) The test signal used for the measurements shall be CSS according to [ITU-T P.501]. The test signal level is S_{SRWsnd}, the level is averaged over the complete test signal.

   The measured power density spectrum SRW reference interface is used as the reference power density spectrum for determining the SRW sending sensitivity.

2) The test arrangement is according to clause 12.1. The SRW sending sensitivity is calculated from each band of the 14 frequencies given in Table A.2 of [ITU-T P.79], bands 4-17.

   For the calculation, the average measured level at the electrical reference point for each frequency band is referred to the average test signal level measured in each frequency band at the SRW reference interface.

3) The sensitivity is expressed in dBV/V, the SRW junction loudness rating $(JLR)_{SRWsnd}$ shall be calculated according to equation A-23d of [ITU-T P.79], bands 4-17, m = 0.175 and the weighting factors W_j for JLR according to Table A.2 of [ITU-T P.79].

12.2.3 Test SRW junction loudness rating in receiving direction

1) The test signal used for the measurements shall be CSS according to [ITU-T P.501]. The test signal is –16 dBm0, measured at the electrical reference point and averaged over the complete test signal sequence.

2) The test arrangement is according to clause 12.1. For the calculation, the averaged level at the SRW reference interface is used. The SRW receiving sensitivity is determined by the bands 4-17 according to Table A.2 of [ITU-T P.79].

   For the calculation, the average signal level of each frequency band is referred to the signal level of the reference signal measured in each frequency band.

3) The sensitivity is expressed in dBV/V, the SRW junction loudness rating in receiving JLR_{SRWrcv} shall be calculated according to equation A-23d of [ITU-T P.79], bands 4-17, m = 0.175 and the weighting factors W_j for JLR according to Table A.2 of [ITU-T P.79].

12.2.4 SRW linearity in sending direction

12.2.4.1 Requirements

The test is aimed to detect any amplitude non-linearities, including AGC or companding. For acoustical signal level variation in the range of –40 dB/+5 dB from the nominal signal level S_{SRWsnd}, the measured JLR_{SRWsnd} shall not deviate more than ±0.5 dB from the JLR_{SRWsnd} measured with the nominal signal level S_{SRWsnd}.

12.2.4.2 Test

1) The test signal used for the measurements shall be CSS according to [ITU-T P.501]. The test signals are in the range of –40 dBV to +5 dBV in steps of 5 dB relative to the nominal signal level S_{SRWsnd}, measured at the SRWR. The test signal level is the average level of the complete test signal.

   The measured power density spectrum at the SRWR is used as the reference power density spectrum for determining the sending sensitivity.

2) The test arrangement is according to clause 12.1. The SRW sending sensitivity is calculated from each band of the 14 frequencies given in Table A.2 of [ITU-T P.79], bands 4-17.

   For the calculation, the average measured level at the electrical reference point for each frequency band is referred to the average test signal level measured in each frequency band at the SRW reference interface.

3) The sensitivity is expressed in dBV/V, the SRW junction loudness rating $(JLR)_{SRWsnd}$ shall be calculated according to equation A-23d of [ITU-T P.79], bands 4-17, m = 0.175 and the weighting factors in the sending direction according to Table A.2 of [ITU-T P.79].

12.2.5 SRW linearity in receiving direction

12.2.5.1 Requirements

The test is aimed to detect any amplitude non-linearities including AGC or companding. For network signal level variations in the range of –40 dB to +5 dB relative to the nominal signal level, the measured JLR_{SRWrcv} shall not deviate more than ±0.5 dB from the JLR_{SRWrcv} measured with nominal signal level.

12.2.5.2 Test

1) The test signal used for the measurements shall be CSS according to [ITU-T P.501]. The test signals are in the range of –40 dBV to +5 dBV in steps of 5 dB relative to the nominal signal level measured at the electrical reference point. The test signal level is the average level of the complete test signal.

   The measured power density spectrum at the electrical reference point is used as the reference power density spectrum for determining the receiving sensitivity.

2) The test arrangement is according to clause 12.1. The SRW receiving sensitivity is calculated from each band of the 14 frequencies given in Table A.2 of [ITU-T P.79], bands 4-17.

   For the calculation, the average measured level at the SRWR for each frequency band is referred to the average test signal level measured in each frequency band at the electrical reference interface.

3) The sensitivity is expressed in dBV/V, the SRW junction loudness rating $(JLR)_{SRWrcv}$ shall be calculated according to equation A-23d of [ITU-T P.79], bands 4-17, m = 0.175 and the weighting factors in the receiving direction according to Table A.2 of [ITU-T P.79].

12.3 SRW sensitivity frequency responses

12.3.1 SRW sending sensitivity frequency response

12.3.1.1 Requirements

The sending sensitivity frequency response is measured from the SRWR to the POI (reference speech codec of the system simulator, output).

The tolerance mask for the sending sensitivity frequency response is shown in Table 12-1, the mask is drawn by straight lines between the breaking points in Table 12-1 on a logarithmic (frequency) – linear (dB sensitivity) scale.

Table 12-1 – Tolerance mask for the SRW sending sensitivity frequency response

Frequency [Hz]	Upper limit	Lower limit
200	0	–2
3 100	0	–2
3 400	0	–3
NOTE – All sensitivity values are expressed in dB on an arbitrary scale.		

12.3.1.2 Test

1) The test signal used for the measurements shall be artificial voice according to [ITU-T P.50]. The test signal is S_{SRWsnd}, the level is averaged over the complete test signal.

   The measured power density spectrum at the SRW reference interface is used as the reference power density spectrum for determining the SRW sending sensitivity.

2) The test arrangement is according to clause 12.1. The SRW sending sensitivity frequency response is determined in one-third octave bands as given by the R.40-series of preferred numbers in [ISO 3] for frequencies from 100 Hz to 4 kHz, inclusive. In each one-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.

3) The sensitivity is determined in dBV/V.

12.3.2 SRW receiving sensitivity frequency response

12.3.2.1 Requirements

The receiving sensitivity frequency response is measured from the electrical reference point (input of the system simulators, POI) to the SRW reference interface.

The tolerance mask for the receiving sensitivity frequency response is shown in Table 12-2, the mask is drawn by straight lines between the breaking points in Table 12-2 on a logarithmic (frequency) – linear (dB sensitivity) scale.

Table 12-2 – Tolerance mask for the receiving sensitivity frequency response

Frequency [Hz]	Upper limit	Lower limit
200	0	–2
3 100	0	–2
3 400	0	–3
NOTE – All sensitivity values are expressed in dB on an arbitrary scale.		

12.3.2.2 Test

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

2) The test arrangement is according to clause 12.1. The SRW receiving sensitivity frequency response is determined in one-third octaves as given by the R.40-series of the preferred numbers in [ISO 3] for frequencies from 100 Hz to 4 kHz inclusive. In each one-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.

3) The sensitivity is determined in dBV/V.

12.4 SRW noise cancellation test in sending

12.4.1 Requirements

The objective of this test is to check whether no noise cancellation is active in the mobile phone. No acoustic echo control shall be active in the mobile phone. The mobile phone noise cancellation is measured from the SRWR to the POI (reference speech codec of the system simulator, output).

The attenuation of the simulated background noise test signal shall not deviate more than ±1 dB for all periods of the test signal.

12.4.2 Test

1) The test arrangement is according to clause 12.1.

2) The test signal used for the measurements shall be pink noise with a duration of 5 s followed by a pause of 3 s, both repeated three times. The test signal during the active parts of the signal is S_{SRWsnd}, the level is averaged over the active parts of the test signal.

3) 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 versus time. The levels are calculated from the time domain using an integration time of 5 ms.

4) The attenuation versus time is determined for each pink noise section.

   NOTE – If a non-linear or time variant behaviour of the phone is observed, the tests as described in clause 11.12.4 can be applied to determine the behaviour of the phone in more detail. Instead of inserting the test signals acoustically, they have to be inserted electrically.

12.5 SRW speech quality during single talk

12.5.1 SRW one-way speech quality in sending

12.5.1.1 Requirement

The nominal values for the speech quality measured from/to the electrical reference point (POI) shall be

$$\text{MOS-LQON} \geq 4.0$$

12.5.1.2 Test

The test uses [ITU-T P.862].

1) The test signals used are the English test sequences as specified in [ITU-T P.501] (2 male speakers, 2 female speakers, two sentences each). The test signal level is S_{SRWsnd}, measured at the SRWR; the test signal level is measured as "active speech level" according to [ITU-T P.56]. The speech activity should be between 30% and 70%.

   The original speech signal is used as the reference signal for the determination of speech quality.

2) The test arrangement is according to clause 12.1. MOS-LQON is determined.

The calculation is made using the signal recorded at the electrical interface.

3) The one-way speech quality is determined as MOS-LQON.

12.5.2 SRW one-way speech quality in receiving

12.5.2.1 Requirement

The nominal values for the speech quality measured from/to the electrical reference point (POI) shall be

$$\text{MOS-LQON} \geq 4.0$$

12.5.2.2 Test

The test uses [ITU-T P.862].

1) The test signals used are the English test sequences as specified in [ITU-T P.501] (2 male speakers, 2 female speakers, two sentences each). The test signal is –16 dBm0; the test signal level is measured as "active speech level" according to [ITU-T P.56]. The speech activity should be between 30% and 70%.

The original speech signal is used as the reference signal for the determination of speech quality.

2) The test arrangement is according to clause 12.1. MOS-LQON is determined. The signal measured at the SRWR is used for the calculation.

3) The one-way speech quality is determined as MOS-LQON.

12.5.3 Verification of disabled echo control

12.5.3.1 Requirements

No acoustic echo control shall be active on the mobile phone. An artificial echo path consisting of an attenuation of 20 dB/40 dB and a delay of 20 ms is inserted at the short-range wireless transmission reference interface. The difference between the echo loss measured with 20 dB echo loss and 40 dB echo loss shall be 20 dB ±0.2 dB.

12.5.3.2 Test

1) For the test, an artificial echo path is inserted at the SRWR. The test set up is shown in Figure 12-3.

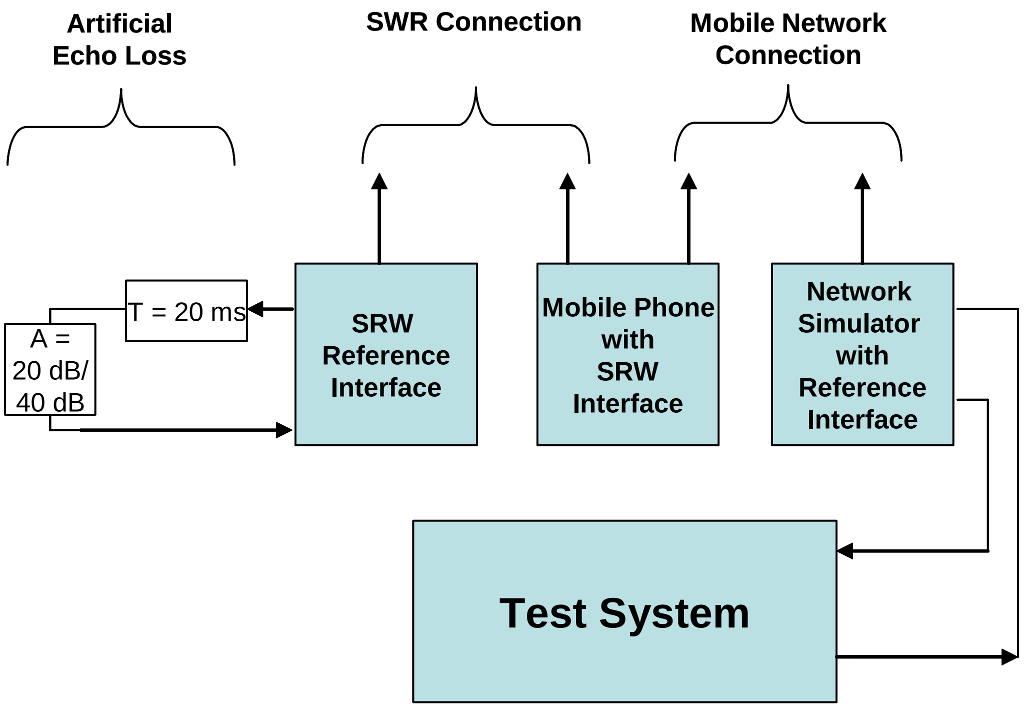


Figure 12-3 – Test setup with artificial echo loss

2) The attenuation between the input of the electrical reference point and the output of the electrical reference point is measured using a speech-like test signal.

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

4) The test signal is a pn sequence according to [ITU-T P.501] with a length of 4096 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 –180º and +180º.

5) TCL_W is calculated according to Annex B, clause B.4 of [ITU-T G.122], (trapezoidal rule). 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) The difference between the echo loss measured with 20 dB echo loss and 40 dB echo loss is determined.

13 Car-to-car communication

13.1 Guidance on subjective testing

Beside objective testing of hands-free telephones, a subjective performance evaluation is also necessary.

The tests described here – in addition to the tests as described in [b-ITU-T P.8xx] – are targeted mainly to "in situ" hands-free tests for optimizing hands-free systems in the target car and under conditions which are not covered by the objective test specification. These tests are neither mandatory nor intended to replace the objective tests described in this Recommendation.

For conducting the tests, the hands-free system under test has to be installed in the target car, which will be referenced in this Recommendation as near-end. Either a landline phone (car-to-land test) or an observing car equipped also with the hands-free system under test (car-to-car test) may serve as the far-end. It is recommended to not only test the hands-free system in a landline connection but also in a car-to-car connection because the latter case can be regarded as a worst-case scenario resulting in lower hands-free quality compared to landline connections.

The evaluation of hands-free performance should be done for different background noise scenarios, such as different driving speeds, fan/defrost settings, etc.

For the main part of the subjective tests, native language should be used. For the recordings, additional languages can be selected.

Since conversational tests are rather time consuming, most hands-free tests are conducted as single talk and double talk tests following a clear given structure.

The evaluation is done at the far-end and/or the near-end, depending on the type of the test category.

To conduct the tests as effectively as possible, it is advantageous to use a tool providing test persons at both ends of the telephone connection with a detailed test procedure and the possibility to easily perform a rating.

The performance evaluation of the hands-free system covers categories such as:

– echo cancellation (echo intensity, speed of convergence, etc.);
– double talk performance (echo during double talk, speech level variation, etc.);
– speech and background noise quality in the sending direction (level, level variation, speech distortion, etc.);
– speech quality in receiving direction (level, level variation, speech distortion);
– stability of the echo canceller for a "closed loop" connection when doing car-to-car hands-free communication.

For the evaluation, some ITU-T Recommendations can serve as guidelines, such as [ITU-T P.800], [ITU-T P.800.1] and others from [b-ITU-T P.8xx]. The judgement is done according to rating scales given for each test case. The offered rating scales are of MOS type, having grades 1 to 5 to be chosen from, where 5 denotes "best" and 1 denotes "worst" quality. Some of the rating scales are designed for a more diagnostic purpose (e.g., "echo duration").

The evaluation has to be done by experts who are experienced with subjective testing of hands-free systems. However, some of the tests described here could be conducted with naive subjects when following the procedures described in the relevant Recommendations of [b-ITU-T P.8xx].

During the tests, the signals on the near-end and the far-end may be recorded to be used later on for third-party listening evaluation.

Tables 13-1 and 13-2 show possible test scenarios and rating categories.

Table 13-1 – Overview of car-to-landline tests

Echo canceller		**Rating**
Single talk	Four typical driving scenarios: (Germany: E.g., 0 km/h, 100 km/h, 130 km/h, 160 km/h) Fan/defrost (off or "worst case") Receiving volume: nominal, maximum, varying Enclosure dislocation	Disturbance caused by echo Echo characteristics: – intensity – duration – frequency of occurrence Background noise variation
Double talk	Four typical driving scenarios: (Germany: E.g., 0 km/h, 100 km/h, 130 km/h, 160 km/h) Receiving volume: nominal, maximum, varying Enclosure dislocation	Disturbance caused by echo Echo characteristics: – intensity – duration – frequency of occurrence Speech level variation Speech intelligibility/listening effort
Conversation	0 km/h	Disturbance caused by echo Echo characteristics: – intensity – duration – frequency of occurrence Speech level variation Speech intelligibility/listening effort
Speech and background noise quality (sending direction)		**Rating**
Stationary noise	4 typical driving scenarios: (Germany: E.g., 0 km/h, 100 km/h, 130 km/h, 160 km/h) Driver's window closed/opened Several fan settings	Speech level Speech level fluctuation Speech sound quality Intelligibility/listening effort Speech naturalness Background noise quality Signal-to-noise ratio
Transient noise	Fan: switching on/off/change setting Indicator noise, wiper noise Passing vehicles	Background noise quality Adaptation to background noise
Speech quality (receiving direction)		**Rating**
Single talk far-end	Maximum receiving volume New call	Speech level Speech sound quality Speech intelligibility/listening effort

Table 13-2 – Overview of car-to-car tests

Echo canceller		**Rating**
Single talk	4 typical driving scenarios (Germany: E.g., 0 km/h, 100 km/h, 130 km/h, 160 km/h) Fan (off or "worst case") Receiving volume: nominal, maximum, varying Enclosure dislocation	Disturbance caused by echo Echo characteristics: – intensity – duration – frequency of occurrence Background noise variation
Double talk	4 typical driving scenarios (Germany: E.g., 0 km/h, 100 km/h, 130 km/h, 160 km/h) Receiving volume: nominal, maximum, varying Enclosure dislocation	Disturbance caused by echo Echo characteristics: – intensity – duration – frequency of occurrence Speech level variation Speech intelligibility/listening effort
System stability (car-to-car)		**Rating**
Stability	Receiving volume: nominal, maximum Test signal: speech or impulse-like excitation	Stability Echo characteristics: – intensity – duration – frequency of occurrence
Speech and background noise quality (sending direction)		**Rating**
Stationary noise	4 typical driving scenarios (Germany: E.g., 0 km/h, 100 km/h, 130 km/h, 160 km/h)	Speech level Speech level fluctuation Speech sound quality Speech intelligibility/listening effort background noise quality Signal-to-noise ratio
NOTE – The driving condition selected should reflect the type of car and the typical range of driving speeds in the country where the car hands-free system is intended to be installed.		

13.2 Test environment and equipment

Figures 13-1 and 13-2 outline the test environment for the scenarios car-to-landline and car-to-car. In both test scenarios, the supervisor is located at the far-end, i.e., landline or the observing car, respectively. The supervisor guides the test procedure and performs most of the test evaluation.

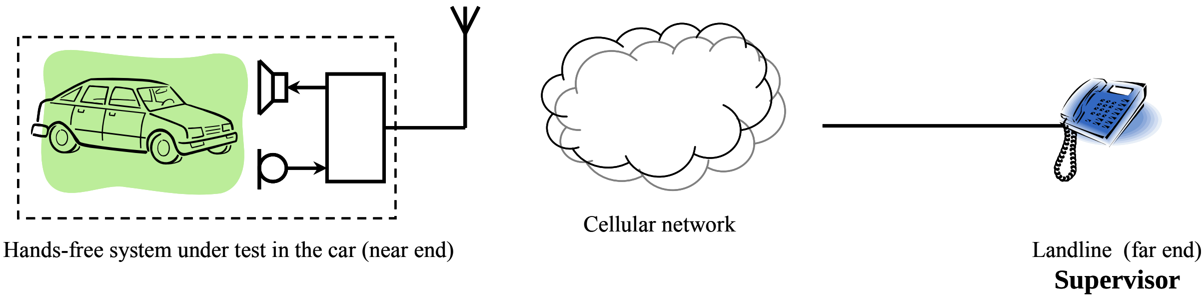


Figure 13-1 – Hands-free test car-to-landline

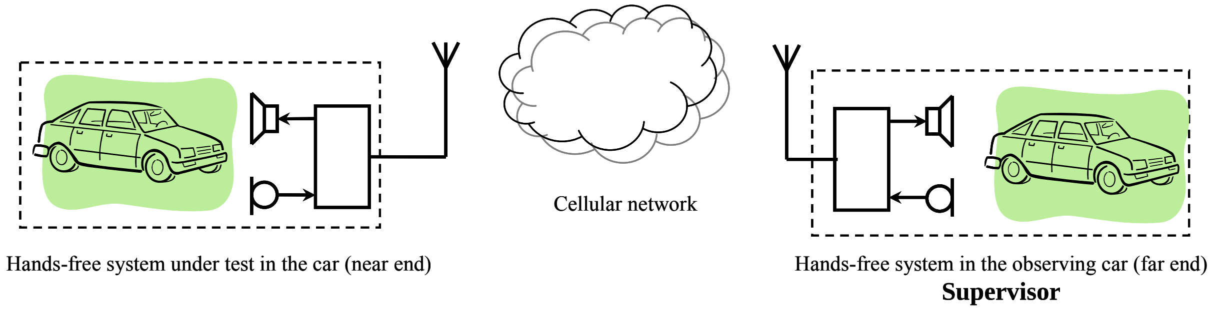


Figure 13-2 – Hands-free test car-to-car

Required test equipment:

– Car equipped with the hands-free telephone system under test.
– For car-to-car tests, both vehicles should be identical in terms of make, model, model year, and hands-free system.
– Appropriate recording equipment is needed at the landline station and the car for documentation purposes and if third-party listening tests are required later on.

Further requirements:

– The tests have to be conducted by expert listeners.
– Male and female test persons should be in the car, quiet and loud talkers should be considered, different positions of the talkers with respect to the microphone should be considered.
– The hands-free tests have to be conducted in areas with good cellular coverage.
– In case of a short-range wireless transmission connection between phone and hands-free system, the phone should be placed in such a way that the short-range wireless transmission coverage and cellular coverage are good.
– Some cellular networks have an influence on the hands-free performance (e.g., echo, AGC, noise reduction, etc.). These networks should not be chosen for the tests, if possible.
– Both the far-end and the near-end participants should be familiar with each other's voice.
– For the duration of the tests, there should be no change in the driver, front-passenger or landline speaker.

NOTE – The network QoS parameters can be monitored by specific systems during the test, e.g., using TEMS investigation.

A description of the testing environment should be provided and should contain the following information:

– Car make, model, model year, tyres, type of road, engine type, interior trim (cloth or leather, sunroof, etc.).
– Mobile telephone type and software version.
– Telematics microphone – description, location, orientation, distance to driver.
– Network provider.
– Hardware and software version of the hands-free device under test.
– Hardware and software version of all devices which are part of the sound system.

Background noise and driving situations

For the tests, several different driving and background noise scenarios should be considered, which correspond to the main operating conditions of the hands-free system. Additionally, some worst case scenarios might be regarded.

For the driving scenarios, this implies that the scenarios might differ between different countries depending, e.g., on the national speed limit for cars.

For example, the following scenarios for different driving noise levels could be used (Germany):

– "low": 0 km/h
– "medium": 100 km/h
– "high": 160 km/h

For countries with speed limits, the maximum allowed speed might be used with, additionally, the climate control switched on to an appropriate level.

Additionally to the driving scenarios, some different settings of the fan/defrost/climate control/recirculation are also useful background noise cases. As a "worst case" scenario, the airstream from the fan might flow directly over the hands-free microphone in the car.

Test documentation

– After each test, a rating of the performance is done by referring to the given rating scales.
– At the beginning of every test, one of the test participants announces the test number (this is done for recording purposes).
– Audio signals should be recorded for documentation purposes:
  • On landline: Telephone audio recording of uplink and downlink signal.
  • In the observing car: Binaural recording.
  • In the car under test: Binaural recording.

Notes on performance rating

For the ratings, the following items should be considered:

– The listeners who do the ratings have to be experienced with hands-free telephone systems.
– Limitations of the network have to be taken into account, e.g., for rating the hands-free system when the network's voice codec quality depends on the telephone connection traffic (e.g., CDMA).
– The offered rating scales are of MOS type, having grades 1 to 5 to be chosen from, where 5 denotes "best" and 1 denotes "worst" quality.
– Some of the rating scales are designed for a more diagnostic purpose (e.g., the "echo duration").

An example for a possible questionnaire applicable for this type of test is given in Annex C.

Annex A

Speech quality measurements

(This annex forms an integral part of this Recommendation)

In this annex, a testing method for determining speech quality in the sending and receiving directions is described. The speech quality expressed in TMOS, and its relation to existing ITU-T terminology according to [ITU-T P.800.1], is as follows.

MOS related to listening-only situations

These MOS scores are applicable to a listening-only situation. Three different cases have to be distinguished.

MOS-LQS

The score has been collected in a laboratory test by calculating the arithmetic mean value of subjective judgments on a 5-point ACR quality scale, as it is defined in [ITU-T P.800].

Subjective tests carried out according to [b-ITU-T P.830], [b-ITU-T P.835] and [b-ITU-T P.840] give results in terms of MOS-LQS.

MOS-LQO

The score is calculated by means of an objective model which aims at predicting the quality for a listening-only test situation. Objective measurements made using the model given in [ITU-T P.862.1] and [b-ITU-T P.862.2] give results in terms of MOS-LQO.

It should be noted that the method recommended by [ITU-T P.862.1] and [b-ITU-T P.862.2] is validated between electrical interfaces only. Currently, no ITU Recommendation exists which covers the measurement of listening quality including acoustical interfaces. Work on a new ITU-T Recommendation, which is intended to include acoustical interfaces, is in progress.

MOS-LQO (acoustical)

This kind of measurement is performed at acoustical interfaces. In order to predict the listening quality as perceived by the user, this measurement includes the actual telephone set products provided by the manufacturer or vendor. In combination with the choice of the acoustical receiver in the lab test ("artificial ear"), there will be more or less degradation of the signal between the handset's receiver and the artificial ear. The same constraints apply for hands-free telephony. Consequently, for more realistic test scenarios, there may be a degradation of the measured MOS value, while for more artificial test scenarios, there may be a negligible difference. The TMOS is a MOS-LQO (acoustical) prediction.

	Listening-only	**Conversational**	**Talking**
Subjective	MOS-LQSy	MOS-CQSy	MOS-TQSy
Objective	MOS-LQOy	MOS-CQOy	MOS-TQOy
Estimated	MOS-LQEy	MOS-CQEy	MOS-TQEy
NOTE – The letter "y" at the end of the above acronyms is a placeholder for the description of the respective audio bandwidth, see the following provisional instructions.			

– N for MOS scores obtained for narrow-band (300-3400 Hz) speech relative to a narrow-band high quality reference. This is applicable, for instance, to narrow-band only subjective tests or to ITU-T P.862.1 scores.

– W for MOS scores obtained for wideband (50-7000 Hz) speech relative to a wideband high quality reference. This is applicable, for instance, to wideband-only subjective tests or to ITU-T P.862.2 scores.
– M for MOS scores obtained for narrow-band or wideband speech relative to a wideband high quality reference in a mixed-bandwidths context. This is applicable, for instance, to mixed-bandwidths subjective tests.

Further information can be found in [ITU-T P.800.1].

Annex B

Principles of relative approach

(This annex forms an integral part of this Recommendation)

The relative approach [b-ETSI EG 202 396-1] is an analysis method developed to model a major characteristic of human hearing. This characteristic is the much stronger subjective response to distinct patterns (tones and/or relatively rapid time-varying structure) than to slowly changing levels and loudness.

The idea behind the relative approach analysis is based on the assumption that human hearing creates a running reference sound (an "anchor signal") for its automatic recognition process against which it classifies tonal or temporal pattern information moment-by-moment. It evaluates the difference between the instantaneous pattern in both time and frequency and the "smooth" or less-structured content in similar time and frequency ranges. In evaluating the acoustic quality of a complex "patterned" signal, the absolute level or loudness is almost without any significance. Temporal structures and spectral patterns are important factors in deciding whether a sound is judged as annoying or disturbing.

Similar to human hearing, and in contrast to other analysis methods, the relative approach algorithm does not require any reference signal for the calculation. Only the signal under test is analysed. Comparable to the human experience and expectation, the algorithm generates an "internal reference" which can be best described as a forward estimation. The relative approach algorithm objectifies pattern(s) in accordance with human perception by resolving or extracting them while largely rejecting pseudo-stationary energy. At the same time, it considers the context of the relative difference of the "patterned" and "non-patterned" magnitudes.

Figure B.1 shows a block diagram of the relative approach. The time-dependent spectral pre-processing can either be done by a filter bank analysis (1/nth octave, typically 1/12th octave) or a hearing model spectrum versus time according to the hearing model of Sottek (see [b-Sottek]). Both of them result in a spectral representation versus time. Both calculate the spectrograph using only linear operation and their outputs are therefore directly comparable. The hearing model analysis parameters are fixed and based on the processing in human ears, whereas the input parameters for the filter bank analysis can vary. The filter bank pre-processing approximates the hearing model version.

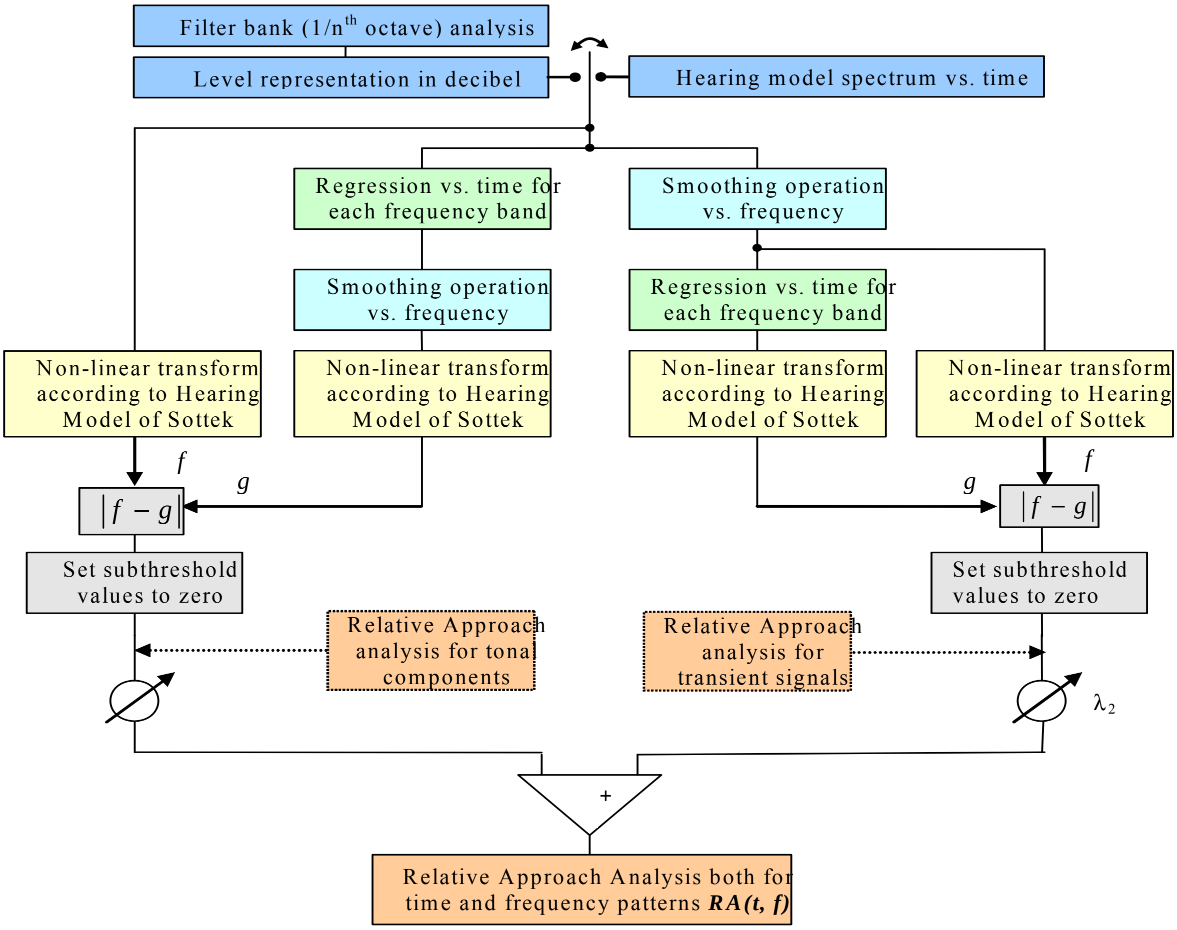


Figure B.1 – Block diagram of relative approach

Two different variants of the relative approach can be applied to the pre-processed signal. The first one applies a regression versus time for each frequency band in order to cover human expectation for each band within the next short period of time. Afterwards, for each time slot, a smoothing versus frequency is performed. The next step is a non-linear transformation according to the hearing model of Sottek (for more information see [b-Sottek]). This output is compared to the source signal which is also hearing model transformed. Non-relevant components for human hearing are finally set to zero. This approach focuses on the detection of tonal components. The second version first smoothes versus frequency within a time slot and then applies the regression versus time. This output signal is transformed non-linearly to the hearing model. It is compared to the output of the smoothing versus frequency which is also non-linearly transformed according to the hearing model. Finally non-relevant components for human hearing are again set to zero. Thus, more transient structures are detected.

Via the factors λ_1 and λ_2, the weighting of the relative approach for tonal and transient signals can be set. Typically $\lambda_1 = 0$ and $\lambda_2 = 1$ are chosen. Thus, the model is tuned to detect time-variant transient structures.

The result of the relative approach analysis is a three-dimensional spectrograph displaying the deviation from the "close to the human expectation" between the estimated and the current signal, and displayed versus time and frequency. The relative approach uses a time resolution of Δt = 6.66 ms. The frequency range from 15 Hz to 24 kHz is divided into 128 frequency bands Δf_m which corresponds to a 1/12th octave resolution. Due to the non-linearity in the relationship between sound pressure and perceived loudness, the term "compressed pressure" in compressed Pascal, (cPa) is used to describe the result of applying the non-linear transform. The relative approach can determine how "close to the human expectation" a signal is, but not if this expectation is of a high or a low quality origin.

Annex C

Example of a questionnaire for subjective testing

(This annex forms an integral part of this Recommendation)

C.1 Performance rating – Overview

The performance evaluation of the hands-free system covers categories such as:

– echo cancellation performance during single talk and double talk (echo during single talk, convergence after enclosure dislocation, echo during double talk, speech level variation during double talk, etc.);

– speech and background noise quality in the sending direction (level, level variation, speech naturalness, etc.);

– speech quality in the receiving direction (level, level variation, speech naturalness);

– stability of the echo canceller for "closed loop" operation when performing car-to-car hands-free communication.

For the evaluation, some ITU-T Recommendations served as guidelines, such as [ITU-T P.800], [ITU-T P.800.1] and [b-ITU-T P.8xx]. The judgement is done according to rating scales given for each test case. The offered rating scales are of MOS type having grades 1 to 5 to be chosen from, where 5 denotes "best" and 1 denotes "worst" quality. Some of the rating scales are designed for a more diagnostic purpose (e.g., "echo duration").

Additional notes on performance rating

For the ratings, the following items should be considered:

– The evaluation has to be done by experts who are experienced with subjective testing of hands-free systems.

– Limitations of the network have to be taken into account and should be documented. For example, for rating the hands-free system when using CDMA, the rating could say "not better than 2 because of CDMA network".

– The offered rating scales are of MOS type having grades 1 to 5 to be chosen from, where 5 denotes "best" and 1 denotes "worst" quality.

– Some of the rating scales are designed for a more diagnostic purpose (e.g., the "echo duration").

C.2 Test categories and rating types

Table C.1 gives an overview of the test categories and the related rating scales.

Table C.1 – Test categories and rating types overview

Test category	Test sub-category	Conversation type	Rating side	Rating type	Rating scales	Test condition/ variation
Speech and background noise quality in sending direction	Speech level	Single talk near-end	Far-end	Speech level	Loudness preference (office)	Stationary background noise scenario: low/medium/high, fan/defrost, window
	Speech quality	Single talk near-end	Far-end	Speech quality	Speech level fluctuation; speech sound quality; speech naturalness; intelligibility/listening effort	Stationary background noise scenario: low/medium/high, fan/defrost, window
	Background noise quality during near-end single talk	Single talk near-end	Far-end	Background noise quality	Signal-to-noise ratio noise quality	Stationary background noise scenario: low/medium/high, fan/defrost
	Transient background noise quality	Idle and single talk near-end	Far-end	Transient background noise quality	Transient noise quality	Transient background noise scenario: fan/defrost start-up, wiper, indicator
		Idle	Far-end	Adaptation to background noise	Adaptation to background noise	Transient background noise scenario: noise jump, e.g., fan/defrost start-up
Speech quality in receiving direction	Speech level	Single talk far-end	Near-end	Speech level (car, max. vol.)	Loudness preference (car) for maximum volume	Only high background noise scenario
		Single talk far-end	Near-end	Speech level (car, nominal vol., new call)	Loudness preference (car) for nominal volume, new call	Low background noise scenario
	Speech quality	Single talk far-end	Near-end	Speech quality	Speech sound quality intelligibility/listening effort	Low background noise scenario

Table C.1 – Test categories and rating types overview

Test category	Test sub-category	Conversation type	Rating side	Rating type	Rating scales	Test condition/ variation
Echo cancellation	Echo during single talk	Single talk far-end	Far-end	Echo	Disturbance caused by Echo Echo characteristics (only to be rated if echo occurs): • intensity, • duration, • frequency of occurrence • echo intelligibility	Stationary background noise scenario: low/medium/high, fan/defrost; volume (car), movement of driver (enclosure dislocation)
	Background noise quality during far-end single talk	Single talk far-end	Far-end	Background noise quality	Comfort noise quality (EC test)	Stationary background noise scenario: medium/high, fan/defrost
	Echo during double talk	Double talk	Far-end	Echo	Disturbance caused by Echo, Echo characteristics (only to be rated if echo occurs): • intensity, • duration, • frequency of occurrence • echo intelligibility	Stationary background noise scenario: Low/medium/high, fan/defrost; volume (car), movement of driver (enclosure dislocation)
	Speech quality at near-end during double talk	Double talk	Near-end	Speech quality	Speech level variation during double talk Intelligibility/listening effort during double talk	Low background noise scenario
	Speech quality at far-end during double talk	Double talk	Far-end	Speech quality	Speech level variation during double talk Intelligibility/listening effort during double talk	Stationary background noise scenario: low/medium/high, fan/defrost, window
System stability	System stability	Special test	Far-end	Echo convergence and stability	System Stability: speed of convergence of echo cancellation and robustness against echo back coupling	Car-to-car, low background noise scenario, EC not adapted at start of test, maximum volume

In addition to the tests given in Table C.1, some conversational tests can be performed as described in [ITU-T P.800] and [b-ITU-T P.831] (conversation opinion scale).

C.3 Speech and background noise quality in sending direction

Speech level (office)

Description:

The rating scale is applied in test cases which evaluate the preferred speech level of the received signal at the far-end (office).

Test category:	Speech and background noise quality in sending direction
Conversation type:	Single talk at near-end
Rating side:	Rated at far-end (office)
Test conditions:	Stationary background noise scenarios of low/medium/high level Additional background noise scenarios: different fan/defrost settings, window open, etc.
Scale type:	Loudness-preference scale, [ITU-T P.800], ACR

The rating scale is given in Table C.2.

Table C.2 – Loudness preference (office)

Rating description		Grade
Much louder than preferred		1
Louder than preferred		3
Preferred		5
Quieter than preferred		3
Much quieter than preferred		1

Speech level fluctuations

Description:

The rating scale is applied in test cases which evaluate level variations in speech in single talk situations.

Test category:	Speech and background noise quality in sending direction
Conversation type:	Single talk at near-end
Rating side:	Rated at far-end
Test conditions:	Stationary background noise scenarios of low/medium/high level Additional background noise scenarios: different fan/defrost settings, window open, etc.
Scale type:	Special scale for HF system diagnostic evaluation

Level fluctuations to be examined for this evaluation are characterized by:

– level fading;
– short drop-outs, e.g., due to missing data packets;
– cut-offs (missing word-ends or syllables);
– chopped voice.

The rating scale is given in Table C.3.

Table C.3 – Speech level fluctuations

Rating description		Grade
No speech level variation audible		5
Slight level variations, just audible or very rarely occurring		4
Moderate speech level variations, may occur frequently		3
Sometimes words or syllables are attenuated or missing		2
Many drop-outs, cut-offs, missing words or syllables, heavily chopped voice		1

Speech quality/speech naturalness

Description:

The rating scale is applied in test cases which evaluate speech naturalness received at the far-end for different background noise scenarios at the near-end. This evaluation includes possible impairments caused by signal distortion and band-limited effects which also degrade the speech naturalness. The best quality case for this evaluation would be a hand-set comparable voice quality.

Test category:	Speech and background noise quality in sending direction
Conversation type:	Single talk at near-end
Rating side:	Rated at far-end
Test conditions:	Stationary background noise scenarios of low/medium/high level Additional background noise scenarios: different fan/defrost settings, window open, etc.
Scale type:	Special scale for HF system diagnostic evaluation (considering also the degradation scale, [ITU-T P.800], degradation category rating)

Properties to be examined for this evaluation are:

– synthetic/robotic sound;

– speech signal distortion characterized by a scratchy sound;

– band limitation or filtering effects characterized by:
  - a shrill, sharp, thin, tinny or muffled sounding speech;
  - an emphasis on high frequencies or low frequencies.

The rating scale is given in Table C.4.

Table C.4 – Speech quality/speech naturalness

Rating description		Grade
Speech sound is comparable to hand-set voice quality; Speech sounds clear and transparent	Natural	5
Minor degradation compared to hand-set, still natural voice Possibly slight distortions and/or slight band limitation effects	.	4
Maybe slight synthetic voice at times and/or low level distortion and/or moderate band limitation effects	.	3
Very noticeable synthetic voice and/or heavy distortion and/or higher degree of band-limitation	.	2
Signal barely recognizable as voice	Unnatural	1

Intelligibility/listening effort

Description:

The rating table is applied to evaluate the effort required to understand the meaning of words and sentences. The applicable test cases are single talk and different background noise scenarios at the near-end and evaluation on the far-end.

Test category:	Speech and background noise quality in sending direction
Conversation type:	Single talk at near-end
Rating side:	Rated at far-end
Test conditions:	Stationary background noise scenarios of low/medium/high level Additional background noise scenarios: different fan/defrost settings, window open, etc.
Scale type:	Listening effort scale, [ITU-T P.800], ACR

The rating scale is given in Table C.5.

The question heading this scale could be, for example:

How would you judge the effort required to understand words and sentences of your remote partner?

Table C.5 – Intelligibility/listening-effort

Rating description		Grade
Every word was clearly understood with no effort required		5
Speech of the other side was understood with no appreciable effort required		4
Some words were hard to understand, moderate effort was required		3
Many words were hard to understand, considerable effort was required		2
No meaning understood with any feasible effort		1

Signal-to-noise ratio for near-end single talk

Description:

The following rating scale is intended to evaluate the noise level compared to the speech level.

The evaluation is usually done in a high environmental noise test condition. The signal-to-noise ratio directly depends on the background noise level of the test scenario. The judgement, therefore, will be of a worse grade with the test scenario changing to higher environmental noise. Therefore, it is problematic for this test category to be used as an absolute evaluation; it is more applicable for comparison of different systems for the same environmental noise scenario.

Test category:	Speech and background noise quality in sending direction
Conversation type:	Single talk at near-end
Rating side:	Rated at far-end
Test conditions:	Stationary background noise scenarios of low/medium/high level Additional background noise scenarios: different fan/defrost settings
Scale type:	Scale is adapted from "degradation category scale" (Annex D of [ITU-T P.800]) and the extended "detectability scale" of the "quantal-response detectability test" (Annex C of [ITU-T P.800])

The rating scale is given in Table C.6.

Table C.6 – Signal-to-noise ratio

Rating description		Grade
Noise very low, just audible		5
Noise audible, noise level clearly lower than speech level, noise is not disturbing		4
Medium noise level, lower than speech level, noise slightly disturbing		3
High noise level, almost same level as speech, clearly disturbing, but call would be continued		2
Noise louder than speech, intolerably disturbing, call would be abandoned		1

Background noise quality

Description:

This scale is for evaluation of the sound quality of the near-end background noise examined at the far-end. For explicit evaluation of transient noise sources, another rating scale is defined later in this annex.

Test category:	Speech and background noise quality in sending direction
Conversation type:	Single talk at near-end
Rating side:	Rated at far-end
Test conditions:	Stationary background noise scenarios of low/medium/high level Additional background noise scenarios: different fan/defrost settings
Scale type:	Special scale for HF system diagnostic evaluation

The following sound characteristics have to be taken into account:

– Are there changes of noise sound and level over time?

– Are there artefacts audible (clicks, pops, rattle) which cannot be matched to a natural source (e.g., road bumps) or do not sound like their natural source?

– Naturalness:
  - does the noise sound like part of the natural background?
  - does the noise sound synthetic (musical tones, watery sound)?
  - does the noise sound distorted?

Artefacts, synthetic sound, level and sound variation over time, and noise not sounding like part of the natural background, result in lower (worse) grades.

The rating scale is given in Table C.7.

Table C.7 – Background noise quality

Rating description		Grade
Comfortable, natural sound, constant in sound and level, no artefacts	Natural	5
Slight distortion/synthetic sound, almost no artefacts, almost constant in sound and level	.	4
Moderate distortion/synthetic sound, or some artefacts/clicks/plops audible, or some moderate variation in sound and level	.	3
Clearly unnatural/distorted/synthetic sounding, or many artefacts/clicks/plops, or often variation in sound and level	.	2
Completely unnatural/distorted/synthetic sound, or permanent artefacts/clicks/plops, or permanent variations in sound and level, very uncomfortable to listen to	Unnatural	1

Noise quality for transient noise sources

Description:

This scale is for evaluation of the sound quality of transient near-end background noise examined at the far-end. Transient noise sources can be, for example, the car's activated wiper, indicator, etc.

Test category:	Speech and background noise quality in sending direction
Conversation type:	Idle and single talk at near-end
Rating side:	Rated at far-end
Test conditions:	Transient noise sources active (indicator, wiper, etc.)
Scale type:	Special scale for HF system diagnostic evaluation

For the evaluation, the naturalness of the noise sound has to be taken into account, for example:

– Does the noise sound like part of the natural background?

– Does the noise sound synthetic?

– Does the noise sound distorted?

The rating scale is given in Table C.8.

Table C.8 – Transient noise quality

Rating description		Grade
Comfortable/natural sound	Natural	5
Almost natural sounding Slight distortion/synthetic sound	.	4
Moderately unnatural sounding Moderately distortion/synthetic sound	.	3
Clearly unnatural/distorted/synthetic sounding	.	2
Completely unnatural/distorted/synthetic sound	Unnatural	1

Adaptation to background noise

Description:

This scale is for evaluation of the speed of adaptation of the noise suppression to the near-end background noise after a noise level jump. The evaluation is done at the far-end.

Test category:	Speech and background noise quality in sending direction
Conversation type:	Idle
Rating side:	Rated at far-end
Test conditions:	Transient background noise: noise jump, e.g., fan/defrost start-up
Scale type:	Special scale for HF system diagnostic evaluation

The test can be conducted, for example, by turning on the defrost/fan in the car to a high setting. When doing the test like this, the time the defrost/fan needs to run up has to be taken into account.

The rating scale is given in Table C.9.

Table C.9 – Adaptation to background noise

Rating description		Grade
Immediate adaptation	Very fast	5
Adaptation time ≤ 1 second		4
Adaptation time 2 … 3 seconds		3
Adaptation time 3 … 10 seconds		2
Adaptation time > 10 seconds	Very slow	1

C.4 Speech quality in receiving direction (in the car under test)

Speech sound quality/speech naturalness (receiving)

Description:

The rating scale is applied in test cases which evaluate speech naturalness received in the car under test (near-end). This evaluation includes possible impairments caused by signal distortion and band-limited effects which also degrade the speech naturalness. The best quality case for this evaluation would be a hand-set comparable voice quality. The evaluation is done in low background noise conditions using a nominal volume setting in the car.

Test category:	Speech quality in receiving direction
Conversation type:	Single talk at far-end
Rating side:	Rated at near-end
Test conditions:	Low background noise condition
Scale type:	Special scale for HF system diagnostic evaluation (considering also the degradation scale, [ITU-T P.800], degradation category rating)

Properties to be examined for this evaluation are:

– synthetic/robotic sound;
– speech signal distortion characterized by a scratchy sound;
– band limitation or filtering effects characterized by:
    • a shrill, sharp, thin, tinny or muffled sounding speech;
    • an emphasis on high frequencies or low frequencies.

For the rating scale, see Table C.4.

Intelligibility/listening effort (receiving)

Description:

The rating table is applied to evaluate the effort required to understand the meaning of words and sentences. The applicable test cases are single talk at the far-end examined in the car under test (near-end) in a low background noise condition using a nominal volume setting in the car.

Test category:	Speech quality in receiving direction
Conversation type:	Single talk at far-end
Rating side:	Rated at near-end
Test conditions:	Low background noise condition
Scale type:	Listening effort scale, [ITU-T P.800], ACR

The question heading this scale could be, for example:

How would you judge the effort required to understand words and sentences of your remote partner?

For the rating scale see Table C.5.

Speech level (receiving, maximum volume)

Description:

The rating table is applied to evaluate the speech level heard from the loudspeakers in the car (near-end) when being in a high background noise condition and having the telephone volume set to maximum.

Test category:	Speech quality in receiving direction
Conversation type:	Single talk at far-end
Rating side:	Rated at near-end
Test conditions:	High background noise condition
Scale type:	Loudness-preference scale, [ITU-T P.800], ACR

The rating scale is given in Table C.10.

Table C.10 – Loudness preference (Car) for maximum volume setting

Rating description		Grade
Much louder than preferred		1
Louder than preferred		3
Preferred (for maximum volume setting)		5
Quieter than preferred		3
Much quieter than preferred		1

Speech level for nominal volume and new call

Description:

The rating table is applied to evaluate the speech level heard from the loudspeakers in the car (near-end) during a new call after having set the volume to nominal in the prior call.

Test category:	Speech quality in receiving direction
Conversation type:	Single talk at far-end
Rating side:	Rated at near-end
Test conditions:	High background noise condition
Scale type:	Loudness-preference scale, [ITU-T P.800], ACR

The rating scale is given in Table C.11.

Table C.11 – Loudness preference for maximum volume and new call

Rating description		Grade
Much louder than preferred		1
Louder than preferred		3
Preferred		5
Quieter than preferred		3
Much quieter than preferred		1

C.5 Echo cancellation performance

A first step in the evaluation of echo is to judge the level of disturbance that it causes. Only if echo is perceived, is an additional evaluation of other echo characteristics performed. The evaluation of these echo characteristics are intended for diagnostic purposes.

The tests are intended to rate the perceived quality according to:

– Amount and nature of echo during single talk.

– Amount and nature of echo during double talk.

– Convergence characteristics of the EC to handle variation of the echo path (e.g., when the driver is moving inside the car).

– Speech quality during double talk situations (e.g., intelligibility and speech level variation); this judgement is done on the near-end and the far-end.

– Stability of the EC in car-to-car communication.

The tests are performed for different background noise scenarios and driving conditions in order to get some information about the EC robustness in high background noise conditions.

The scales given in this clause can be applied for steady state conditions and initial convergence tests.

Disturbance caused by echo

Description:

The rating scale is applied to evaluate the disturbance caused by echo examined at the far-end. The test scenarios should include different background noise conditions at the near-end.

Test category:	Echo cancellation
Conversation type:	Single talk at far-end Double talk
Rating side:	Rated at far-end
Test conditions:	Stationary background noise scenarios of low/medium/high level Additional background noise scenarios: different fan/defrost settings Enclosure dislocation due to movement of the driver Volume setting in the car: nominal, maximum
Scale type:	Conversation impairment scale, [ITU-T P.800] and [b-ITU-T P.831]

The other echo rating scales (echo intensity, duration, frequency of occurrence and intelligibility) below are intended for diagnostic purposes. They are only used if echo is perceived.

The rating scale is given in Table C.12.

For the evaluation, the participant has to answer a question such as:

How would you judge the degradation/impairment/disturbance due to echo of your own voice during the test?

Table C.12 – Disturbance caused by echo

Rating description		Grade
Imperceptible		5
Perceptible but not annoying		4
Slightly annoying		3
Annoying		2
Very annoying		1

Echo intensity

Only for diagnostic purposes.

Description:

This rating is intended for diagnostic purposes and only has to be done if echo is perceived.

The scale is applied for evaluation of the echo level occurring in far-end single talk and in double talk test cases. The evaluation is done at the far-end. The test scenarios should include different background noise conditions at the near-end.

Test category:	Echo cancellation
Conversation type:	Single talk at far-end Double talk
Rating side:	Rated at far-end
Test conditions:	Stationary background noise scenarios of low/medium/high level Additional background noise scenarios: different fan/defrost settings Enclosure dislocation due to movement of the driver Volume setting in the car: nominal, maximum
Scale type:	Special scale for HF system diagnostic evaluation

The rating scale is given in Table C.13.

Table C.13 – Echo intensity

Rating description		Grade
–		–
Slight		4
Moderate		3
Loud		2
Very loud		1

Echo duration

Only for diagnostic purposes.

Description:

This rating is intended for diagnostic purposes and only has to be done if echo is perceived.

The scale is applied for evaluation of the echo duration occurring in far-end single talk and in double talk test cases. The evaluation is done at the far-end. The test scenarios should include different background noise conditions at the near-end.

Test category:	Echo cancellation
Conversation type:	Single talk at far-end Double talk
Rating side:	Rated at far-end
Test conditions:	Stationary background noise scenarios of low/medium/high level Additional background noise scenarios: different fan/defrost settings Enclosure dislocation due to movement of the driver Volume setting in the car: nominal, maximum
Scale type:	Special scale for HF system diagnostic evaluation

The rating scale is given in Table C.14.

Table C.14 – Echo duration

Rating description		Grade
Very short		
Short		
Moderate		
Long		
Very long/permanent		

Frequency of echo occurrence

Only for diagnostic purposes.

Description:

This rating is intended for diagnostic purposes and only has to be done if echo is perceived.

The scale characterizes the number of echo events occurring during the echo test, the test cases include far-end single talk and double talk. The evaluation is done at the far-end. The test scenarios should include different background noise conditions at the near-end.

Test category:	Echo cancellation
Conversation type:	Single talk at far-end Double talk
Rating side:	Rated at far-end
Test conditions:	Stationary background noise scenarios of low/medium/high level Additional background noise scenarios: different fan/defrost settings Enclosure dislocation due to movement of the driver Volume setting in the car: nominal, maximum
Scale type:	Special scale for HF system diagnostic evaluation

The rating scale is given in Table C.15.

Table C.15 – Frequency of echo occurrence

Rating description		Grade
Only once during the test		
Only twice during the test		
Infrequently several times		
Echo occurs more often than not		
Permanent		

Echo intelligibility

Only for diagnostic purposes.

Description:

This rating is intended for diagnostic purposes and only has to be done if echo is perceived.

The scale is applied for characterizing the type of sound of the echo occurring in far-end single talk and in double talk test cases (pure artefacts or the echoed voice of the talker). The evaluation is done on the far-end. The test scenarios should include different background noise conditions at the near-end.

Test category:	Echo cancellation
Conversation type:	Single talk at far-end Double talk
Rating side:	Rated at far-end
Test conditions:	Stationary background noise scenarios of low/medium/high level Additional background noise scenarios: different fan/defrost settings Enclosure dislocation due to movement of the driver Volume setting in the car: nominal, maximum
Scale type:	Special scale for HF system diagnostic evaluation

The rating scale is given in Table C.16.

Table C.16 – Echo intelligibility

Rating description		Grade
Pure artefacts		
Hardly recognizable as voice		
Distorted voice		
Slightly distorted voice		
Clear voice		

Comfort noise quality (EC test)

Description:

The scale is for evaluation of the near-end background noise sound quality received at the far-end during far-end single talk. The evaluation gives information about the quality of comfort noise injection. Transient noise should be avoided at the near-end during the test.

Test category:	Echo cancellation
Conversation type:	Single talk at far-end
Rating side:	Rated at far-end
Test conditions:	Stationary background noise scenarios of low/medium/high level Additional background noise scenarios: different fan/defrost settings
Scale type:	Special scale for HF system diagnostic evaluation

The following sound characteristics have to be taken into account:

– Are there changes of noise sound and level over time (e.g., when changing from natural near-end background noise to comfort noise injection and vice versa)?

– Are artefacts audible (clicks, pops, rattle)?

– Naturalness:
  - Does the noise sound like part of the natural background?
  - Does the noise sound synthetic (musical tones, watery sound)?
  - Does the noise sound distorted?

Artefacts, synthetic sound, level and sound variation over time, and noise not sounding like part of the natural background will result in lower (worse) grades.

The rating scale is given in Table C.17.

Table C.17 – Comfort noise quality (EC test)

Rating description		Grade
No difference between comfort noise and natural background noise perceivable Comfortable, natural, constant in sound and level, no artefacts	Natural	5
Slight difference between comfort noise and natural background noise perceivable Slightly distorted/synthetic sound, almost no artefacts, almost constant in sound and level		4

Table C.17 – Comfort noise quality (EC test)

Rating description		Grade
Moderate difference between comfort noise and natural background noise perceivable Moderate distorted/synthetic sound, or some artefacts/clicks/plops audible, or some moderate variation in sound and level		3
Clear difference between comfort noise and natural background noise perceivable Clearly unnatural/distorted/synthetic sounding, or many artefacts/clicks/plops, or often variation in sound and level		2
Comfort noise does not sound like the natural background noise at all Very unnatural/distorted/synthetic sound, or permanent artefacts/clicks/plops, or permanent variations in sound and level, very uncomfortable to listen to	Unnatural	1

Speech level variation during double talk

Description:

The scale is applicable to evaluate speech level variations during double-talk. The evaluation is done at both ends, near-end and far-end. When the evaluation is done on the far-end then the test scenarios should include different background noise conditions at the near-end.

Test category:	Echo cancellation
Conversation type:	Double talk
Rating side:	Rated at near-end Rated at far-end
Test conditions:	For rating at the near-end: – low background noise scenario For rating at the far-end: – stationary background noise scenarios of low/medium/high level – additional background noise scenarios: different fan/defrost settings – enclosure dislocation due to movement of the driver – volume setting in the car: nominal, maximum
Scale type:	Special scale for HF system diagnostic evaluation

Level variations are characterized by:

– switching of an attenuation exactly during the double talk phases;

– level fading;

– short drop-outs or cut-offs (missing word-ends or syllables);

– chopped voice.

The rating scale is given in Table C.18.

Table C.18 – Speech level variation during double talk

Rating description		Grade
No speech level variation audible		5
Slight level variations, just audible or very rarely occurring		4
Moderate speech level variations may occur frequently, sometimes words or syllables might be attenuated or missing, or moderate constant attenuation being switched during the double talk phases		3
Many drop-outs, cut-offs, missing words or syllables, heavily chopped voice, or high constant attenuation being switched during the double talk phases		2
Not possible to hear the other end at all during double talk		1

Intelligibility/listening effort during double talk

Description:

The rating scale is applied to evaluate the effort required to understand the meaning of words and sentences during double talk. The evaluation is done at both ends, near-end and far-end. When the evaluation is done on the far-end then the test scenarios should include different background noise conditions at the near-end.

Test category:	Echo cancellation
Conversation type:	Double talk
Rating side:	Rated at near-end Rated at far-end
Test conditions:	For rating at the near-end: – low background noise scenario For rating at the far-end: – stationary background noise scenarios of low/medium/high level – additional background noise scenarios: different fan/defrost settings – enclosure dislocation due to movement of the driver – volume setting in the car: nominal, maximum
Scale type:	Listening effort scale, [ITU-T P.800], ACR

The rating scale is given in Table C.19.

Table C.19 – Intelligibility/listening effort during double talk

Rating description		Grade
Every word was clearly understood during double talk with no effort required		5
Speech of the other side was understood during double talk with no appreciable effort required		4
Some words were hard to understand during double talk, moderate listening effort was required		3
Many words were hard to understand during double talk, considerable listening effort was required		2
No meaning understood with any feasible effort during double talk		1

C.6 Hands-free system stability tests (car-to-car)

The evaluation for system stability is intended to examine the convergence characteristic of the echo cancellation for "closed loop" operation when performing car-to-car hands-free communication. For the according tests, the hands-free system under test is installed in both cars, and neither system has the echo cancellation filter adapted when starting the test.

System stability

Description:

The scale is applied for evaluation of the convergence of the echo cancellation and the robustness against back-coupling of echo in car-to-car communication. In one test case, single talk at the far-end is performed; in another test case, an impulse-like noise signal is generated close to the microphone at the far-end.

In both cars, the EC filter is not adapted at the start of the test. The evaluation is done on the far-end.

As a suggestion, an appropriate test procedure could be as follows. Both cars are at standstill and have the doors open and the volume set to nominal. Then single talk is performed in both cars, one after the other, to give the EC filters the chance to adapt to this situation (or not to be adapted when the doors are closed afterwards). After that, in both cars, the volume is set to maximum. For generating an impulse-like noise, for example, the doors of the cars could be slammed. Another possibility would be to close the doors quietly and to generate the impulse like noise by clapping the hands close to the microphone.

Test category:	System stability
Conversation type:	Single talk at far-end Impulse-like noise at far-end
Rating side:	Rated at far-end
Test conditions:	Initial state of EC filters: not adapted Volume setting in the car: maximum
Scale type:	Special scale for HF system diagnostic evaluation

The rating scale is given in Table C.20.

Table C.20 – System stability

Rating description		Grade
No echo is audible		5
Some echo can be heard, but disappears very quickly		4
The echo disappears slowly, the recurrences are audible for a few seconds		3
The echo disappears very slowly, the recurrences are audible for more than 10 seconds		2
The echo builds up like in an unstable feedback system		1

Appendix I

A method to determine the listening speech quality

(This appendix does not form an integral part of this Recommendation)

One possible method to determine the listening speech quality is described in this appendix:

I.1 One-way speech quality in sending

1) The test signals used are the German test sequences as specified in [ITU-T P.501] (two male speakers, two female speakers, two sentences each). The test signal is equalized at the MRP, the test signal level is –4.7 dBPa at the MRP, the test signal level is measured as "active speech level" according to [ITU-T P.56]. The speech activity should be between 30% and 70%. Finally, the level at the HATS-HFRP is adjusted to –28.7 dBPa.

   The original speech signal is used as the reference signal for the determination of the speech quality.

2) The test arrangement is according to clause 7.1. TMOS is determined using the settings "high quality handset" and "narrow-band" with TOSQA2001.

   The calculation is made using the signal recorded at the electrical interface.

3) The one-way speech quality is determined as TMOS.

I.2 One-way speech quality in receiving

1) The test signals used are the German test sequences as specified in [ITU-T P.501] (two male speakers, two female speakers, two sentences each). The test signal is –16 dBm0, the test signal level is measured as "active speech level" according to [ITU-T P.56]. The speech activity should be between 30% and 70%.

   The original speech signal is used as the reference signal for the determination of the speech quality.

2) The test arrangement is according to clause 7.1. For the measurement, the artificial head is free-field equalized according to [ITU-T P.581]. The equalized output signal of the inboard ear is used for the measurements. TMOS is determined using the settings "high quality handset" and "narrowband" with TOSQA2001.

   The calculation is made using the signal recorded at the DRP of the inboard ear of the artificial head.

3) The one-way speech quality is determined as TMOS.

   NOTE 1 – TOSQA2001 has only been validated with German language test material.

   NOTE 2 – This test method does not apply to systems including artificial bandwidth extension.

Bibliography

[b-ITU-T G.722.2] Recommendation ITU-T G.722.2 (in force), *Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB).* <http://web.itu.int/rec/T-REC-G.722.2>

[b-ITU-T O.41] Recommendation ITU-T O.41 (in force), *Psophometer for use on telephone-type circuits*. <http://web.itu.int/rec/T-REC-O.41>

[b-ITU-T P.8xx] Recommendation ITU-T P.8xx-series (in force), *Acoustics – Determination of occupational noise exposure and estimation of noise-induced hearing impairment*. <http://web.itu.int/rec/T-REC-P>

[b-ITU-T P.830] Recommendation ITU-T P.830 (in force), *Subjective performance assessment of telephone-band and wideband digital codecs*. <http://web.itu.int/rec/T-REC-P.830>

[b-ITU-T P.831] Recommendation ITU-T P.831 (in force), *Subjective performance evaluation of network echo cancellers*. <http://web.itu.int/rec/T-REC-P.831>

[b-ITU-T P.835] Recommendation ITU-T P.835 (in force), *Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm*. <http://web.itu.int/rec/T-REC-P.835>

[b-ITU-T P.840] Recommendation ITU-T P.840 (in force), *Subjective listening test method for evaluating circuit multiplication equipment*. <http://web.itu.int/rec/T-REC-P.840>

[b-ITU-T P.862.2] Recommendation ITU-T P.862.2 (in force), *Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs*. <http://web.itu.int/rec/T-REC-P.862.2>

[b-ISO 1999] ISO 1999:in force, *Acoustics – Determination of occupational noise exposure and estimation of noise-induced hearing impairment*. <http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=6759>

[b-ETSI EG 202 396-1] ETSI EG 202 396-1 (2008), *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*. <http://pda.etsi.org/pda/queryform.asp>

[b-3GPP TS 46.010] 3GPP TS 46.010 (in force), *Full-rate speech; Transcoding*. <http://www.3gpp.org/ftp/Specs/html-info/46010.htm>

[b-TIA-127-A] TIA-127-A (in force), *Enhanced Variable Rate Codec Speech Service Option 3 for Wideband Spread Spectrum Digital Systems*. <http://tia.nufu.eu/std/TIA-127-A>

[b-TIA-127-B] TIA-127-B (in force), *Software Distribution for Enhanced Variable Rate Codec, Speech Service Options 3, 68, and 70 for Wideband Spread Spectrum Digital Systems*. <http://tia.nufu.eu/std/TIA-127-C>

[b-TIA-127-C] TIA-127-C (in force), *Software Distribution for Enhanced Variable Rate Codec, Speech Service Options 3, 68, and 70 for Wideband Spread Spectrum Digital Systems*. <http://tia.nufu.eu/std/TIA-127-C>

[b-TIA-893] TIA-893 (in force), *Selectable Mode Vocoder (SMV) – Service Option for Wideband Spread Spectrum Communications Systems*. <http://tia.nufu.eu/std/TIA-893-1[E]>

[b-TIA-1016] TIA-1016 (in force), *Source-Controlled Variable-Rate Multimode Wideband Speech Codec (VMR-WB) Service Option 62 for Spread Spectrum Systems*. <http://tia.nufu.eu/std/TIA-1016>

[b-TIA/EIA/IS-95-A] TIA/EIA/IS-95-A (in force), *Mobile Station – Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System*. <http://tia.nufu.eu/std/TIA|EIA|IS-95-A >

[b-TIA/EIA-95-B] TIA/EIA-95-B (in force), *Mobile Station-Base Station Compatibility Standard for Wideband Spread Spectrum Cellular Systems*. <http://tia.nufu.eu/std/TIA|EIA-95-B>

[b-TIA/EIA-136-x] TIA/EIA-136-series (in force), *TDMA Cellular Packet Data*. <http://tia.nufu.eu/std/TIA|EIA|IS-136+Series>

[b-TIA/EIA-136-410] TIA/EIA-136-410 (in force), *TDMA Cellular PCS – Radio Interface Enhanced Full-Rate Voice Codec*. <http://tia.nufu.eu/std/TIA|EIA-136-410>

[b-Berger] Berger, J., *Results of objective speech quality assessment including receiving terminals using the advanced TOSQA2001*, ITU-T Contribution, Dec. 2000, COM 12-20-E. <http://www.itu.int/md/T01-SG12-C-0020/>

[b-Fingscheidt] Fingscheidt, T., and Suhadi, S. (2007), *Quality Assessment of Speech Enhancement Systems by Separation of Enhanced Speech, Noise, and Echo*, INTERSPEECH 2007, Antwerpen, Belgium, Aug. <http://www.interspeech2007.org/Technical/detail.php?ses=WeB.P3>

[b-Sottek] Sottek, R., Genuit, K. (2005), *Models of Signal Processing in Human Hearing*, AEU - International Journal of Electronics and Communications, Vol. 59, No. 3, pp. 157-165. <http://dx.doi.org/10.1016/j.aeue.2005.03.016>

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	**Telephone transmission quality, telephone installations, local line networks**
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects and next-generation networks
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2009