

International Telecommunication Union

**ITU-T**

TELECOMMUNICATION  
STANDARDIZATION SECTOR  
OF ITU

**P.1110**

(03/2017)

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

Communications involving vehicles

---

**Wideband hands-free communication in motor  
vehicles**

Recommendation ITU-T P.1110

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
Voice terminal characteristics	Series	P.30
		P.300
Reference systems	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 speech quality	Series	P.80
Methods for objective and subjective assessment of speech and video quality	Series	P.800
Audiovisual quality in multimedia services	Series	P.900
Transmission performance and QoS aspects of IP end-points	Series	P.1000
<b>Communications involving vehicles</b>	<b>Series</b>	<b>P.1100</b>
Models and tools for quality assessment of streamed media	Series	P.1200
Telemeeting assessment	Series	P.1300
Statistical analysis, evaluation and reporting guidelines of quality measurements	Series	P.1400
Methods for objective and subjective assessment of quality of services other than speech and video	Series	P.1500

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

# Recommendation ITU-T P.1110

## Wideband hands-free communication in motor vehicles

### Summary

Recommendation ITU-T P.1110 describes performance requirements and test methods for wideband hands-free communication in motor vehicles. This Recommendation covers:

- built-in hands-free systems,
- aftermarket hands-free car kits,
- corded headsets, and
- wireless headsets

to be used in motor vehicles for communication.

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

For testing purposes, 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.501, ITU-T P.502, ITU-T P.340, and ITU-T P.380.

### History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T P.1110	2009-12-14	12	<a href="http://handle.itu.int/11.1002/1000/10659">11.1002/1000/10659</a>
2.0	ITU-T P.1110	2015-01-13	12	<a href="http://handle.itu.int/11.1002/1000/12416">11.1002/1000/12416</a>
3.0	ITU-T P.1110	2017-03-01	12	<a href="http://handle.itu.int/11.1002/1000/13175">11.1002/1000/13175</a>

### Keywords

Hands-free, headset, motor vehicles, quality of service, QoS, wideband.

---

\* To access the Recommendation, type the URL <http://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID. For example, <http://handle.itu.int/11.1002/1000/1830-en>.

## 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/>.

© ITU 2017

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

## Table of Contents

	<b>Page</b>
1	Scope..... 1
2	References..... 1
3	Definitions ..... 2
3.1	Terms defined elsewhere ..... 2
3.2	Terms defined in this Recommendation ..... 2
4	Abbreviations and acronyms ..... 4
5	Conventions ..... 6
6	How to use this Recommendation ..... 6
6.1	Determining compliance with this Recommendation ..... 6
6.2	Testing at different stages of the development cycle ..... 7
7	Test arrangement ..... 7
7.1	Test arrangement in a car ..... 9
8	Digital interfaces for development, debugging and test ..... 13
8.1	Interfaces and access points..... 13
8.2	Test set-up and tests..... 15
9	Test signals and test signal levels ..... 16
9.1	Signals ..... 16
9.2	Background noise signals ..... 17
10	Measurement parameters and requirements for microphones used in speakerphone hands-free systems..... 19
10.1	Microphone measurements in anechoic conditions..... 19
10.2	Microphone measurements in the car..... 22
11	Measurement parameters and requirements for hands-free terminals..... 27
11.1	Preparation measurements..... 27
11.2	Delay..... 27
11.3	Loudness ratings ..... 29
11.4	Sensitivity frequency responses ..... 32
11.5	Speech quality during single talk ..... 36
11.6	Listening speech quality stability ..... 37
11.7	Idle channel noise ..... 38
11.8	Out-of-band signals ..... 40
11.9	Distortion in send ..... 42
11.10	Distortion in receive ..... 42
11.11	Echo performance..... 43
11.12	Double talk performance ..... 52
11.13	Background noise transmission..... 59
12	Verification of the transmission performance of short-range wireless transmission enabled phones..... 63

	<b>Page</b>
12.1	Interface definition and calibration ..... 63
12.2	SRW loudness ratings..... 67
12.3	SRW sensitivity frequency responses ..... 69
12.4	SRW noise cancellation test in send direction ..... 70
12.5	SRW speech quality during single talk ..... 71
13	Car-to-car communication ..... 74
14	Guidance on subjective testing ..... 74
15	Corded headsets ..... 79
15.1	Connector type..... 79
15.2	Connector wiring and electrical specifications..... 79
15.3	Headset receive characteristics..... 79
15.4	Headset transmit characteristics ..... 79
15.5	Standard behaviour in the presence of corded headsets ..... 79
16	Wireless headsets..... 79
16.1	Wireless headset types..... 79
16.2	Test methodology for verification of standard behaviour ..... 79
16.3	Standard behaviour in the presence of wireless headsets ..... 79
Annex A	– Speech quality measurements..... 81
Annex B	– Principles of Relative Approach..... 82
Annex C	– Example for a questionnaire for subjective testing..... 84
C	Performance rating..... 84
C.1	Overview ..... 84
C.2	Test categories and rating types ..... 84
C.3	Speech and background noise quality in send direction..... 86
C.4	Speech quality in receive direction (in the car under test) ..... 92
C.5	Echo cancellation performance ..... 94
C.6	Hands-free system stability tests (car-to-car) ..... 101
Annex D	– Standard set of user scenarios..... 102
Annex E	– System stability with insufficient far-end echo loss ..... 104
Annex F	– Microphone related sound field reproduction technique – Multi-point noise simulation method (MPNS)..... 106
F.1	Loudspeaker setup in the car ..... 106
F.2	Microphone setup in the car ..... 106
F.3	Background noise recordings and reference noise ..... 107
F.4	Equalization..... 107
F.5	Accuracy of the equalization ..... 108
Appendix I	– A method to determine the listening speech quality ..... 109
I.1	One-way speech quality in sending ..... 109
I.2	One-way speech quality in receiving ..... 109
Bibliography	..... 110

# Recommendation ITU-T P.1110

## Wideband hands-free communication in motor vehicles

### 1 Scope

The aim of this Recommendation is the definition of user scenarios, test methods and requirements for wideband hands-free communication in motor vehicles. This Recommendation covers:

- built-in hands-free systems,
- aftermarket hands-free car kits,
- corded headsets, and
- wireless headsets

to be used in motor vehicles for communication.

Furthermore, the compatibility between narrowband and wideband implementations is addressed.

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

For testing purposes, 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 [ITU-T P.501], [ITU-T P.502], [ITU-T P.340], [ITU-T P.380], [b-ETSI ES 202 739], and [b-ETSI ES 202 740].

Although important subjective testing is outside the scope of this Recommendation, guidance on how to conduct subjective tests can be found in clause 14.

### 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.122] Recommendation ITU-T G.122 (1993), *Influence of national systems on stability and talker echo in international connections.*
- [ITU-T O.41] Recommendation ITU-T O.41 (1994), *Psophometer for use on telephone-type circuits.*
- [ITU-T P.56] Recommendation ITU-T P.56 (1993), *Objective measurement of active speech level.*
- [ITU-T P.57] Recommendation ITU-T P.57 (2009), *Artificial ears.*
- [ITU-T P.58] Recommendation ITU-T P.58 (1996), *Head and torso simulator for telephonometry.*
- [ITU-T P.79] Recommendation ITU-T P.79 (2007), *Calculation of loudness ratings for telephone sets.*

- [ITU-T P.340] Recommendation ITU-T P.340 (2000), *Transmission characteristics and speech quality parameters of hands-free terminals.*
- [ITU-T P.380] Recommendation ITU-T P.380 (2003), *Electro-acoustic measurements on headsets.*
- [ITU-T P.501] Recommendation ITU-T P.501 (2009), *Test signals for use in telephony.*
- [ITU-T P.502] Recommendation ITU-T P.502 (2000), *Objective test methods for speech communication systems using complex test signals.*
- [ITU-T P.581] Recommendation ITU-T P.581 (2000), *Use of head and torso simulator (HATS) for hands-free terminal testing.*
- [ITU-T P.800] Recommendation ITU-T P.800 (1996), *Methods for subjective determination of transmission quality.*
- [ITU-T P.800.1] Recommendation ITU-T P.800.1 (2006), *Mean Opinion Score (MOS) terminology.*
- [ITU-T P.830] Recommendation ITU-T P.830 (1996), *Subjective performance assessment of telephone-band and wideband digital codecs.*
- [ITU-T P.831] Recommendation ITU-T P.831 (1998), *Subjective performance evaluation of network echo cancellers.*
- [ITU-T P.835] Recommendation ITU-T P.835 (2003), *Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm.*
- [ITU-T P.840] Recommendation ITU-T P.840 (2003), *Subjective listening test method for evaluating circuit multiplication equipment.*
- [ITU-T P.863] Recommendation ITU-T P.863 (2011), *Perceptual objective listening quality assessment.*
- [ITU-T P.863.1] Recommendation ITU-T P.863.1 (2013), *Application guide for Recommendation ITU-T P.863.*
- [IEC 61260] IEC 61260 (1995), *Electroacoustics – Octave-band and fractional-octave-band filters.*
- [ISO 1999] ISO 1999:1990, *Acoustics – Determination of occupational noise exposure and estimation of noise-induced hearing impairment.*
- [ISO 3745] ISO 3745:2003, *Acoustics – Determination of sound power levels of noise sources using sound pressure – Precision methods for anechoic and hemi-anechoic rooms.*

### **3 Definitions**

#### **3.1 Terms defined elsewhere**

None.

#### **3.2 Terms defined in this Recommendation**

This Recommendation defines the following terms:

**3.2.1 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.



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

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

**3.2.4 diffuse-field equalization:** Equalization of the head and torso simulator (HATS) sound pick-up, equalization of the difference, in dB, between the spectrum level of the acoustic pressure at the ear-drum reference point (DRP) and the spectrum level of the acoustic pressure at the HATS reference point (HRP) in a diffuse sound field with the HATS absent using the reverse nominal curve given in Table 3 of [ITU-T P.58].

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

**3.2.6 free-field equalization:** The transfer characteristic 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.

**3.2.7 free-field reference point:** A point located in the free sound field, at least at 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).

**3.2.8 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 [b-ITU-T P.51].

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

**3.2.10 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.

**3.2.11 headset hands-free terminal:** A device that includes telephone receiver and transmitter, which is typically secured to the head or ear of the wearer.

**3.2.12 inboard ear:** Ear closest to the centreline of the vehicle.

**3.2.13 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.

**3.2.14 MOS-LQO (mean opinion score – listening-only quality objective):** The score 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.863] give results in terms of MOS-LQO. For further information about MOS terminology, see Annex A.

**3.2.15 MOS-TQO (mean opinion score – talking quality objective):** The score 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 not yet standardized.

**3.2.16 motor vehicle:** Any vehicle equipped with a motor where a hands-free system can be installed.

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

**3.2.18 nominal setting of the volume control:** When a receive volume control is provided, the setting which is closest to the nominal RLR of 2 dB.

**3.2.19 receive 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 e.m.f. to measured sound pressure.)

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

**3.2.21 short-range wireless transmission (SRW):** Radio transmission link used to connect the hands-free system typically to a mobile phone, which is connected to the mobile network.

**3.2.22 speakerphone hands-free terminal:** A device that is not attached to the user and does not require the use of hands for holding a telephone receiver and transmitter during communication.

**3.2.23 wideband speech:** Voice service with enhanced quality compared to PCM 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 <sub>H,R</sub>	Attenuation Range in Receive direction
A <sub>H,R,dt</sub>	Attenuation Range in Receive direction during Double Talk
A <sub>H,S</sub>	Attenuation Range in Send direction
A <sub>H,S,dt</sub>	Attenuation Range in Send direction during Double Talk
BGN	Background Noise
CDMA	Code Division Multiple Access
CSS	Composite Source Signal
D/A	Digital/Analogue
DI	Digital Interface
DRP	ear-Drum Reference Point
DTX	Discontinuous Transmission
DUT	Device Under Test
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
HF System	Hands-Free System
HFT	Hands-Free Terminal

HVAC	Heating, Ventilation and Air Conditioning
JLR	Junction Loudness Rating
$L_{R,min}$	minimum activation Level (Receive direction)
$L_{S,min}$	minimum activation Level (Send direction)
MOS	Mean Opinion Score
MOS-LQOn	Mean Opinion Score-Listening-only Quality, Narrowband
MPNS	Multi-Point Noise Simulation Method
MRP	Mouth Reference Point
NC	Noise Criterion
NR	Noise Reduction
OHC	Overhead Console
PCM	Pulse Code Modulation
POI	Point of Interconnection
QoS	Quality of Service
RF	Radio Frequency
RLR	Receive Loudness Rating
SLR	Send Loudness Rating
SNR	Signal-to-Noise Ratio (also S/N)
SRW	Short-Range Wireless transmission
SRWR	SRW Reference Point
S/N	Signal-to-Noise ratio (also SNR)
TCL	Terminal Coupling Loss
TCLw	weighted Terminal Coupling Loss
TEMS	Terrestrial Ecosystem Monitoring Site
TMOS	TOSQA Mean Opinion Score
TOSQA	Telecommunication Objective Speech Quality Assessment
$T_{as}$	Access-Specific delay
$T_{ass}$	Access-Specific delay in Send
$T_{asr}$	Access-Specific delay in Receive
$T_{imps}$	Implementation dependent send delay HFT
$T_{impr}$	Implementation dependent receive delay HFT
$T_r$	Receive delay hands-free terminal
$T_{r,R}$	built-up Time (Receive direction)
$T_{r,S}$	built-up Time (Send direction)
$T_{rtd-HF}$	Round Trip Delay hands-free terminal
$T_{rtdimp}$	Implementation dependent round trip delay HFT
$T_{SRWr}$	Mobile phone receive delay from the cellular antenna to the SRW antenna

$T_{SRWs}$	Mobile phone send delay from the SRW antenna to the cellular antenna
$T_{SRWrd}$	Round trip delay SRW consisting of the sum of $T_{SRWr}$ and $T_{SRWs}$
$T_s$	Send delay hands-free terminal

## 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 dB point)

**dBm0p**: weighted dBm0, according to [ITU-T O.41]

**dBm0(C)**: C weighted dBm0, according to [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

**dB SPL**: sound pressure level relative to 20  $\mu$ Pa, expressed in dB; (94 dB SPL=0 dB Pa)

**dBV(P)**: P-weighted voltage relative to 1 V, expressed in dB, according to [ITU-T O.41]

**dB**: relative power level of a signal in a transmission path referred to the level at a reference point on the path (0 dB point)

**N**: Newton

**V<sub>rms</sub>**: 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

### 6.1 Determining compliance with this Recommendation

To claim compliance with this Recommendation, the following must be true:

- The hands-free terminal (HFT) passes all of the requirements of clause 11 using the test procedures specified,
- The HFT passes all noise related requirements in clause 11 for each of the user scenarios defined in Table D.1,
- If the HFT is intended to be used in multiple vehicles (e.g., aftermarket hands-free car kits), then the HFT must meet the above criteria on a minimum of three vehicles that are representative (e.g., microphone type/placement, noise) of all vehicles before compliance is claimed.

NOTE – 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.

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

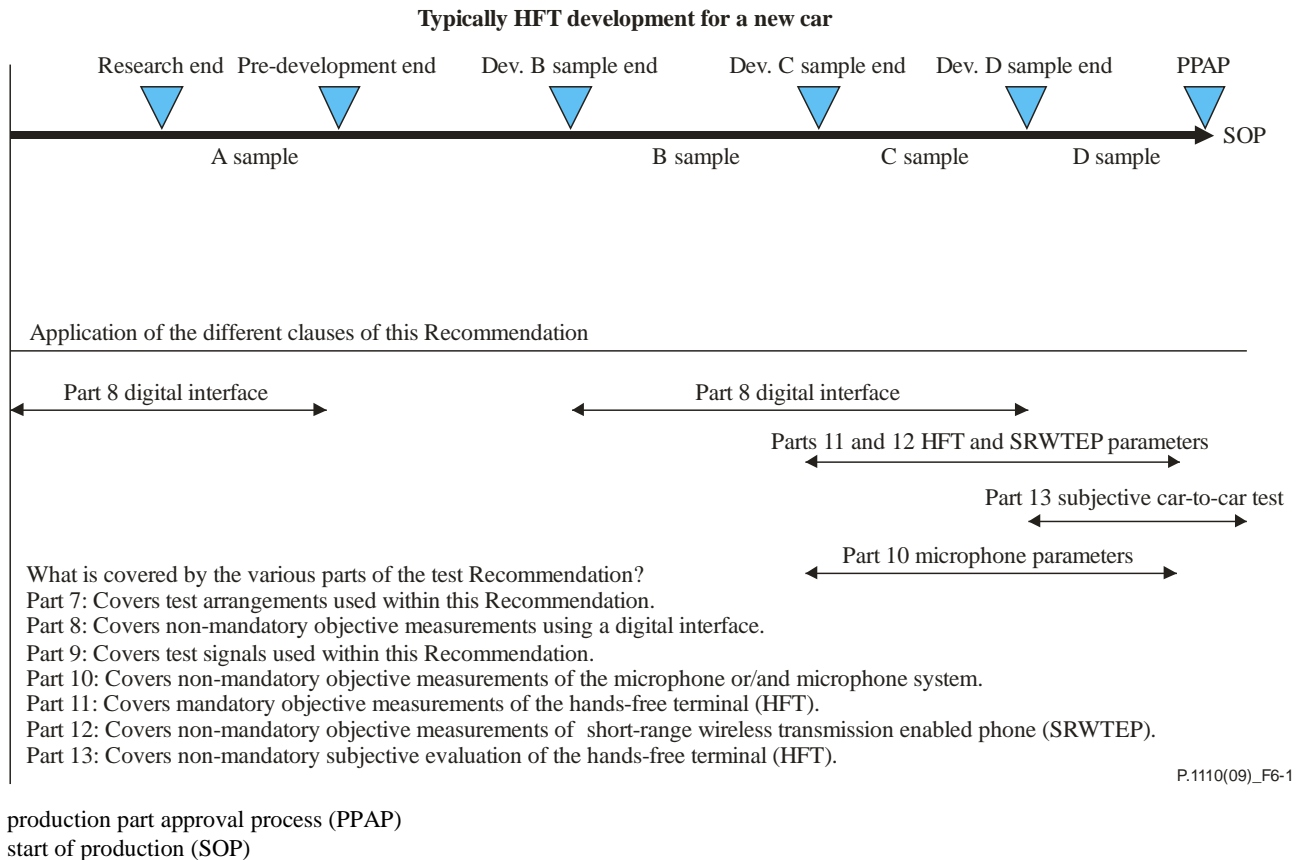
To claim compliance with these sections of the Recommendation, all requirements of the clauses in the relevant sections are passed using the procedures specified.

For the definition of the use of "shall" in this context see Note in the foreword to this Recommendation.

NOTE – This Recommendation is not intended to be used for validating tolerances in production.

## 6.2 Testing at different stages of the development cycle

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**

NOTE 1 – Additional subjective testing should be performed. There is currently work ongoing in ITU-T Study Group 12 defining test methods and test plans.

NOTE 2 – The word "non-mandatory" in the context of Figure 6-1 is only used from the car-makers point of view and does not imply that these clauses are non-mandatory in a general sense.

## 7 Test arrangement

The acoustical interface for all hands-free terminals (speakerphones and headsets) is realized by using an artificial head and torso simulator (HATS) – according to [ITU-T P.58]. The properties of the artificial head shall conform to [ITU-T P.58] for send as well as for receive 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 will be found in each clause. Table 7-1 provides an overview of the wideband speech codecs used for the tests.

**Table 7-1 – Overview of speech codecs used**

System	Codec
GSM 850, 900, 1800, 1900	AMR-WB (ITU-T G.722.2) at 12.65 kbit/s
UMTS (WCDMA)	AMR-WB (ITU-T G.722.2) at 12.65 kbit/s
CDMA2000 (IS-2000)	VMR-WB (TIA-1016) at $\leq 13.3$ kbit/s EVRC-WB (TIA-127-C) at $\leq 8.55$ kbit/s
TD-SCDMA	AMR-WB (ITU-T G.722.2) at 12.65 kbit/s
Voice over data network (VoDN)	ITU-T G.722 ITU-T G.722.2 ITU-T G.729.1 ITU-T G.711.1

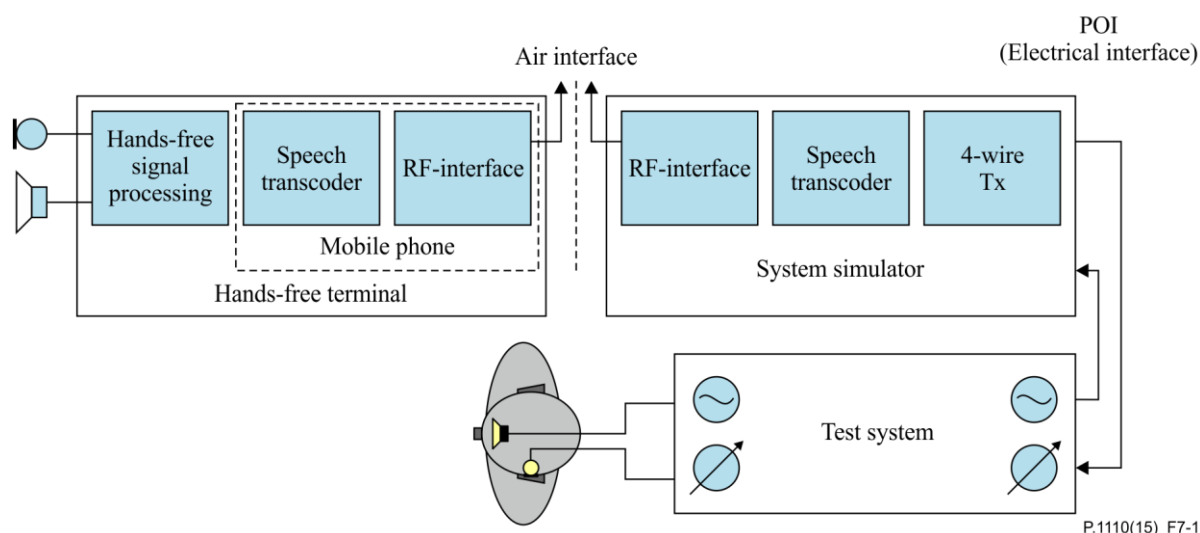
NOTE – The applicability of this Recommendation for code division multiple access (CDMA) systems has not yet been proven. Care should be taken when testing with these types of codecs to ensure that the test results observed represent the system under test, rather than the codec itself.

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

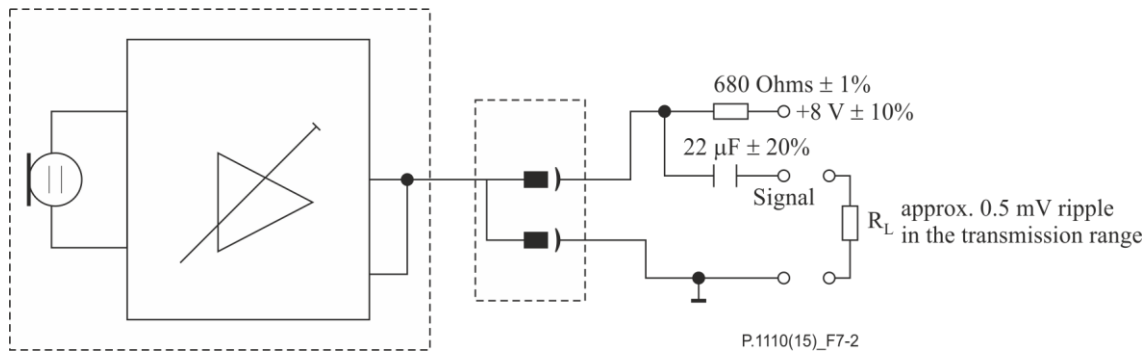
NOTE 1 – Different codecs as well as the variation of the bitrate of codecs with variable bitrates will influence the speech quality. In order to take into account "real life" conditions, bitrates used in the real network should be used for testing and optimization.

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 only use such phones for the tests which do not introduce additional speech signal processing.



**Figure 7-1 – Test arrangement for hands-free terminal**

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



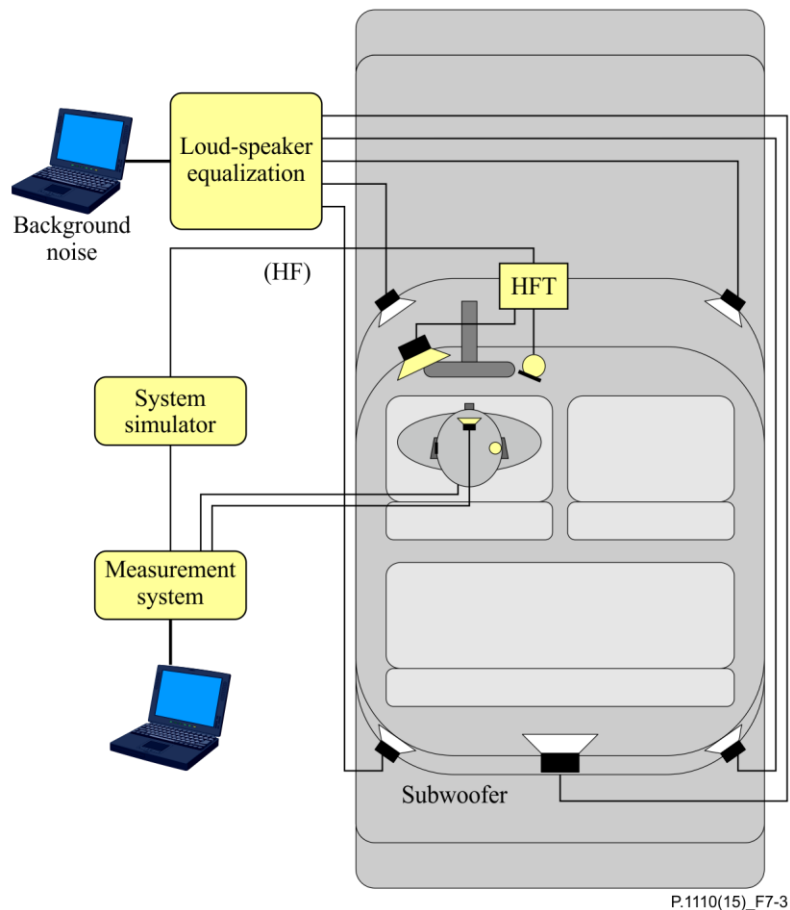
**Figure 7-2 – Test arrangement for hands-free microphones and microphone arrangements**

Care has to be taken such 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 the narrowband.  $R_L$  shall be  $> 10 \text{ k}\Omega$ .

## 7.1 Test arrangement in a car

### 7.1.1 Microphone related simulation

The transmission performances of car hands-free terminals are measured in a car cabin. In order to simulate a realistic driving situation, background noise is inserted using a 4-loudspeaker arrangement with subwoofer while measurements with background noise are conducted. This method is not a real sound-field reproduction but a simplified method mainly targeted to single microphone solutions. The simulation arrangement is shown in Figure 7-3. The test arrangement conforms to [b-ETSI ES 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  $\pm 1 \text{ dB}$ . The third octave power density spectrum between 100 Hz and 10 kHz shall not deviate more than  $\pm 3 \text{ dB}$  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 ES 202 396-1].



**Figure 7-3 – Test arrangement with background noise simulation**

For microphone arrays, distributed microphones or directional microphones a more sophisticated background noise simulation technology is recommended. This method is based on a sound-field simulation technique described in [b-ETSI TS 103 224]. Using this technology a more accurate equalization can be performed and the sound-field generated can be reproduced correctly in magnitude and phase up to a frequency of 2-3 kHz. A detailed description how to apply the method can be found in Annex F.

The background noise playback system is time-synchronized to the recording process in the measurement system in order to guarantee reproducibility of recordings in the presence of noise.

### 7.1.2 Positioning of the hands-free terminals

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

If not stated otherwise, the HATS, according to [ITU-T P.58], is positioned in the driver's seat for the measurement. The position has to be in line with an 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. Detailed information can be found in [b-SAE J941]. The size of these people 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 provided 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 mouth reference point (MRP) is defined by the test lab.

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

NOTE – Different positions of the artificial head may highly 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 signal-to-noise ratio (SNR) and/or the speech quality in the send direction may be worse.

### 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 headset hands-free terminals the average sound pressure level, calibrated at the MRP, is –4.7 dBPa.

In the case of speakerphone hands-free terminals, the sound pressure level is calibrated at the head and torso simulator – hands-free reference point (HATS-HFRP) so that the average level at the HATS-HFRP is –25.7 dBPa. The sound pressure level at the MRP has to be corrected correspondingly. The detailed description for equalization at the MRP and level correction at the HATS-HFRP can be found in [ITU-T P.581].

When testing with vehicle noise, the output level of the mouth is increased to account for the "Lombard effect". The Lombard effect refers to the change in speaking behaviour caused by acoustic noise. The level is increased by 3 dB for every 10 dB that the long-term A-weighted noise level exceeds 50 dB(A) [b-Kettler]. This relationship is shown in the following formula:

$$I(N) = \begin{cases} 0 & \text{for } N < 50 \\ 0.3(N - 50) & \text{for } 50 \leq N < 77 \\ 8.0 & \text{for } N \geq 77 \end{cases}$$

where:

$I$  = The dB increase in mouth output level due to noise level

$N$  = The long-term A-weighted noise level measured near the driver's head position

As an example, if the vehicle noise measures 70 dB(A), then the output of the mouth would be increased by 6 dB. No gain is applied for noise levels below 50 dB(A). The maximum amount of gain that can be applied is 8 dB. Vehicle noise levels are measured using a measurement microphone positioned near the driver's head position. The 3 dB speech level increase according to [ITU-T P.340] and applicable for all hands-free tests in the send direction is taken into account independently (see clause 9 on test signal levels).

### 7.1.4 Artificial ear

For speakerphone hands-free terminals, the ear signals of both ears of the artificial head are used. The artificial head is free-field or diffuse-field equalized (see clause 7.1.6.1); 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].

NOTE – In case of special insert type headsets which do not fit to the ear canal of the 3.3 or 3.4 artificial ear, a type 2 artificial ear, as defined in [ITU-T P.57], fitted with an ear canal adapter suitable for the headset under test, may be used.

### **7.1.5 Influence of the transmission system**

Measurements may be influenced by signal processing (different speech codecs, DTX, comfort noise insertion), 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 documenting this behaviour.

### **7.1.6 Calibration and equalization**

The following clauses describe preparations that have to be completed before running the tests.

#### **7.1.6.1 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):

- Diffuse-field equalization of the artificial head is used.

Equalization (for headset hands-free terminals only):

- Diffuse-field equalization of the artificial head.

#### **7.1.6.2 Reference measurement**

For the compensation of the different power density spectra of the measurement signals, it is required to reference the measured power density spectra to the power density spectra of the test signal. This is denoted as a reference measurement.

- in send direction, the reference spectrum is recorded at the MRP,
- in receive direction, the reference spectrum is recorded at the electrical interface.

### **7.1.7 System simulator settings**

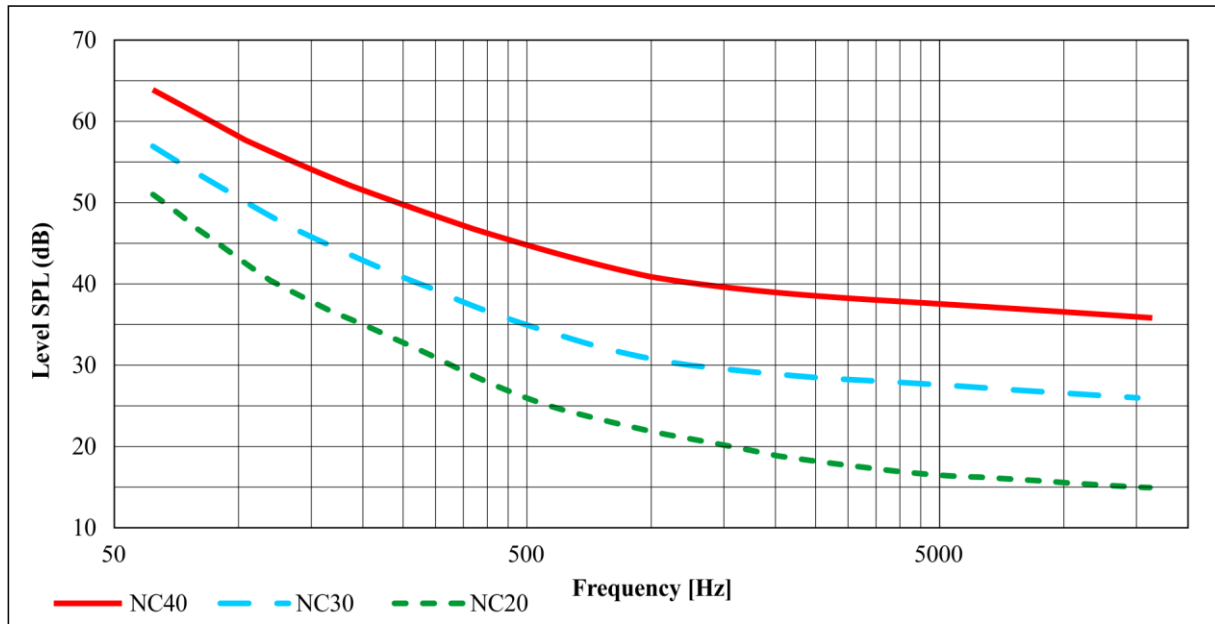
All settings of the system simulator have to ensure that the audio signal is not disturbed by any processing, and that the transmission of the HF signal is error-free. DTX shall be switched-off. For all networks, the RF-level shall be set to maximum. The settings shall be reported in the test protocol.

For measurements in global system for mobile (GSM) or universal mobile telecommunications system (UMTS) networks, the adaptive multi-rate wideband (AMR-WB) codec with a bit rate of 12.65 kbit/s is used.

### **7.1.8 Environmental conditions**

Unless specified otherwise, the background noise level in the vehicle at all measurement locations shall be less than  $-54$  dBPa(A) in conjunction with NC40 [ISO 3745].

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



P.1110(17)\_F7-4

**Figure 7-4 – NC-criteria for test environment**

#### 7.1.8.1 Verification of environmental conditions

This test is intended to be used in order to verify the environmental conditions as defined in this Recommendation.

- 1) For the measurements no test signal is used.
- 2) A free-field measurement microphone is positioned in the test room outside the car.
- 3) The room noise is in the frequency range between 20 Hz and 20 kHz. The measurement duration is 5 seconds 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.
- 4) For checking the room noise level the measured spectrum is A-weighted.
- 5) For checking the NC-criteria the octave levels of the room noise are determined from 63 Hz to 10 kHz.

## 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 hands-free (HF) devices.

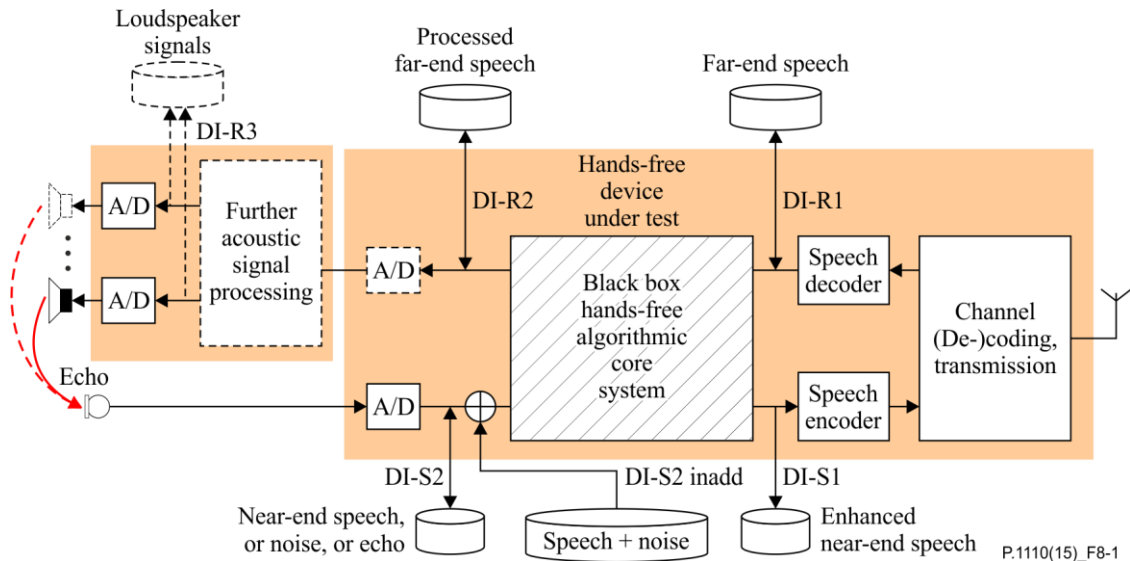
### 8.1 Interfaces and access points

Digital interfaces allow to record and investigate 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 send as well as for the receive 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 of the digital interfaces that are useful for development, debugging, and test.



**Figure 8-1 – Digital interfaces for the HF system**

The digital interfaces (DI) are called DI- $\{R|S\}n$  with 'R' standing for receive path and 'S' standing for send path. The number  $n$  is used to distinguish between different digital interfaces in the send and receive path, respectively.

DI-R1 can be used to record transmitted far-end speech (READ) or to test the HF device under test (DUT) 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 receive path. Here only READ access should be realized.

In some systems further digital signal processing may be used, connected digitally or analogue 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 send path, DI-S2 is the access point of highest 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 analogue/digital (A/D) 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 A/D conversion in the send 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 send input path to the HF system DI-S2inadd (ADD), while the HF system is in real-time operation in the send and receive 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 is necessary at all. A reduction in test effort is achieved by not testing with either background noise simulation or with real driving noise.

Finally, DI-S1 allows access to the HF system output signal in the send direction (READ). This signal gives important information about the core HF system's functionality: namely acoustic echo cancellation and noise reduction (NR). In end-to-end test simulations (near-end to far-end), it is sometimes hard to relate problematic HF equipment behaviour to the HF system stand-alone, or to its interaction with speech codec or network-sided voice enhancement devices. Investigation of signals recorded via DI-S1 may give an answer to this question.

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

- 16-bit linear pulse code modulation (PCM),
- [ITU-T G.722] at 64 kbit/s.

The sampling frequency of the digital interfaces should be 16 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 11. 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 send direction**

In a similar manner to clause 11.5.1, the one-way speech quality in send direction can be measured with stored near-end test signals (see Annex E) 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 of interconnection (POI) can be used in order to perform the test for requirement 11.5.1.1 yielding mean opinion score (MOS) LQO-W(POI), and second, the measurement can be done via the DI-S1 interface yielding MOS LQO-W(S1). Here the requirement is:

$$\text{MOS-LQOw}(S1) \geq \text{MOS-LQOw}(\text{POI}) \geq 3.6$$

The value of  $\text{DELTA} = \text{MOS-LQOw}(S1) - \text{MOS-LQOw}(\text{POI})$  can be caused by:

- the codecs and the network;
- an interaction between the HF signal processing and the codecs or 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 send direction in double talk. The test aims to help optimize 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 send direction.

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

The processing steps for the test are as follows:

- 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 clause 11.11).
- The HF system is to be processed in real-time with the speech input signals on both sides (interface DI-R1 in receive, and DI-S2inadd in send). It must be ensured that different talkers are always used for the receive and send directions. In 25 per cent of the test cases, two female voices shall be applied, in 25 per cent of the test cases two male voices shall be applied, and in 50 per cent of the test cases different genders in receive and send 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 send direction is stored via DI-S1.
- Using the echo (DI-S2), the near-end speech (DI-S2inadd), the output of the HF system in send (DI-S1), and the signal at the electrical reference POI in send, 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 POI can be extracted using the signal separation methodology as described in [b-Fingscheidt], using a Blackman window of 1024 samples with a frame shift of  $\leq 128$  samples [b-Steinert]. In a similar manner to clause 8.2.3, the requirement is stated as:

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

The MOS-LQOw analysis is performed based on [ITU-T P.863]. The value of  $\text{DELTA} = \text{MOS-LQO-w}(S1) - \text{MOS-LQO-w}(\text{POI})$  can be considered to be the speech degradation 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 are found in [ITU-T P.501]. Detailed information about the test signal used is found in the corresponding clause of this Recommendation. Wherever possible, the speech signals described in [ITU-T P.501], clause 7.3 are used.

NOTE 1 – For single talk measurements, in cases where it can be shown that the hands-free signal processing does not affect the measurement results when using a shorter version of the single talk sequence of clause 7.3.2 ([ITU-T P.501]), a shorter sequence consisting of two sentences may be used. In such event, the following two sentences (one male, one female voice) covering the low pitch frequency of male voices and the typically higher energy in the high frequency range for female voices should be used:

- "The last switch cannot be turned off" (sentence 1),

- "The hogs were fed chopped corn and garbage" (sentence 6).

When the composite source signal (CSS), according to [ITU-T P.501], is used, shaping of the wideband CSS spectrum is applied. The shaping response characteristics described in Figure 6 of [ITU-T P.501] is applied but with extension of the 5 dB/octave shaping response characteristics from 4 kHz to 8 kHz.

For wideband hands-free terminals, all test signals – which are used in receive direction – have to be band-limited. The band limitation is achieved by bandpass filtering in the frequency range between 50 Hz and 8 kHz using bandpass filtering providing 24 dB/octave. In send 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 receive direction, the band-limited test signal is measured, in send direction no band-limitation is applied.

The average signal levels for the measurements are as follows:

- –16 dBm<sub>0</sub> in receive direction (typical signal level in networks),
- –4.7 dBPa in the send direction at the MRP (typical average speech levels) (equivalent to –28.7 dBPa at the HATS-HFRP). This level applies for **headset hands-free** terminals,
- –1.7 dBPa in the send direction at the MRP for **speakerphone hands-free** terminals (typical average speech levels) (equivalent to –25.7 dBPa at the HATS-HFRP).

NOTE 2 – If different networks signal levels are to be used in a test, this is stated in the individual test. The "Lombard Effect" (increased talker speech level due to high background noise) is considered in 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 open roof in a cabriolet, have to be taken into account. Specific driving situations, e.g., driving with open window, may be taken into account as well. In general, it is recommended to conduct all tests during constant driving conditions simulating a 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 right 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.

### 9.2.1 Recording of driving noise

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

NOTE 1 – In case of microphone arrays, the simple background noise simulation technology based on [b-ETSI ES 202 396-1] is not recommended. For microphone arrays it is recommended to use the background noise recording technology as described in Annex F or to record the electrical output signals of all microphones and insert them electrically as described below. With this methodology, structure-borne noise and wind noise coupled to the microphone can also be included.

Background noise recordings are collected from the vehicle being tested and used in noise related tests. Table D.1 lists the standard set of user scenarios that noise related requirements must be tested with to be considered compliant with this Recommendation. These user scenarios are important because they define what it means to be compliant, ensure that performance is tested for some common usage scenarios, and allow for reasonable comparisons across vehicle platforms. If the main goal of testing is to directly compare different hands-free systems, then it is important to more tightly control the experimental variables listed in Table D.1 (e.g., use identical vehicles, identical routes for noise collection, identical noise recordings for testing different algorithms).

NOTE 2 – The HVAC (see Table D.1) settings may be automatically adjusted by the car during a hands-free call in order to limit the background noise. For the background noise recordings the HVAC active a hands-free call shall be used.

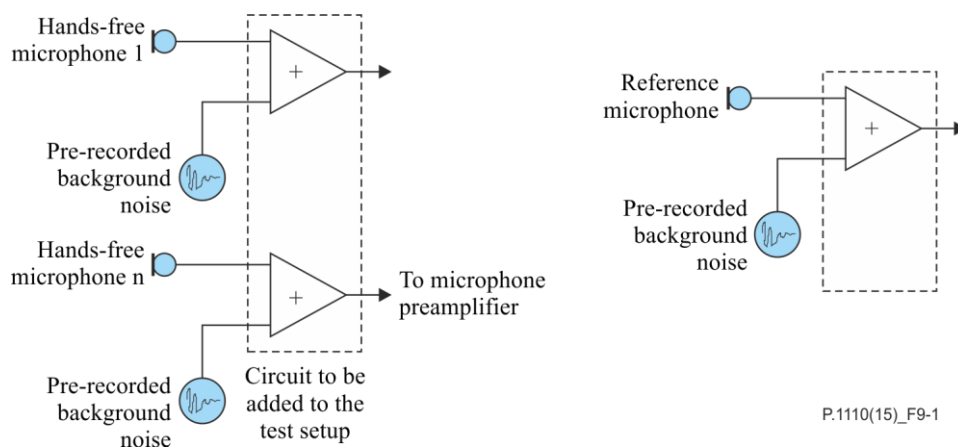
### 9.2.2 Playback of the recorded background noise

Three ways of background noise playback are recommended:

- 1) The test lab employs a 4-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 disturbances of the transmission paths between loudspeakers and hands-free microphone and the artificial head at the driver's seat. Details can be found in [b-ETSI ES 202 396-1] and Annex F.
- 2) Background noise can be inserted electrically to the microphone signal and to the reference microphone positioned close to the hands-free microphone. Therefore, background noise signals recorded at the electrical output of the hands-free microphone(s) and at the reference microphone are inserted at the electrical access point which was used for the recording. Appropriate electronics allowing the mix of the previously recorded background noise signal(s) with the microphone signal(s) at this access point have to be provided; see Figure 9-1. The test lab has to ensure the right calibration of the two signals.
- 3) Background noise can be digitally recorded at the DI-S2 interface in Figure 8-1 and later digitally inserted (added) as described in clause 8 via interface DI-S2inadd in Figure 8-1.

The type of background noise simulation chosen shall be reported in the test report.

NOTE 1 – 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.



**Figure 9-1 – Set-up for analogue electrical insertion of the pre-recorded background noise signal at the hands-free microphone(s) and the reference microphone**

NOTE 2 – Structure-borne noise is also covered with this arrangement, which is part of the microphone recording.



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

This clause is applicable to single microphones but not to the output of microphone arrays.

The requirements in this section provide general guidelines on how to test microphones for a car hands-free terminal. The requirements are chosen to provide sufficient freedom for implementers in order to achieve the best overall system performance. Therefore, especially the frequency response tolerance masks are chosen wider than for the overall hands-free terminal. Care must be taken to implement a microphone response characteristic which fits optimally to the overall system taking into account the equalization capabilities of the car hands-free signal processing.

### **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**

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.

Microphone sensitivity at 1 kHz shall be  $300 \text{ mV/Pa} \pm 3 \text{ dB}$  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 the undisturbed free sound field.
- 2) The microphone is positioned in a distance of 1 m in the acoustic centre line of the loudspeaker.
- 3) The microphone is oriented toward the loudspeaker with its direction of maximum sensitivity.
- 4) The sensitivity is determined in mV/Pa.

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

### **10.1.2 Microphone frequency response in anechoic conditions**

#### **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.

If chosen to measure the microphone sensitivity in anechoic conditions, the requirement has to be met in conjunction with the requirement for the microphone frequency response measured in the car in clause 10.2.3.

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

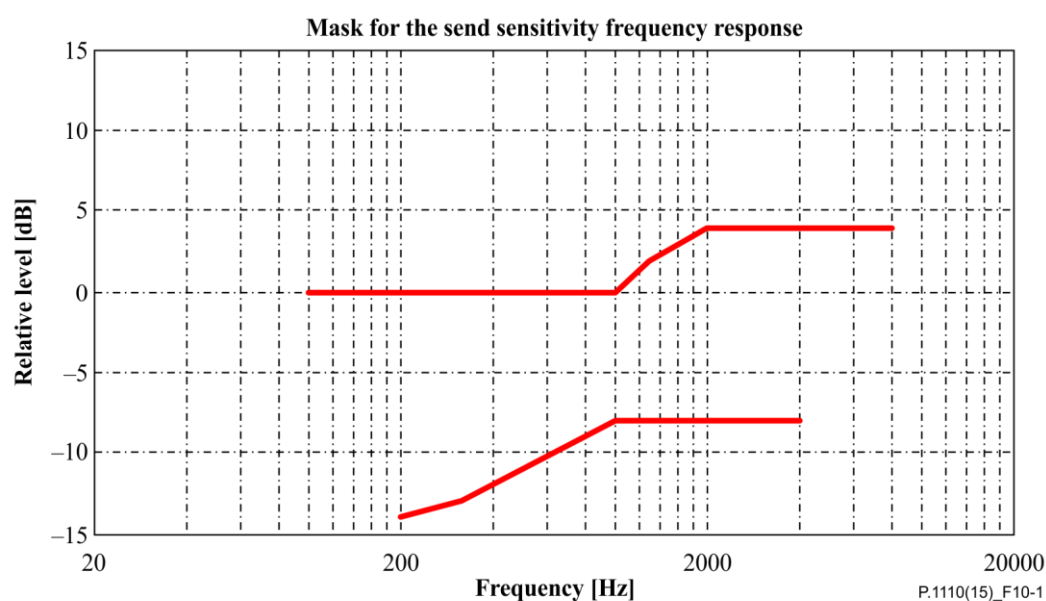
Frequency (Hz)	Upper limit	Lower limit
100	0	
125	0	
200	0	-14
315	0	-13
400	0	-12
500	0	-11
630	0	-10
1 000	0	-8
1 300	2	-8
1 600	3	-8
2 000	4	-8
3 100	4	-8
4 000	4	-8
8 000	4	

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

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

NOTE 2 – Ideally, the response characteristics of the microphone should be flat in the frequency range of wideband transmission (100 Hz – 7 kHz). However, especially in the presence of background noise, a bandwidth limitation may be desirable. No explicit recommendation can be given here since such limitation would depend on level and spectral content of the background noise and ideally should be adaptive. If, however, a bandwidth limitation is introduced, it should be made at both high and low frequencies.

NOTE 3 – Table 10-1 applies wider tolerances than Table 10-2 to achieve the freedom of adapting the microphones frequency response to the needs in the car.



**Figure 10-1 – Microphone send frequency response mask (Fig. is informative)**

### **10.1.2.2 Test**

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

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

### **10.1.3 Microphone directional characteristics**

The directional characteristic of a microphone is 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 direction of highest sensitivity and the sensitivity at the angle 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, also lower front to back ratios can 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 the undisturbed free sound field.
- 2) The microphone is positioned in a distance of 1 m in the acoustic centre line of the loudspeaker.
- 3) The first measurement is done with the microphone oriented to the loudspeaker with its direction of maximum sensitivity. The second measurement is done with the microphone oriented toward the loudspeaker with 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 [b-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 dB SPL) at the position of the microphone shall be less than 1 per cent in the wideband frequency range.

#### **10.1.4.2 Test**

- 1) The test signal is a sine wave with a frequency of 300 Hz, 500 Hz, 1 kHz, 2 kHz at a level of 0 dBPa.
- 2) The microphone is positioned in an acoustic centre line of the loudspeaker.
- 3) The microphone is oriented toward the loudspeaker with its direction of maximum sensitivity.

- 4) The total harmonic distortion is expressed in per cent.

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

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

### **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 per cent in the wideband 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 shall be higher than +12 dBPa (106 dB SPL) 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% total harmonic distortion.
- 2) The microphone is positioned in an acoustic centre line of the loudspeaker.
- 3) The microphone is oriented toward the loudspeaker with its direction of maximum sensitivity.
- 4) The maximum sound pressure level is expressed in dB SPL or dB Pa.

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

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

#### **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 8 kHz, A-weighting has to be applied.
- 4) The self-noise is expressed in dBV(A).

Care has to be taken such 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 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 talker, as within the near field of a sound source (in a vehicle this is up to 80-100 cm)<sup>1</sup> 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 a better performance than a microphone array, which is placed further away.
- There has to be a direct path between the talker'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 talker's mouth. If different seating positions or several talkers are to be covered by one microphone, a compromise for the microphone position has to be found as the direction of the highest sensitivity might not cover all positions or talkers. However, this often means a significantly reduced performance in comparison to an optimal alignment of the microphone for a single talker. In this case, the application of additional microphones might be considered to achieve an optimal speech quality.
- A direct airstream to the microphone, e.g., from the air conditioning, has to be avoided as the speech signal might be highly disturbed by wind buffeting.
- Saturation of the microphone by loudspeakers nearby, 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-borne sound to the microphone has to be avoided.

NOTE 1 – When the microphone 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 in the car, it is recommended to measure the microphone 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:

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

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

However, depending on specific electrical/acoustical designs, arrangements inside the car or others the sensitivity requirement may be different. Therefore, this requirement has to be adapted to the individual arrangement in 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  $-10 \text{ dBPa}$  measured at the MRP.

---

<sup>1</sup> 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 critical distance.

- 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.
- 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 levels 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 – 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

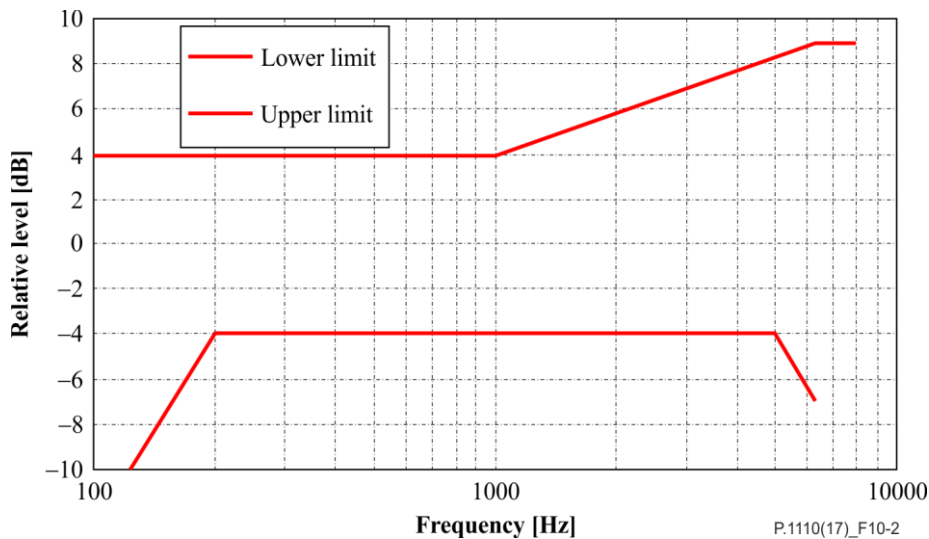
The microphone frequency response is measured from the MRP to the output of the test circuit, according to Figure 7-2.

The tolerance mask for the sensitivity frequency response in send 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 send sensitivity frequency response**

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

NOTE – The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) – logarithmic (Hz) scale. All sensitivity values are expressed in dB on an arbitrary scale.



**Figure 10-2 – Send frequency response mask (Fig. is informative)**

NOTE – Ideally, the response characteristics of the microphone should be flat in the frequency range of wideband transmission (100 Hz – 7 kHz). However, especially in the presence of background noise, a bandwidth limitation may be desirable. No explicit recommendation can be given here since such limitation would depend on the level and spectral content of the background noise and ideally should be adaptive. If, however, a bandwidth limitation is introduced, it should be made at both, the high and low frequencies.

### 10.2.3.2 Test

- 1) The test arrangement is according to clause 7.1.
- 2) The test signal, the British-English single talk sequence, described in clause 7.3.2 of [ITU-T P.501] is used. Alternatively, a periodic noise signal or CS signal according to [ITU-T P.501] can be used. The correct activation of the measurement object during the measurement has to be ensured by the test lab. The artificial mouth is equalized at the MRP, the test signal level shall be  $-1.7$  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  $-25.7$  dBPa.
- 3) The measured power density spectrum at the MRP is used as the reference power density spectrum for determining the frequency response in send direction.
- 4) The sensitivity frequency response is determined in third octave intervals as given by [IEC 61260] for frequencies of 100 Hz and 8 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 send direction measured at the output of the test circuit according to Figure 7-2 in quiet conditions shall be less than  $-72$  dBV(A). Spectral peaks in the frequency domain shall not exceed the averaged spectrum by more than 10 dB.

NOTE 1 – It is recognized that fan noise, which varies from car to car and depends upon the relative positioning of the microphone and fan, 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.

NOTE 2 – A maximum idle channel noise of  $-66$  dBV(A) may be acceptable if the overall system performance is not impaired by the additional microphone noise.

### 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 four 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  $-25.7$  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 8 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 one 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 the fast Fourier transform (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 A-weighting.
- 4) Spectral peaks are measured in the frequency domain. The frequency spectrum of the A-weighted idle channel noise is measured by a spectral analysis having a noise bandwidth of 8.79 Hz (determined using FFT 8 k samples/48 kHz sampling rate with Hanning window or equivalent). The idle channel noise spectrum is stated in dB. A smoothed average idle channel noise spectrum is calculated by a moving average (arithmetic mean) 1/3rd octave wide across the idle noise channel spectrum stated in dB (linear average in dB of all FFT bins in the range from  $2^{(-1/6)}f$  to  $2^{(+1/6)}f$ ). Peaks in the idle channel noise spectrum are compared against a smoothed average idle channel noise spectrum.

### 10.2.5 SNR improvement (in the car)

#### 10.2.5.1 Requirements

The SNR improvement provided by the microphone should be  $\geq 3$  dB.

#### 10.2.5.2 Test

- 1) The test arrangement is according to clause 7.1. The hands-free microphone is used for the measurements.
- 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. The background noise has to be inserted at minimum five seconds before the actual measurement starts in order to allow adaptation of possible background noise reduction algorithms.
- 3) The send signal (background noise only) is recorded at the output of the test circuit according to Figure 7-2. The A-weighted output level LBG<sub>Nhftmic</sub> is measured between 100 Hz and 8 kHz. The result is scaled in dBV/Pa(A).
- 4) The speech signal (without background noise) is simulated by a composite source signal according to [ITU-T P.501] with a duration of  $\geq 2$  CS sequences. The A-weighted output level LSh<sub>ftmic</sub> is measured between 100 Hz and 8 kHz. The result is scaled in dBV/Pa(A). The (unweighted) frequency response characteristic of the hands-free microphone is



determined from this signal, according to the procedure described in clause 10.2.3.2, and stored.

- 5) The  $SNR_{hftmic}$  is determined by  $SNR_{hftmic} = L_{Shftmic} - L_{BGNhftmic}$ .
- 6) An omnidirectional microphone with a flat frequency response characteristic is positioned close to the car hands-free microphone. The measured spectrum of the omnidirectional microphone is weighted with the frequency response characteristic measured in step 4.
- 7) Steps 1-4 are repeated using the omnidirectional microphone. The measured levels are  $L_{BGNomni}$  and  $L_{Somni}$ .
- 8) The  $SNR_{omni}$  is determined by  $SNR_{omni} = L_{Somni} - L_{BGNomni}$ .
- 9) The SNR improvement is calculated by  $SNR_{hftmic} - SNR_{omni}$ .

## 11 Measurement parameters and requirements for hands-free terminals

NOTE – The measurements described in this clause do not cover bandwidth extension technologies. If such techniques are used, they should be disabled for the tests.

### 11.1 Preparation measurements

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

### 11.2 Delay

#### 11.2.1 Requirements

In general the delay  $T_{rtd}$  consists of an access specific delay  $T_{as}$  and the implementation dependent delay  $T_{rtdimp}$ .

The access-specific specifications define the access specific delays  $T_{as}$  which have to be taken into account when measuring  $T_{rtd}$ .

The HFT implementation dependant delay  $T_{rtdimp}$  consists of:

- the HFT signal processing in send and receive
- the air-paths from mouth to microphone and from the loudspeakers to the ear

The HFT delays in the send and receive directions are defined as:

- The HFT delay in the send (uplink) direction  $T_s$  is the delay between the first acoustic event at the MRP to the last bit of the corresponding speech frame at the UE antenna
- The HFT delay in the receive (downlink) direction  $T_r$  is the delay between the first bit of a speech frame at the UE antenna and the first acoustic event at the DRP corresponding to that speech frame

$T_{rtd} = T_s + T_r$  (the delay in the send direction  $T_s$  plus the delay in the receive direction  $T_r$ ) shall be less than 170 ms.

For hands-free implementations using SRW in the connection the round trip delay consist of different components. In such connections  $T_{as} = T_{SRW}$ . Consequently the roundtrip delay  $T_{rtd}$  in such connection consists of the implementation dependent delay and the short range wireless delay:  $T_{rtd} = T_{rtdimp} + T_{rtdsrw}...$

In case a SRW connection is used the implementation dependent delay  $T_{rtdimp} = T_{imps} + T_{impr}$  shall be less than 70ms.

NOTE 1 – The access specific delays can be found in the relevant specification for the access technologies used and need to be calculated based on the information provided in these specifications. When deriving  $T_{as}$  from these specification it is assumed that any speech signal processing is deactivated and does not introduce any additional delay. For 3GPP UMTS circuit-switched speech and 3GPP LTE MTSI-based speech, definitions, test methods, performance objectives and requirements are found in [b-3GPP TS 26.131] and [b-3GPP TS 26.132].

NOTE 2 – Regarding the user effect of mouth-to-ear delay to the conversational quality in handset mode, guidance is found in ITU-T Recommendation [b-ITU-T G.114].

NOTE 3 – When providing state of the art low delay implementations the delay introduced by the hands-free signal processing  $T_{rdimp}$  should not exceed 70 ms.

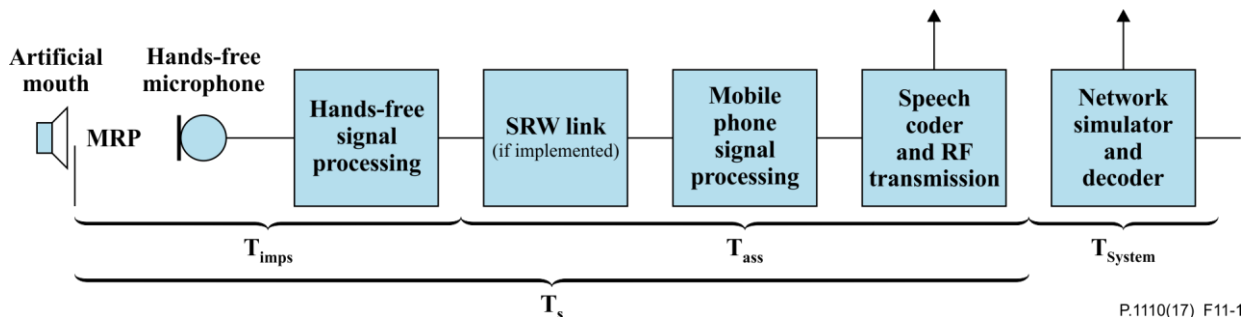
## 11.2.2 Delay in send direction

### 11.2.2.1 Test

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

$$T_s + T_{system}$$

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



**Figure 11-1 – Different blocks contributing to the delay in send direction**

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

- 1) For the measurements, a CSS, according to [ITU-T P.501], is used. The pseudo-random noise (pn)-part of the CSS shall be longer than the maximum expected delay. It is recommended to use a pn sequence of 16 k samples with 48 kHz sampling rate or equivalent. The test signal level is  $-4.7$  dBPa at the MRP, for headset hands-free terminals. For speakerphone hands-free terminals, the test signal level is adjusted to  $-25.7$  dBPa at the HATS-HFRP (see [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 according 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.

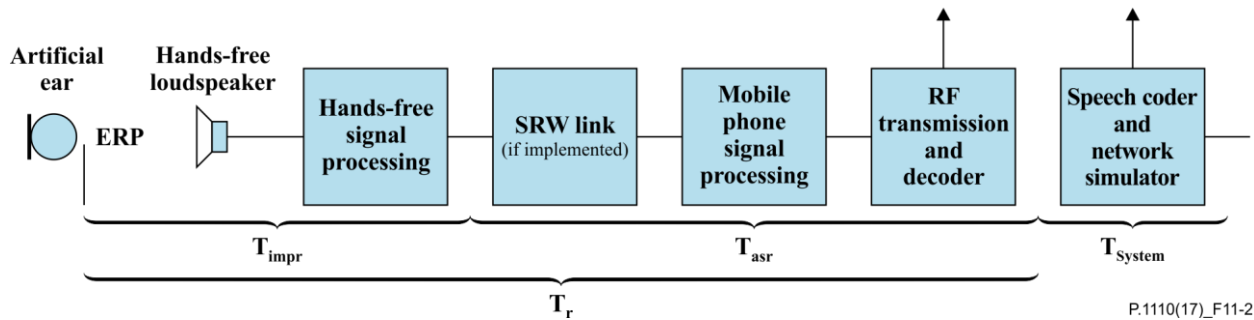
### 11.2.3 Delay in receive direction

#### 11.2.3.1 Test

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

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

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



**Figure 11-2 – Different blocks contributing to the delay in receive direction**

The system delay  $T_{\text{system}}$  depends on the transmission system and on the network simulator used. The delay  $T_{\text{system}}$  must be known.

- 1) For the measurements, a CSS, according to [ITU-T P.501], is used. The pseudo-random noise (pn)-part of the CSS shall be longer than the maximum expected delay. It is recommended to use a pn sequence of 16 k samples with 48 kHz sampling rate or equivalent. 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 or diffuse-field equalized (see clause 7.1.6.1), 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 measured sound pressure is diffuse-field equalized according to [ITU-T P.58].
- 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 send loudness rating (SLR)/receive loudness rating (RLR) from/to the electrical reference point (POI) shall be:

For speakerphone hands-free terminals:

- SLR =  $13 \text{ dB} \pm 4 \text{ dB}$ ,
- RLR =  $2 \text{ dB} \pm 4 \text{ dB}$ .

For headset hands-free terminals:

- SLR = 8 dB  $\pm$  4 dB,
- RLR = 2 dB  $\pm$  4 dB.

For binaural headset hands-free terminals:

- SLR = 8 dB  $\pm$  4 dB,
- RLR (bin) = 8 dB  $\pm$  4 dB for each earphone.

If a user-specific volume control is provided, the requirements 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 signal-to-noise ratio (S/N)  $\geq$  6 dB for all signal and noise conditions. This will allow sufficient loudness of the speech signal in receive 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 so as 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 send direction

- 1) The test signal used for the measurements shall be the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501]. The test signal is equalized at the MRP, the test signal level is  $-4.7$  dBPa at the MRP for headset hands-free terminals. 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  $-25.7$  dBPa.

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

- 2) The test arrangement is according to clause 7.1. The send sensitivity is calculated from each band of the 20 frequencies given in Table A.2 of [ITU-T P.79], bands 1-20.

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 terms of dBV/Pa, and the SLR shall be calculated according to Annex A of [ITU-T P.79].

### 11.3.3 Test loudness rating in receive direction

- 1) The test signal used for the measurements shall be the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501]. The test signal is  $-16$  dBm<sub>0</sub>, 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 diffuse-field equalized (see clause 7.1.6.1) according to [ITU-T P.581]. The equalized output signals of both artificial ears are used for the measurement. The equalized output signal of each artificial ear is power-averaged on the total time of analysis; the "right" and "left" signals are voltage-summed for each 1/3 octave band.

For headset hands-free terminals, the sound pressure is measured at the DRP of the right ear and corrected to the ear reference point (ERP) according to [ITU-T P.57]. The receive sensitivity is determined by the bands 1-20 according to Table A.2 of [ITU-T P.79].

For binaural headset hands-free terminals, the sound pressure is measured separately at the DRP of the right ear and respectively at the left ear and corrected to the ERP, according to

[ITU-T P.57]. The receive sensitivity is determined by the bands 1-20 according to Table A.2 of [ITU-T P.79].

In case of headset measurements the tests are repeated five times, in conformance with [ITU-T P.380]. The results are averaged (averaged value in dB, for each frequency).

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 terms of dBPa/V and the RLR shall be calculated according to Annex A of [ITU-T P.79], without the  $L_E$  factor.
- 4) For speakerphone hands-free terminals, the correction 8 dB, according to [ITU-T P.340], is used for the correction of the measurement results.
- 5) The test is repeated for maximum volume control setting.

### **11.3.4 Variation of loudness rating in send 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  $\pm 0.5$  dB from the SLR measured with nominal signal level.

NOTE – Under certain conditions, the use of automatic gain control (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 the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501]. The test signals are  $-7.7$  dBPa and  $+1.3$  dBPa, measured at the MRP, for headset hands-free terminals. 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 send sensitivity.

- 2) The test arrangement is according to clause 7.1. The send sensitivity is calculated from each band of the 20 frequencies given in Table A.2 of [ITU-T P.79], bands 1-20.  
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 terms of dBV/Pa and the SLR shall be calculated according to Annex A of [ITU-T P.79].
- 4) For both signal levels, the measured result is compared to the measured with nominal signal level.

### **11.3.5 Variation of loudness rating in receive direction**

#### **11.3.5.1 Requirements**

With nominal volume control setting for network signal level variations of  $\pm 5$  dB from the nominal signal level, the measured RLR shall not deviate more than  $\pm 0.5$  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 the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501]. 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 diffuse-field equalized (see clause 7.1.6.1) according to [ITU-T P.581]. The equalized output signals of both artificial ears are used for the measurement. The equalized output signal of each artificial ear is power-averaged on the total time of analysis; the "right" and "left" signals are voltage-summed for each 1/3 octave band.

For headset hands-free terminals, the sound pressure is measured at the DRP of the right ear and corrected to the ERP, according to [ITU-T P.57]. The receive sensitivity is determined by the bands 1-20 according to Table A.2 of [ITU-T P.79].

For binaural headset hands-free terminals, the sound pressure is measured separately at the DRP of the right ear and respectively at the left ear and corrected to the ERP, according to [ITU-T P.57]. The receive sensitivity is determined by the bands 1-20 according to Table A.2 of [ITU-T P.79].

In case of headset measurements the tests are repeated five times, in conformance with [ITU-T P.380]. The results are averaged (averaged value in dB, for each frequency).

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 terms of dBPa/V and the RLR shall be calculated according to Annex A of [ITU-T P.79], without the  $L_E$  factor.
- 4) For speakerphone hands-free terminals, the correction 8 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 measured with nominal signal level.

## 11.4 Sensitivity frequency responses

### 11.4.1 Send sensitivity frequency response

#### 11.4.1.1 Requirements

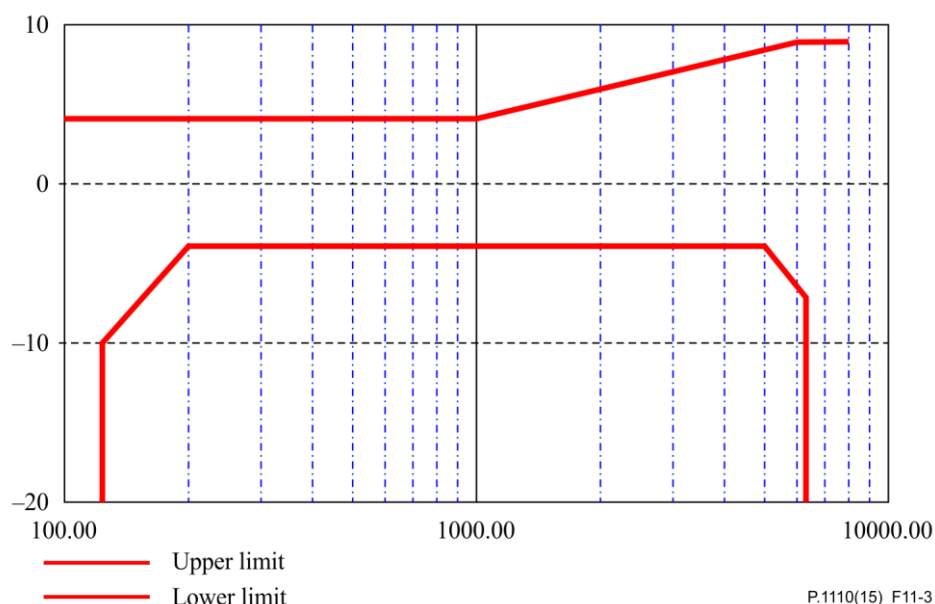
The send sensitivity frequency response is measured from the MRP to the output of the speech codec at the electrical point (output of the system simulators, POI).

The tolerance mask for the send 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, as shown in Figure 11-3.

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

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

NOTE – The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) – logarithmic (Hz) scale. All sensitivity values are expressed in dB on an arbitrary scale.



NOTE – Ideally, the response characteristics of the microphone should be flat in the frequency range of wideband transmission (100 Hz – 7 kHz). However, especially in the presence of background noise, a bandwidth limitation may be desirable. No explicit recommendation can be given here since such limitation would depend on the level and spectral content of the background noise and ideally should be adaptive. If, however, a bandwidth limitation is introduced, it should be made at both, the high and low frequencies.

**Figure 11-3 – Tolerance mask for the send sensitivity frequency response (Fig. is informative)**

#### 11.4.1.2 Test

- 1) The test signal used for the measurements shall be the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501]. The test signal is equalized at the MRP, the test signal level is  $-4.7$  dBPa at the MRP for headset hands-free terminals. 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  $-25.7$  dBPa.

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

- 2) The test arrangement is according to clause 7.1. The sensitivity frequency response is determined in third octave intervals as given by [IEC 61260] for frequencies of 100 Hz and 8 kHz inclusive. In each third octave band, the level of the measured signal is referred to the level of the reference signal averaged over the complete test sequence length.

3) The sensitivity is determined in dBV/Pa.

## 11.4.2 Receive sensitivity frequency response

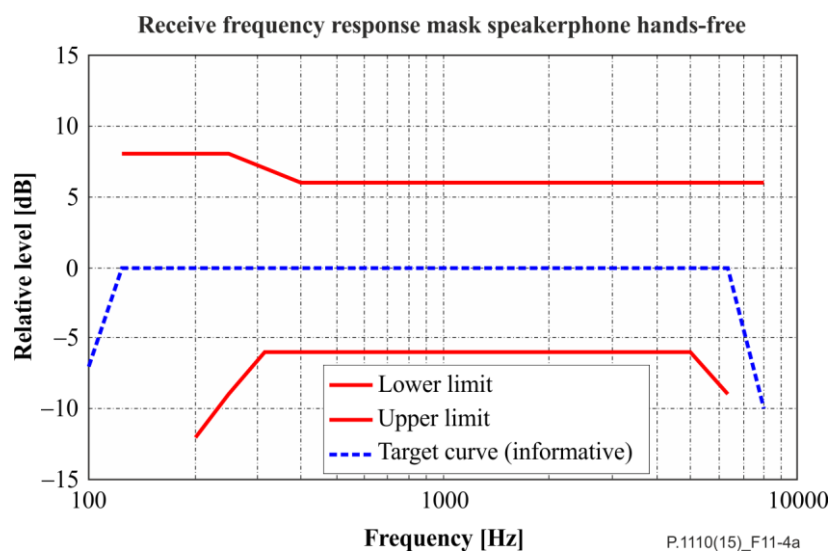
### 11.4.2.1 Requirements

The receive sensitivity frequency response is measured from the electrical reference point (input of the system simulators, POI) to the DRP when headset hands-free terminals are measured. The HATS is diffuse-field equalized. For speakerphone hands-free terminals, the sound pressure of the free-field or diffuse-field equalized (see clause 7.1.6.1) HATS is measured.

The tolerance mask for the receive 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 11-2a – Tolerance mask for the receive sensitivity frequency response for speakerphone hands-free**

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



**Figure 11-4a – Receive frequency response mask speakerphone hands-free (Fig. is informative)**

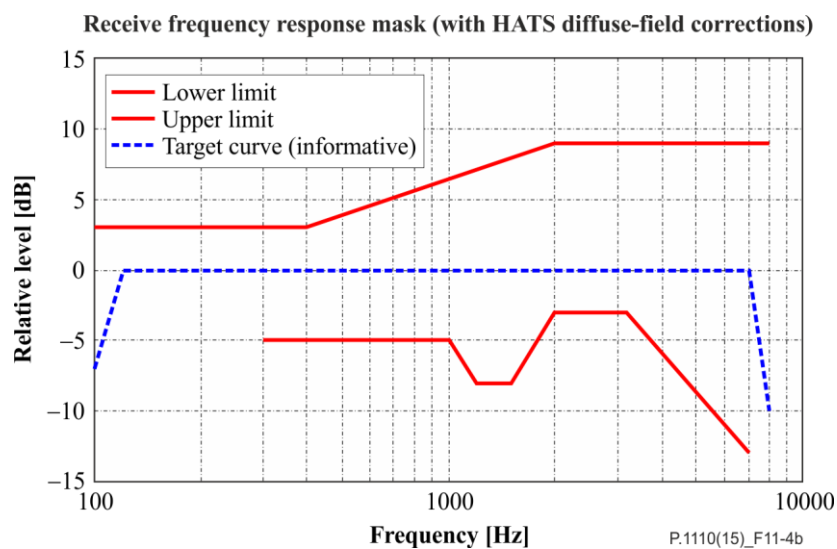


**Table 11-2b – Tolerance mask for the receive sensitivity frequency response for headset hands-free**

Frequency	Upper limit	Lower limit
100 Hz	3 dB	
120 Hz	3 dB	-5 dB
200 Hz	3 dB	-5 dB
400 Hz	3 dB	-5 dB
1 010 Hz	(Note 1)	-5 dB
1 200 Hz	(Note 1)	-8 dB
1 500 Hz	(Note 1)	-8 dB
2 000 Hz	9 dB	-3 dB
3 200 Hz	9 dB	-3 dB
7 000 Hz	9 dB	-13 dB
8 000 Hz	9 dB	

NOTE 1 – The limit curves shall be determined by straight lines joining successive coordinates given in the table, where frequency response is plotted on a linear dB scale against frequency on a logarithmic scale. The mask is a floating or "best fit" mask.

NOTE 2 – The basis for the target frequency responses in send and receive is the orthotelephonic reference response which is measured between two subjects in 1 m distance under free field conditions and is assuming an ideal receive characteristic. Under these conditions, the overall frequency response shows a rising slope. In contrast to other standards, this Recommendation no longer uses the ERP as the reference point for receive but the free-field. With the concept of diffuse-field based receive measurements, a rising slope for the overall frequency response is achieved by a flat target frequency response in send and a diffuse-field based receive frequency response.



**Figure 11-4b – Receive frequency response mask (with HATS diffuse-field corrections) (Fig. is informative)**

#### 11.4.2.2 Test

- 1) The test signal used for the measurements shall be the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501]. The test signal is -16 dBm<sub>0</sub>, 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 or diffuse-field equalized (see clause 7.1.6.1) according to [ITU-T P.581]. The equalized output signals of both artificial ears are used for the measurement. The equalized output signal of each artificial ear is power-averaged on the total time of analysis; the "right" and "left" signals are voltage-summed for each 1/3 octave band frequency band.

For headset hands-free terminals, the sound pressure is measured at the DRP. The HATS is diffuse-field equalized, as described in [ITU-T P.581]. The diffuse-field correction, as defined in [ITU-T P.58], is applied. The receive sensitivity frequency response is determined in third octaves as given in [IEC 61260] for frequencies from 100 Hz to 8 kHz, inclusive. In each third octave band, the level of the measured signal is referred to the level of the reference signal, averaged over the complete test sequence length.

For binaural headset hands-free terminals, the sound pressure is measured separately at the DRP of the right ear and respectively at the left ear. The HATS is diffuse-field equalized, as described in [ITU-T P.581]. The diffuse-field correction, as defined in [ITU-T P.58], is applied. The receive sensitivity frequency response is determined separately for each earphone in third octaves as given in [IEC 61260] for frequencies from 100 Hz to 8 kHz, inclusive. In each third octave band, the level of the measured signal is referred to the level of the reference signal, averaged over the complete test sequence length.

In case of headset measurements, the tests are repeated five times, in conformance with [ITU-T P.380]. The results are averaged (averaged value in dB, for each frequency).

- 3) The sensitivity is determined in dBPa/V.

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

## **11.5 Speech quality during single talk**

### **11.5.1 One-way speech quality in send direction**

#### **11.5.1.1 Requirement**

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

$$\text{MOS-LQO}_w \geq 3.6$$

NOTE – At this time, this requirement is non-normative because there is currently no normative test method (see clause 11.5.1.2).

#### **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 receive direction**

#### **11.5.2.1 Requirement**

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

$$\text{MOS-LQO}_w \geq 3.6$$

NOTE – This requirement is non-normative at this point in time because there is currently no normative test method (see clause 11.5.2.2).

### 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 Listening speech quality stability

### 11.6.1 Listening speech quality stability in send direction

Listening quality stability during a call (if the position or transmission characteristics change during the call (or during several different calls)) takes into account degradations generated on the signal by the transmission link impairment and the phone position.

In case of systems using SRW transmission, it is the purpose to verify the integration of SRW radio network by evaluating change of speech quality over time. This will help detect problems with the radio frequency (RF) coverage inside the car cabin and verify the error concealment (packet loss, bit errors) caused by the weak RF link or interference with other radios.

A guidance to proceed is:

- 1) Check RF coverage from the SRW unit to possible mobile positions (protocol analyser).
- 2) Identify weak and bad reception areas inside the vehicle (protocol analyser).  
Check bit error rate, rate of packet loss, etc.
- 3) Use the speech quality measurement to rate the quality of error concealment in the weak areas identified in step 2 above and compare with measurements from areas with good coverage.
- 4) Identify possible issues from interference with parallel SRW links or other networks. This will check how the SRW can handle interference and change to undisturbed channels. Speech quality measurements can be used to see the performance of handling these problems.

It is recommended to use a known phone as a reference.

#### 11.6.1.1 Requirement

For stability indicator about listening speech quality, THRESHOLD1 = 0.1 and the linear weighting function applies in order to express stability (ST-MOS) on a 0 to 100 scale. By definition, stability equals 100 when no variations occur and stability ST-MOS equals 0 when MOS-LQOw variation is equal to or more than 0.4.

ST-MOS shall be  $\geq 90$

The requirement applies only when nominal MOS-LQOw is fulfilled.

#### 11.6.1.2 Test

Several measurements of the MOS-LQOw score performed according to clause 12.5.1.1 are performed in series within the same call or for different calls, within the test arrangement defined below. Typically, for the same call, a measurement each 20 seconds is enough. The results are reported in terms of statistics.

The assessment of listening speech quality stability is performed in five steps:

- 1) To measure the MOS-LQOw periodically over the duration of one communication. N measurements provide N MOS-LQOw values.
- 2) For each MOS-LQOw (i) value i from 2 to N, MOS-LQOw\_GAP (i) is calculated as the absolute difference with the previous value MOS-LQOw (i-1):
  - MOS-LQOw\_GAP (i) = |MOS-LQOw (i) - MOS-LQOw (i-1)|

- 3) In order to take into account the subjective perception and measurement accuracy, MOS-LQOw\_GAP (i) is set to 0 when the difference is equal to or lower than THRESHOLD1:
  - if  $\text{MOS-LQOw\_GAP (i)} > (2 * \text{THRESHOLD1})$ , then  $\text{MOS-LQOw\_GAP (i)} = \text{MOS-LQOw\_GAP (i)}$
  - if  $\text{THRESHOLD1} < \text{MOS-LQOw\_GAP (i)} \leq 2 * \text{THRESHOLD1}$ , then  $\text{MOS-LQOw\_GAP (i)} = [\text{MOS-LQOw\_GAP (i)} * 2] - (2 * \text{THRESHOLD1})$
  - if  $\text{MOS-LQOw\_GAP (i)} \leq \text{THRESHOLD1}$ , then  $\text{MOS-LQOw\_GAP (i)} = 0$
- 4) The instability (INS\_MOS-LQOw) associated to the MOS-LQOw over the entire N measurements is defined by mean value of MOS-LQOw\_GAP (i).
  - $\text{INS\_MOS-LQOw} = 1 / (N - 1) \sum \text{MOS-LQOw\_GAP (i)}$  with  $i = [2:N]$
- 5) A linear weighting function is applied in order to express stability ST-MOS-LQOw on a 0 to 100 scale.

This formulation is used to determine the listening quality stability (ST-MOS) as:

$$\text{ST-MOS} = 100 - (250 * \text{INS\_MOS})$$

and

$$\text{ST-MOS} = 0 \text{ if } [100 - (250 * \text{INS\_MOS})] < 0$$

When ST-MOS is calculated within a single call, the call shall be longer than three minutes (recommended duration being between three and five minutes). The duration of each measurement depends on the length of the speech samples used for the test as described in this clause.

(As an example, if a sample is 15 seconds and the analysis is done every 20 seconds, a minimum of 10 values will be measured.)

## 11.6.2 Listening speech quality stability in receive direction

### 11.6.2.1 Requirement

For stability indicator about listening speech quality, THRESHOLD1 = 0.1 and the linear weighting function applies in order to express stability (ST-MOS) on a 0 to 100 scale. By definition, stability equals 100 when no variations occur and stability ST-MOS equals 0 when MOS-LQOw variation is equal to or more than 0.4.

$$\text{ST-MOS shall be } \geq 90$$

The requirement applies only when nominal MOS-LQOw is fulfilled.

### 11.6.2.2 Test

The test procedure is described in clause 11.6.1.2.

## 11.7 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), a background noise level of  $-64$  dBPa(A) shall not be exceeded.

## 11.7.1 Idle channel noise in send direction

### 11.7.1.1 Requirements

The maximum idle channel noise in send direction, measured at the electrical reference point (POI) in quiet conditions, shall be less than  $\leq -64$  dBm0(A).

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

NOTE 1 – Care should be taken that the measured spectral peaks result from the hands-free terminal under test and not from other sources e.g. noise produced by the car in quiet conditions.

NOTE 2 – In case spectral peaks higher than 10 dB above the average noise floor are produced by the hands-free terminal but considered to be inaudible due to the very low noise floor produced by the hands-free terminal in average, it is the responsibility of the test lab to demonstrate the desired performance subjectively.

### 11.7.1.2 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 four 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  $-4.7$  dBPa for headset hands-free terminals, measured at the MRP and  $-25.7$  dBPa for speakerphone hands-free terminals, 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 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 for the time window is one 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 A-weighting
- 4) Spectral peaks are measured in the frequency domain. The frequency spectrum of the A-weighted idle channel noise is measured by a spectral analysis having a noise bandwidth of 8.79 Hz (determined using FFT 8 k samples/48 kHz sampling rate with Hanning window or equivalent). The idle channel noise spectrum is stated in dB(A). A smoothed average idle channel noise spectrum is calculated by a moving average (arithmetic mean) 1/3rd octave wide across the idle noise channel spectrum stated in dB (linear average in dB of all FFT bins in the range from  $2^{(-1/6)}f$  to  $2^{(+1/6)}f$ ). Peaks in the idle channel noise spectrum are compared against a smoothed average idle channel noise spectrum.

## 11.7.2 Idle channel noise in receive direction

### 11.7.2.1 Requirements

The requirements for the maximum noise produced by the hands-free terminal, in case no signal is applied to the receive direction, is 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 shall occur.

NOTE 1 – Care should be taken that the measured spectral peaks result from the hands-free terminal under test and not from other sources e.g. noise produced by the car in quiet conditions.

NOTE 2 – In case spectral peaks higher than 10 dB above the average noise floor are produced by the hands-free terminal but considered to be inaudible due to the very low noise floor produced by the hands-free terminal in average, it is the responsibility of the test lab to demonstrate the desired performance subjectively.

### 11.7.2.2 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 four composite source signals according to [ITU-T P.501]. The level of the activation level is adjusted to  $-16$  dBm<sub>0</sub>, 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 diffuse-field equalized (see clause 7.1.6.1), according to [ITU-T P.581]. The equalized output signal of the right ear is used for the measurement. For headset hands-free terminals, the sound pressure is measured at the DRP and diffuse-field equalized.
- 3) The idle channel noise is measured at the DRP in the frequency range between 50 Hz and 10 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 one 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.
- 4) The idle channel noise is A-weighted.
- 5) Spectral peaks are measured in the frequency domain. The frequency spectrum of the A-weighted idle channel noise is measured by a spectral analysis having a noise bandwidth of 8.79 Hz (determined using FFT 8 k samples/48 kHz sampling rate with Hanning window or equivalent). The idle channel noise spectrum is stated in dB(A). A smoothed average idle channel noise spectrum is calculated by a moving average (arithmetic mean) 1/3rd octave wide across the idle noise channel spectrum stated in dB (linear average in dB of all FFT bins in the range from  $2^{(-1/6)}f$  to  $2^{(+1/6)}f$ ). Peaks in the idle channel noise spectrum are compared against a smoothed average idle channel noise spectrum.
- 6) In case of headset measurements the tests are repeated five times, in conformance with [ITU-T P.380]. The results are averaged (averaged value in dB, for each frequency).

## 11.8 Out-of-band signals

For the measurements, it is desirable to have a background noise level of less than  $-74$  dBPa(A), a background noise level of  $-64$  dBPa(A) shall not be exceeded.

### 11.8.1 Discrimination against out-of-band signals in send direction

#### 11.8.1.1 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 which 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 8 kHz and 10 kHz with a level of  $-4.7$  dBPa at the MRP is used. The total level – measured in a frequency range from 200 Hz to 7 kHz – is measured at the electrical reference point (POI) and shall be at least 35 dB below the level of the reference signal. The reference level is determined using the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501], band-limited in the frequency range between 100 Hz and 8 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.

#### **11.8.1.2 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 four CSSs, according to [ITU-T P.501]. The level of 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 8 kHz to 10 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 (the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501], band-limited from 100 Hz to 8 kHz,  $-4.7$  dBPa at the MRP) is determined at the electrical reference point (POI) as well.

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

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

### **11.8.2 Spurious out-of-band signal in receive direction**

#### **11.8.2.1 Requirements**

The test signal used is the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501], band-limited in a frequency range between 100 Hz and 7 kHz with a level of  $-12$  dBm0 in receive direction. The level of the out-of-band signal is measured in a frequency range between 8.6 and 16 kHz at the hands-free loudspeaker. The out-of-band signal with the in-band signal present shall not be more than 10 dB above the signal level (reference signal level) in the same frequency range with no in-band signal present.

For the measurements, it is desirable to have a background noise level of less than  $-74$  dBPa(A), a background noise level of  $-64$  dBPa(A) shall not be exceeded.

#### **11.8.2.2 Test**

- 1) The test signal shall be the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501]. 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, currently no test can be recommended due to the limitations of the specification for HATS and ear simulators.

- 3) The level of the out-of-band signal is determined between 8.6 and 16 kHz. The reference level is determined by measuring the signal level without the in-band signal present in the same frequency range.

NOTE 1 – This measurement method does not apply to systems including artificial bandwidth extension.

NOTE 2 – In case of fluctuating noise levels, multiple reference noise measurements should be made and the worst-case result should be used as a reference.

## **11.9 Distortion in send**

The distortion in send, up to 8 kHz, is measured from the MRP to the electrical reference point (input of the system simulator, POI).

NOTE – It is recognized that for some systems, including AGC or companding techniques, the distortion test does not lead to valid results.

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

### **11.9.1 Requirements**

The harmonic distortion in send shall not be higher than 3%.

### **11.9.2 Test**

- 1) The test signal is a sinusoidal signal with a frequency of 300 Hz, 500 Hz, 1 kHz and 2 kHz. The test signal level is  $-4.7$  dBPa. In order to guarantee a reliable activation of the hands-free terminal, a sequence of four 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.

## **11.10 Distortion in receive**

The distortion in receive is measured from the POI to the artificial ear up to 15 kHz.

NOTE 1 – It is recognized that for some systems including AGC, companding techniques or bandwidth extension, the distortion test does not lead to valid results.

NOTE 2 – The current [ITU-T P.58] specifies the transmission characteristics of HATS for sound pick-up only up to 10 kHz. Therefore, the distortion test is performed using a measurement microphone.

For the measurements, it is desirable to have a background noise level of less than  $-74$  dBPa(A), a background noise level of  $-64$  dBPa(A) shall not be exceeded.

### **11.10.1 Requirements**

The distortion in receive is measured from a linear access point (access point at the hands-free system where no non-linear and time-variant signal processing except the speech coder is present) to the measurement microphone positioned at the point where, for all other measurements, the centre of the HATS is located.



The harmonic distortion shall be less than 3% when producing a sound pressure level needed to achieve an  $S/N \geq 6$  dB (see clause 7, preparation measurements) 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.

### 11.10.2 Test

- 1) The test signal is a sinusoidal signal with a frequency of 300 Hz, 500 Hz, 1 kHz, 2 kHz and 3 kHz. The test signal level is the level measured at the linear access point when inserting a test signal with  $-16$  dBm<sub>0</sub> at the POI. In order to guarantee a reliable activation of the hands-free terminal, a sequence of four 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  $-16$  dBm<sub>0</sub>, 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.

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

## 11.11 Echo performance

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, weighted terminal coupling loss (TCL<sub>w</sub>), 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), a background noise level of  $-64$  dBPa(A) shall not be exceeded.

### 11.11.1 Terminal coupling loss

#### 11.11.1.1 Requirements

Terminal coupling loss (TCL) in quiet environments is measured as unweighted echo loss and shall be at least 46 dB for nominal setting of the volume control. For maximum setting of the volume control, TCL shall be higher than 46 dB. The implemented echo control mechanism should provide a sufficient echo loss for all typical environments and typical impulse responses.

NOTE 1 – A TCL of  $\geq 50$  dB is recommended as a performance objective when measured under free-field conditions at the nominal setting of the volume control. Depending on the idle channel noise in the sending direction, it may not always be possible to measure an echo loss  $\geq 50$  dB.

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

NOTE 2 – There may be implementations where echo problems may be observed although the TCL 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.

### 11.11.1.2 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) The test signal is the compressed real speech signal described in clause 7.3.3 of [ITU-T P.501]. The signal level shall be  $-10$  dBm0.
- 3) The first 17.0 seconds of the test signal (six sentences) are discarded from the analysis to allow for convergence of the acoustic echo canceller. The analysis is performed over the remaining length of the test sequence (last six sentences).
- 4) TCL is calculated as unweighted echo loss from 100 Hz to 8 kHz. For the calculation, the averaged test signal level at each frequency band is referred to the averaged measured echo signal level 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. The echo loss is calculated by the equations

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

and

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

where:

$A_0$  is the output/input power ratio at frequency  $f_0 = 100$  Hz;

$A_1$  the ratio at frequency  $f_i$ ; and

$A_N$  the ratio at frequency  $f_N = 8000$  Hz.

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

## 11.11.2 Echo level vs. time

### 11.11.2.1 Requirements

This test aims to verify that the system will maintain sufficient echo attenuation during single talk. When measuring using the CS-signal the measured echo attenuation during single talk shall not decrease by more than 6 dB from the maximum measured during the test. When measuring using the British-English single talk sequence the echo level variation shall be less than 6 dB.

NOTE 1 – The echo path is kept constant during this test, and the test should begin five 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.

### 11.11.2.2 Test

- 1) The test arrangement is according to clause 7.1.
- 2) The test signal consists of 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 seconds which represents eight 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 the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501]. One sequence of a male and one sequence of a female voice 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.

- 3) When using the CS-signal the measurement result is displayed as attenuation vs. time. The exact synchronization between input and output signal has to be guaranteed.
- 4) When using the speech signal the measurement is displayed as level versus time.

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).

### 11.11.3 Spectral echo attenuation

#### 11.11.3.1 Requirements

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

**Table 11-3 – Spectral echo attenuation mask**

Frequency (Hz)	Upper limit
100 Hz	$-41$ dB
1 300 Hz	$-41$ dB
3 450 Hz	$-46$ dB
5 200 Hz	$-46$ dB
7 500 Hz	$-37$ dB
8 000 Hz	$-37$ dB

NOTE 1 – All sensitivity values are expressed in dB.  
 NOTE 2 – The limit at intermediate frequencies lies on a straight line drawn between the given values on a log (frequency) – linear (dB) scale.

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

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

#### 11.11.3.2 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 seconds. 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.

## 11.11.4 Initial convergence without background noise

### 11.11.4.1 Requirements

The initial convergence (echo attenuation vs. time) during single talk immediately after activating the hands-free terminal with maximum volume control setting shall conform to the requirement in Figure 11-5.

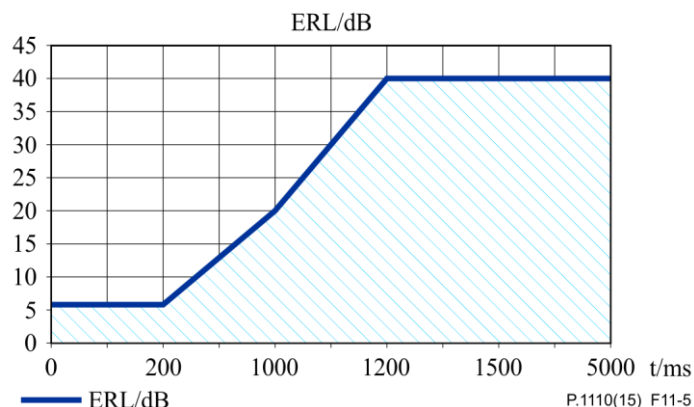


Figure 11-5 – Initial convergence, ERL versus time

### 11.11.4.2 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 five seconds. 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 vs. time, measured signal and reference signal have to be synchronized in time.

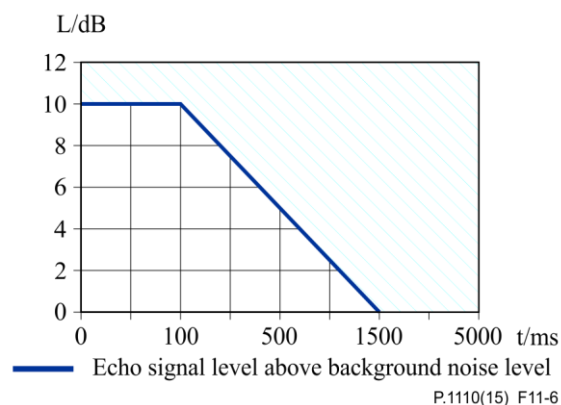
NOTE 1 – The analysis of the CSS is performed only on the active signal parts, the pauses between the bursts of the composite source signal are not analysed. The analysis time is reduced by the time constant of the level analysis due to the integration time of 35 ms.

NOTE 2 – The required performance should be achieved for different speech signals.

## 11.11.5 Initial convergence with background noise

### 11.11.5.1 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 shall conform to the requirement in Figure 11-6.



**Figure 11-6 – Initial convergence with background noise, requirement on echo signal level versus time**

### 11.11.5.2 Test

- 1) The test arrangement is described in clause 7.1.
- 2) The background noise defined by the manufacturer/test house is played back at least five seconds 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  $-16$  dBm<sub>0</sub>. The echo signal is analysed over a period of at least five seconds. The analysis integration time is 35 ms. In addition, the test is repeated with the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501]. The first male sentence and the first female sentence are used. The starting point of each signal is as defined by the start of the male sentence. The average test signal level is  $-16$  dBm<sub>0</sub>. The echo signal is analysed over a period of at least five seconds. The analysis integration time is 35 ms.
- 5) The measurement is displayed as echo level vs. 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 speech signals should be achieved for different starting points of the speech signal.

### 11.11.6 Echo performance with time variant echo path

#### 11.11.6.1 Requirements

This test aims to verify that the system will maintain sufficient echo attenuation during time variant echo path. The measured echo level measured with a time varying echo path shall not increase by more than 6 dB compared to the echo level observed under steady state conditions. The time-variant echo-path may be realized as follows:

- A rotating 30 cm × 40 cm reflecting surface (e.g., piece of cardboard, wood, or plastic) is positioned on the co-driver's 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 seat. The reflecting surface then pivots 90° such that the most forward edge of the reflecting surface

rotates out towards the co-driver's 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.

- Alternatively the time variant echo path is realized by opening and closing the door at the driver's side during the measurement. Care has to be taken to quietly open and close the door in order not to impair the measurement by noises produced when opening and closing the door or by warning signals produced by the car.
- As a third alternative, the test is conducted by positioning a person on the co-driver's seat, which is quietly moving one arm close to the microphone during the measurement.

The type of echo path variation chosen shall be reported.

#### **11.11.6.2 Test**

- 1) Before conducting the test, the echo canceller shall be fully converged.
- 2) The test arrangement is according to clause 7.1.
- 3) The measurement is started by introducing the time-variant echopath.
- 4) The test signal consists of periodically repeated composite source signal according to [ITU-T P.501], with an average level of –5 dBm<sub>0</sub> as well as an average level of –25 dBm<sub>0</sub>. The echo signal is analysed during a period of at least 2.8 seconds which represents eight 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.
- 5) The measurement result is displayed as attenuation vs. 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).

### **11.11.7 Echo performance with time variant echo path and speech**

#### **11.11.7.1 Requirements**

This test aims to verify that the system will maintain sufficient echo attenuation during time variant echo path when applying speech. The measured echo level measured with a time varying echo path shall not increase by more than 6 dB compared to the echo level observed under steady-state conditions. The time-variant echopath is realized as described in 11.11.6.1.

#### **11.11.7.2 Test**

- 1) Before conducting the test, the echo canceller shall be fully converged.
- 2) The test arrangement is according to clause 7.1. The test signal used is the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501]. The first male sentence and the first female sentence are used. The average test signal level is –16 dBm<sub>0</sub>. The echo signal is analysed during the complete test signal. The integration time for the level analysis shall be 35 ms. The measurement result is calculated as level vs. time. The echo level is determined under steady state conditions and stored as a reference.
- 3) Now, a second measurement is started by introducing the time-variant echopath.
- 4) The test is repeated with the reflecting surface moving and the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501]. Again, the first male sentence and the first female sentence are used. The average test signal level is –16 dBm<sub>0</sub>. The echo signal is

analysed during the complete test signal. The integration time for the level analysis shall be 35 ms. The measurement result is calculated as level vs. time.

- 5) The difference of the echo level between the reference and the measured echo loss with the time-variant echopath is determined.
- 6) The measurement result is displayed as attenuation vs. time. The exact synchronization between the two measured signals has to be guaranteed.

### 11.11.8 Switching characteristics

#### 11.11.8.1 Activation in send direction

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

The activation level described in the following is always referred to the test signal level at the MRP.

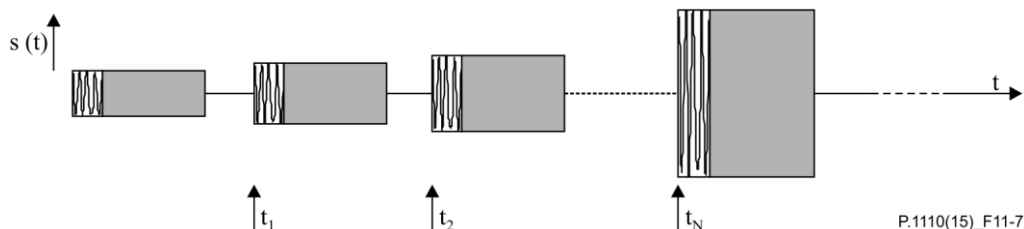
##### 11.11.8.1.1 Requirements

The minimum activation level  $L_{S,min}$  shall be  $\leq -20$  dBPa.

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

##### 11.11.8.1.2 Test

The structure of the test signal is shown in Figure 11-7. The test signal consists of CSS components according to [ITU-T P.501] with increasing level for each CSS burst.



**Figure 11-7 – 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 send direction**

	CSS duration/ Pause duration	Level of the first CS signal (active signal part at the MRP)	Level difference between two periods of the test signal
CSS to determine switching characteristic in Send Direction	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 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 vs. 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 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 corresponding CSS-burst.

### 11.11.8.2 Activation in receive direction

The activation in receive direction is mainly determined by the built-up time  $T_{r,R,min}$  and the minimum activation level receive direction ( $L_{R,min}$ ). 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 in the following 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 receive direction, a measurement microphone is used for this test and positioned close to the loudspeaker of the hands-free terminal.

#### 11.11.8.2.1 Requirements

The minimum activation level  $L_{R,min}$  shall be  $\leq -35.7$  dBm0 (measured during the active signal part).

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

#### 11.11.8.2.2 Test

The signal construction is shown in Figure 11-7. The test signal settings are as follows:

**Table 11-5 – Settings of the CSS in receive direction**

	<b>CSS duration/ Pause duration</b>	<b>Level of the first CS signal (active signal part at the POI)</b>	<b>Level difference between two periods of the test signals</b>
CSS to determine switching characteristics in Receive 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 vs. time. The integration time of the level analysis used shall 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 and the average level of the word must be adapted to the CS-Signal level of the corresponding CSS-burst.

### 11.11.8.3 Attenuation range in send direction

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

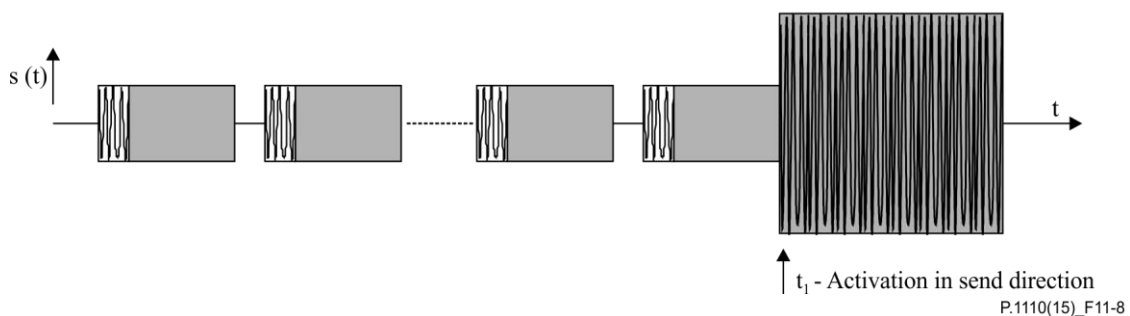
#### 11.11.8.3.1 Requirements

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

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

#### 11.11.8.3.2 Test

The structure of the test signals is shown in Figure 11-8. It consists of periodically repeated composite source signal bursts used for activating the receive direction and the voiced sound used to measure the send direction.



**Figure 11-8 – 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 shall be constant during the measurement.

#### 11.11.8.3.3 Test

- 1) The test arrangement is according to clause 7.1.
- 2) The test signal used is according to Figure 11-8, the receive direction is activated first. The measurement parameters are as follows:

**Table 11-6 – Signal levels for double talk tests in send and receive direction**

	Receive direction	Send direction (at the MRP)
Average signal level (including 101.38 ms pauses)	-16 dBm0	-
Active signal part for speakerphone HFT	-14.7 dBm0	0 dBPa
Active signal part for headset HFT	-14.7 dBm0	-3 dBPa

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

- 3) The level is determined as level vs. 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 send direction ( $t_1$  in Figure 11-8) until complete activation in send direction.

#### 11.11.8.4 Attenuation range in receive direction

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

##### 11.11.8.4.1 Requirements

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

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

##### 11.11.8.4.2 Test

The structure of the test signal is shown in Figure 11-8. Again, CSS bursts are used for activating the opposite direction (send direction) and the voiced sound is used to measure the receive direction. The test signals are synchronized in time at the acoustical interface. The delay of the test arrangement shall be constant during the measurement.

- 1) The test arrangement is according to clause 7.1.
- 2) The test signal shown in Figure 11-8 is used, the send direction is activated first.

The measurement parameters are as follows:

**Table 11-7 – Signal levels for double talk tests in send and receive direction**

	Receive direction	Send direction (at the MRP)
Average level speakerphone HFT (including 101.38 ms pauses)	–	–1.7 dBPa
Average level headset HFT (including 101.38 ms pauses)	–	–4.7 dBPa
Active signal part for speakerphone HFT	–14.7 dBm0	0 dBPa
Active signal part for headset HFT	–14.7 dBm0	–3 dBPa

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

- 3) The level is determined as level vs. 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 receive direction ( $t_1$  in Figure 11-8) and the complete activation in receive direction.

## 11.12 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 a good echo attenuation which is realized by a high attenuation range in this case.

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

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

### 11.12.1 Attenuation range in send direction during double talk: $A_{H,S,dt}$

#### 11.12.1.1 Requirements

Based on the level variation in send 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
	<i>Full duplex capability</i>	<i>Partial duplex capability</i>			<i>No duplex capability</i>
$A_{H,S,dt}$ [dB]	$\leq 3$	$\leq 6$	$\leq 9$	$\leq 12$	$> 12$

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

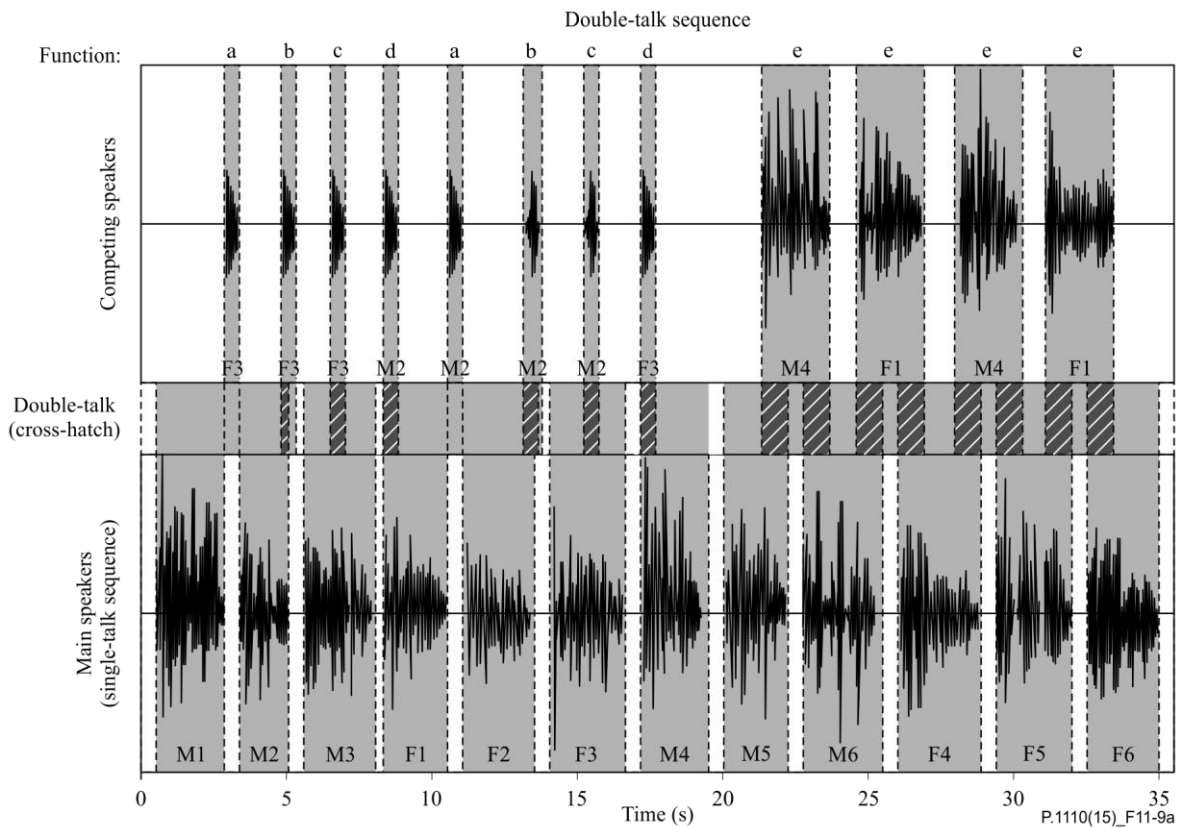
The requirements apply for nominal signal levels in send and receive directions as well as for the level combinations +6 dB (re. nominal level) in send/−6 dB (re. nominal level) in receive and +6 dB (re. nominal level) in receive/−6 dB (re. nominal level) in send. Furthermore, the test is conducted with 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.

#### 11.12.1.2 Test

The test signal to determine the attenuation range during double talk is shown in Figure 11-9a. The test signal to determine the attenuation range during double talk is the double talk speech sequence as defined in clause 7.3.5 of [ITU-T P.501]. The competing speaker is always inserted as the double talk sequence sdt(t) in send and is used for analysis.



**Figure 11-9a – Double talk test sequence with overlapping speech sequences in send and receive directions**

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

The settings for the test signals are as follows:

**Table 11-9 – Timing of the double talk sequences**

	Receive direction	Send direction
Average signal level for speakerphone HFT	-16 dBm0	-1.7 dBPa
Average signal level for headset HFT	-16 dBm0	-4.7 dBPa

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

- 1) The test arrangement is according to clause 7.1. Before the actual test, a training sequence for the echo canceller consisting of the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501] with a level of -16 dBm0 is applied to the electrical reference point.
- 2) When determining the attenuation range in send direction, the signal measured at the electrical reference point is referred to the test signal inserted.
- 3) The attenuation range during double talk is determined as described in [ITU-T P.502]. The double talk performance is analysed for the sequence of words and the sequence of sentences produced by the competing speaker. The requirement has to be met for the sequence of words and the sequence of sentences produced by the competing speaker.
- 4) The test is repeated for all level combinations as defined in the requirements.

### 11.12.2 Attenuation range in receive direction during double talk: $A_{H,R,dt}$

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

#### 11.12.2.1 Requirements

Based on the level variation in receive 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**

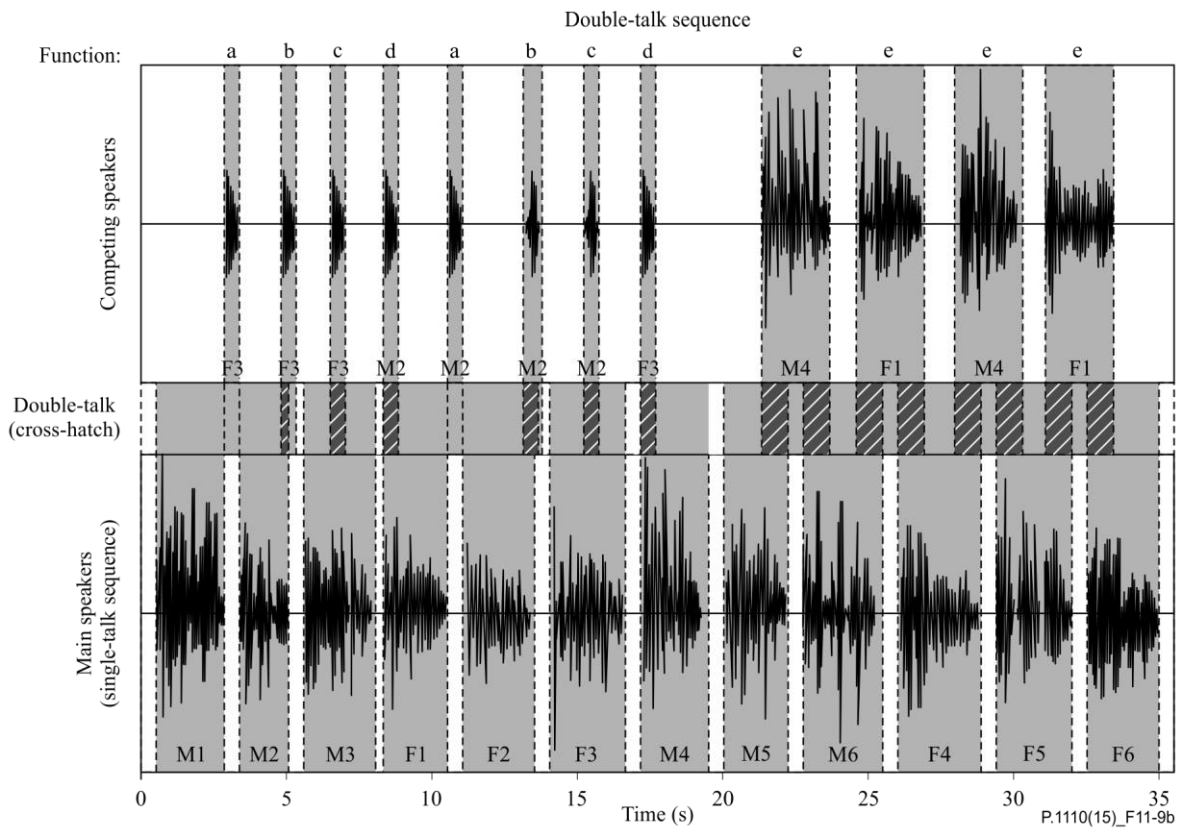
Category (according to ITU-T P.340)	1	2a	2b	2c	3
	<i>Full duplex capability</i>	<i>Partial duplex capability</i>			<i>No duplex capability</i>
$A_{H,R,dt}$ [dB]	$\leq 3$	$\leq 5$	$\leq 8$	$\leq 10$	$> 10$

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

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.

#### 11.12.2.2 Test

The test signal to determine the attenuation range during double talk is shown in Figure 11-9b. The test signal to determine the attenuation range during double talk is the double talk speech sequence as defined in clause 7.3.5 of [ITU-T P.501]. The competing speaker is always inserted as the double talk sequence  $sdt(t)$  in receive and is used for analysis. The test signals are synchronized in time at the acoustical interface. The delay of the test arrangement shall be constant during the measurement.



**Figure 11-9b – Double talk test sequence with overlapping speech sequences in receive and send directions**

The settings for the test signals are as follows:

**Table 11-11 – Timing of the double talk sequences**

	Receive direction	Send direction
Average signal level for speakerphone HFT	-16 dBm0	-1.7 dBPa
Average signal level for headset HFT	-16 dBm0	-4.7 dBPa

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

- 1) The test arrangement is according to clause 7.1.
- 2) When determining the attenuation range in receive direction, the signal measured at the loudspeaker of the hands-free terminal is referred to the test signal inserted. 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 attenuation range during double talk is determined as described [ITU-T P.502]. The double talk performance is analysed for the sequence of words and the sequence of sentences produced by the competing speaker. The requirement has to be met for the sequence of words and the sequence of sentences produced by the competing speaker.
- 4) The test is repeated for all level combinations as defined in the requirements.

### 11.12.3 Detection of echo components during double talk

#### 11.12.3.1 Requirements

The echo attenuation during double talk is based on the parameter talker echo loudness rating ( $TEL_{R_{dt}}$ ). 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 below 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
	<i>Full duplex capability</i>	<i>Partial duplex capability</i>			<i>No duplex capability</i>
<b>Echo loss [db]</b>	$\geq 27$	$\geq 23$	$\geq 17$	$\geq 11$	$< 11$

#### 11.12.3.2 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 signal is described in [ITU-T P.501]. The signal settings used are shown in Figure 11-10 below. A detailed description can be found in [ITU-T P.501].

The signals are fed simultaneously in send and receive directions. The level in send direction is  $-4.7$  dBPa at the MRP (nominal level) for headset hands-free terminals,  $-25.7$  dBPa at the HATSHFRP for speakerphone hands-free terminals, the level in receive direction is  $-16$  dBm0 at the electrical reference point (nominal level).

Send direction		Receive direction	
$f_0^{(1)}$ (Hz)	$\pm \Delta f^{(1)}$ (Hz)	$f_0^{(2)}$ (Hz)	$\pm \Delta f^{(2)}$ (Hz)
125	$\pm 2.5$	180	$\pm 2.5$
250	$\pm 5$	270	$\pm 5$
500	$\pm 10$	540	$\pm 10$
750	$\pm 15$	810	$\pm 15$
1 000	$\pm 20$	1 080	$\pm 20$
1 250	$\pm 25$	1 350	$\pm 25$
1 500	$\pm 30$	1 620	$\pm 30$
1 750	$\pm 35$	1 890	$\pm 35$
2 000	$\pm 40$	2 160	$\pm 35$
2 250	$\pm 40$	2 400	$\pm 35$
2 500	$\pm 40$	2 650	$\pm 35$
2 750	$\pm 40$	2 900	$\pm 35$
3 000	$\pm 40$	3 150	$\pm 35$
3 250	$\pm 40$	3 400	$\pm 35$

Send direction		Receive direction	
$f_0^{(1)}$ (Hz)	$\pm\Delta f^{(1)}$ (Hz)	$f_0^{(2)}$ (Hz)	$\pm\Delta f^{(2)}$ (Hz)
3 500	$\pm 40$	3 650	$\pm 35$
3 750	$\pm 40$	3 900	$\pm 35$
4 000	$\pm 40$	4 150	$\pm 35$
4 250	$\pm 40$	4 400	$\pm 35$
4 500	$\pm 40$	4 650	$\pm 35$
4 750	$\pm 40$	4 900	$\pm 35$
5 000	$\pm 40$	5 150	$\pm 35$
5 250	$\pm 40$	5 400	$\pm 35$
5 500	$\pm 40$	5 650	$\pm 35$
5 750	$\pm 40$	5 900	$\pm 35$
6 000	$\pm 40$	6 150	$\pm 35$
6 250	$\pm 40$	6 400	$\pm 35$
6 500	$\pm 40$	6 650	$\pm 35$
6 750	$\pm 40$	6 900	$\pm 35$
7 000	$\pm 40$		

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

**Figure 11-10 – 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 (send direction). The measured signal consists of the double talk signal which was fed in by the artificial mouth and the echo signal. The echo signal is filtered by comb filter using mid-frequencies and bandwidth according to the signal components of the signal in receive direction (see [ITU-T P.501]). The filter will suppress frequency components of the double talk signal.
- 4) In each frequency band which is used in receive direction, the echo attenuation can be measured separately. The requirement for category 1 is fulfilled if, in any frequency band, the echo signal is either below the signal noise or below the required limit. If echo components are detectable, the classification is based on Table 11-12 above. The echo attenuation is to be achieved for each individual frequency band from 200 Hz to 6950 Hz according to the different categories.

NOTE – Some HFTs may fail this requirement due to perceptually-based spectral filters which allow low levels of the double-talk signal to leak into the analysis window used for measuring echo. If it can be demonstrated that failures are not caused by echo, then the DUT is considered compliant with this requirement.

#### 11.12.4 Sent speech attenuation during double talk

##### 11.12.4.1 Requirements

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

Based on the level variation in send direction during double talk  $A_{H,S,dt}$ , the behaviour of hands-free terminals can be classified according to Table 11-13.



**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
	<i>Full duplex capability</i>	<i>Partial duplex capability</i>			<i>No duplex capability</i>
$A_{H,S,dt}$ [dB]	$\leq 3$	$\leq 6$	$\leq 9$	$\leq 12$	$> 12$

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

The requirements apply for nominal signal levels in send and receive directions as well as for the level combinations +6 dB (re. nominal level) in send/−6 dB (re. nominal level) in receive and +6 dB (re. nominal level) in receive/−6 dB (re. nominal level) in send. 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 11.12.1 in order to guarantee that no switching device with short reaction time is classified as a duplex or partially duplex system.

#### 11.12.4.2 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 signal used is described in [ITU-T P.501]. The signal settings are shown in Figure 11-10 below. A detailed description can be found in [ITU-T P.501].

The signals are fed simultaneously in send and receive directions. The level in send direction is −4.7 dBPa at the MRP (nominal level) for headset hands-free terminals, −25,7 dBPa at the HATSHFRP for speakerphone hands-free terminals, the level in receive direction is −16 dBm0 at the electrical reference point (nominal level).

- 3) The test signal is measured at the electrical reference point (send direction). The measured signal consists of the double talk signal, which was fed in by the artificial mouth, and the echo signal. The double talk signal (send signal) is filtered by comb filter using mid-frequencies and bandwidth according to the signal components of the signal in send 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 send 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 send 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 6900 Hz according to the different categories.
- 5) The test is repeated for all level combinations as defined in the requirements.

### 11.13 Background noise transmission

#### 11.13.1 SNR improvement provided by the HFT algorithm

For further study.

## 11.13.2 Background noise transmission after call set-up

### 11.13.2.1 Requirements

The analysis based on the Relative Approach [b-Sottek] (see Annex B) shall not indicate remarkable characteristics exceeding 6 cp/cPa. The first transmitted signal peak in send direction shall not cause higher excitation than 15 cp/cPa between 200 Hz and 7 kHz.

### 11.13.2.2 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 shall 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 shall be taken not to produce any disturbances or unwanted noise by touching the terminals housing while answering the incoming call.
- 4) The transmitted signal in send direction is recorded at the POI starting at least one second before the call is answered and for at least seven seconds after the call is established. The analysis range is chosen to eight seconds including an initial pause of one second before the call was established.
- 5) The recorded signal is analysed using the Relative Approach [b-Sottek].

## 11.13.3 Speech quality in the presence of background noise

### 11.13.3.1 Requirements

Speech quality for wideband systems can be tested based on [b-ETSI EG 202 396-3]. The test method is applicable for wideband transmission systems. The test method described leads to three MOS-LQO quality numbers:

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

According to specifications of manufacturer/test lab, a realistic background noise is played back. For the background noises chosen, the following requirements apply:

Background noises in quiet, low fan (see Annex D):

- N-MOS-LQOw  $\geq 3.9$
- S-MOS-LQOw  $\geq 4.6$
- G-MOS-LQOw  $\geq 3.9$

Background noises for driving speed  $\leq 80$  km/h:

- N-MOS-LQOw  $\geq 3.0$
- S-MOS-LQOw  $\geq 3.0$
- G-MOS-LQOw  $\geq 3.0$

For this test, the speech level is adjusted at the MRP to take into account the Lombard effect. The level adjustment is calculated according to clause 7.1.3.

Background noises for driving speed  $\leq 130$  km/h:

- N-MOS-LQOw  $\geq 2.6$
- S-MOS-LQOw  $\geq 3.0$
- G-MOS-LQOw  $\geq 2.9$

For this test, the speech level is adjusted at the MRP to take into account the Lombard effect. The level adjustment is calculated according to clause 7.1.3.

NOTE – It is recommended to test the terminal performance with different types of background, e.g., open window, different types of road surfaces and other relevant conditions. Especially, time variant conditions should be taken into account.

#### **11.13.3.2 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 background noise shall be applied for at least five second in order to adapt noise reduction algorithms in advance the test.
- 3) The near end speech signal consists of eight sentences of speech (two male and two female talkers, two sentences each). Appropriate speech samples can be found in [ITU-T P.501]. The preferred language is French since the objective method was validated with the French language. Three signals are required for the tests:
  - The clean speech signal is used as the undisturbed reference (see [b-ETSI EG 202 396-3]).
  - The speech plus undisturbed background noise signal is recorded at the terminal's microphone position using an omnidirectional measurement microphone with a linear frequency response between 50 Hz and 12 kHz.
  - The send signal is recorded at the electrical reference point.

N-MOS-LQOw, S-MOS-LQOw and G-MOS-LQOw are calculated as described in [b-ETSI EG 202 396-3].

#### **11.13.4 Quality of background noise transmission (with far-end speech)**

##### **11.13.4.1 Requirements**

This test is conducted by applying the composite source signal in receive direction. During and after the end of composite source signal bursts (representing the end of far-end speech simulation), the signal level in send direction shall not vary more than 10 dB (during transition to transmission of background noise without far-end speech).

##### **11.13.4.2 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.
- 3) First, the measurement is conducted without inserting the signal at the far end. At least 10 seconds of noise are recorded. The background signal level vs. time is calculated using a time constant of 35 ms. This is the reference signal.
- 4) Next, the same measurement is conducted but with inserting the CS-Signal at the far end. The exact 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 shall be applied for at least five seconds in order to allow adaptation of the noise reduction algorithms. After at least five seconds, a composite source signal according to [ITU-T P.501] is applied in receive direction with a duration of  $\geq 2$  CSS periods. The test signal level is  $-16$  dBm<sub>0</sub> at the electrical reference point.
- 5) The send signal is recorded at the electrical reference point. The test signal level versus time is calculated using a time constant of 35 ms.
- 6) The level variation in send direction is determined during the time interval when the CS-Signal is applied and after it stops. The level difference is determined from the difference

of the recorded signal levels vs. time between the reference signal and the signal measured with the far-end signal.

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

#### 11.13.5.1 Requirements

This test is conducted by applying a simulated speech signal in send direction. During and after the end of the simulated speech signal (composite source signal bursts), the signal level in send direction shall not vary more than 10 dB.

#### 11.13.5.2 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. The background noise shall be applied for at least five seconds in order to allow adaptation of the noise reduction algorithms.
- 3) The near-end speech is simulated using the composite source signal according to [ITU-T P.501] with a duration of  $\geq 2$  CSS periods. The test signal level is  $-4.7$  dBPa at the MRP for headset hands-free terminals,  $-25.7$  dBPa for hands-free. For speakerphone hands-free terminals, the HATS-HFRP correction has to be applied.
- 4) The send signal is recorded at the electrical reference point. The test signal level versus time is calculated using a time constant of 35 ms.
- 5) First, the measurement is conducted without inserting the signal at the near end. The signal level is analysed vs. time. Next, the same measurement is conducted but with inserting the CS-Signal at the near end. The level variation is determined by the difference between the background noise signal level without inserting the CS-Signal and the maximum level of the noise signal during and after the CSS-bursts in send direction.

### 11.13.6 "Comfort noise" injection

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

#### 11.13.6.1 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 A-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
2 000	10	-10
2 000	6	-6
4 000	6	-6
8 000	6	-6

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

### 11.13.6.2 Test

- 1) The test arrangement is according to clause 7.1. Background noise is played back.
- 2) The test signal is applied in receive direction consisting of an initial pause of 10 seconds and a periodical repetition of the composite source signal in receive direction (duration  $\geq 10$  seconds) with nominal level to enable comfort noise injection.
- 3) The transmitted signal is recorded in send direction at the electrical reference point.
- 4) The power density spectrum measured in send direction during the initial pause of the test signal (8k FFT/48 kHz sampling rate or equivalent, averaged over  $\geq 5$  seconds) is referred to the power density spectrum determined during the period with the periodical repetition of the composite source signals in receive direction (8k FFT/48 kHz sampling rate or equivalent, averaged over  $\geq 5$  seconds). 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 send direction is determined during the initial pause of the test signal in receive direction and referred to the level of the transmitted signal in send direction determined during the application of the test signal in receive direction. Both levels are calculated using A-weighting.

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

NOTE – In general these tests are targeted to verify that no additional signal processing is active when sending the appropriate command for deactivating the signal processing. However, in addition it is useful to verify that the signal processing is active when NOT sending this command. When the signal processing is active this can typically be observed in different measurement results for the following parameters:

- Delay in SND and in RCV,
- Junction loudness rating (JLR) in SND and in RCV,
- Sensitivity frequency response in SND and in RCV,
- Noise reduction test in SND,
- Verification of disabled echo control.

### 12.1 Interface definition and calibration

The principle of the SRW testing is shown in Figure 12-1:

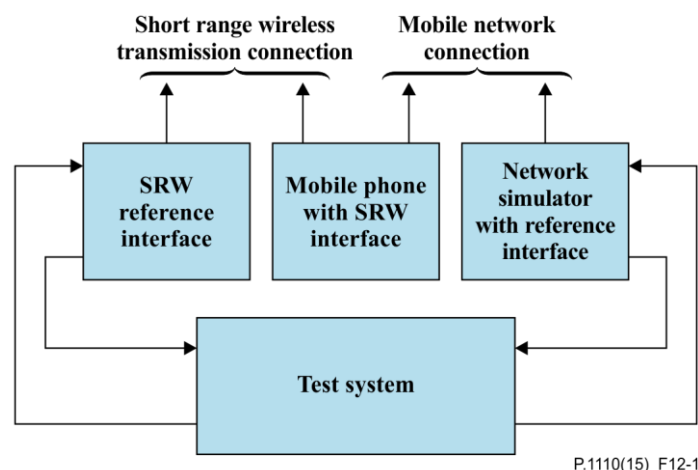
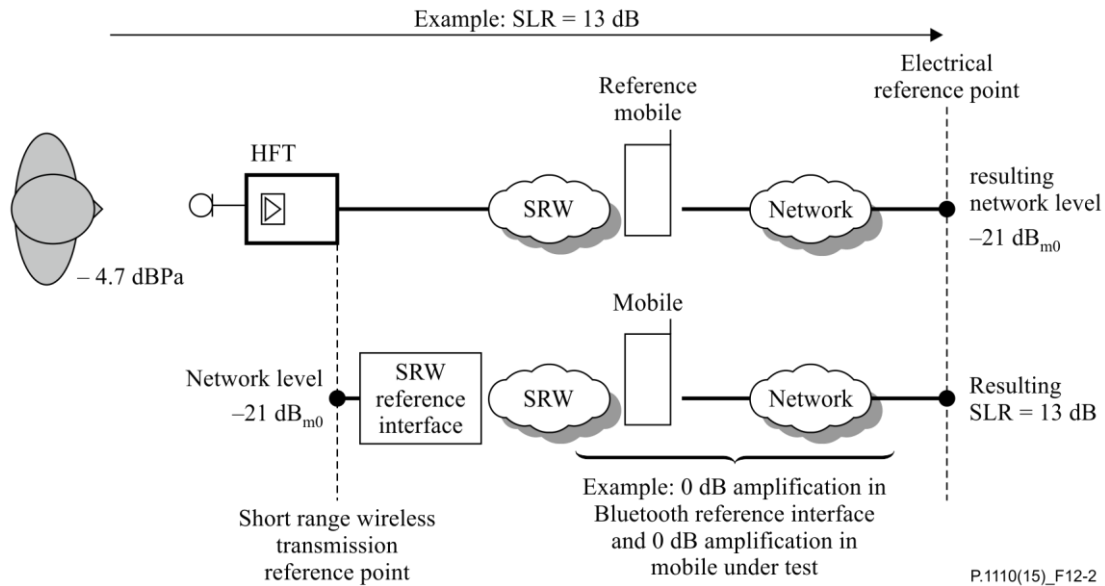


Figure 12-1 – SRW test set-up

The mobile phone under test is connected to the SRW reference interface at the SRW reference point (SRWR) and to the network (system) Simulator. The SRW reference interface, as well as the network simulator reference interface, is calibrated to the electrical inputs and outputs of the test system. While the 0 dB<sub>r</sub> point is clearly defined for the network reference interface, the calibration of the SRW 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 SRW reference interface**

The SRW reference interface is calibrated using a mobile phone with known characteristics which do not introduce any amplification or attenuation in the SRW link. No additional signal processing except the SRW 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 closest 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, a 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 SRW reference interface.

With an arbitrary mobile phone linked to the system simulator and at the same time connected to the SRW transmission reference interface via the SRW link, the level at the SRW 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 SRW 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 distance between MRP-HFRP – is applied at the SRW 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 SRW transmission link and the phone.

NOTE 2 – Evaluation boards from SRW transmission chipset vendors may be used for implementation of the SRW reference interface.

In case the SRW technology uses profiles, the profile(s) used during testing shall be selected as follows:

- If a profile pertaining specifically to car hands-free is supported by the DUT, this profile shall be used during testing.
- Otherwise, a profile pertaining to hands-free devices in general (in one SRW technology this is called the "Handsfree profile") shall be used during testing.

In case the relevant profile has provisions for sending commands to the DUT, the following commands shall be used during testing:

- Command A: Command from the SRW reference interface to the DUT which includes an indicator of the ability of the simulated hands-free device to exchange volume control status over the SRW link, here denoted remote audio volume control, RAVC.
- Command B: Command from the SRW reference interface to the DUT asking the DUT to enable or disable speech processing functions (e.g. a command that asks the DUT to disable its noise reduction and echo cancelling, NREC).

The behaviour of the mobile phone with SRW interface shall be verified as described in Table 12-0.

**Table 12-0 – Required behaviour of mobile phones which support exchange of commands A and/or B over the SRW link**

	Command A	Command B	Action
Case 1	RAVC indicator = 0	Processing indicator = 0	It shall be verified that <ul style="list-style-type: none"> <li>– Command B is acknowledged by the DUT</li> <li>– The speech processing is <i>disabled</i> in the DUT.</li> <li>– The JLR requirements are met for all DUT volume control settings.</li> </ul>
Case 2	RAVC indicator = 1	Processing indicator = 0	

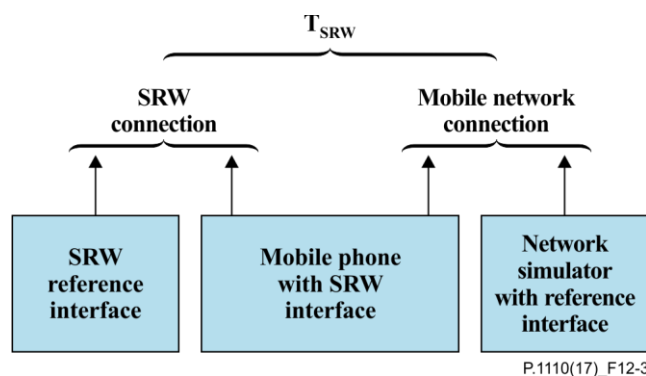
NOTE 1 – For the present table; RAVC indicator = 1 denotes that the HF device signals support for RAVC to the DUT. Processing indicator = 0 denotes a request from HF device that the DUT shall disable its speech processing.

NOTE 2 – Case 1 is considered to represent a typical behaviour of car hands-free devices.

### 12.1.1 SRW delay

#### 12.1.1.1 Requirements

The delay in send direction is measured from the SRWR to POI (reference speech codec of the system simulator, output) and in receive direction from the POI to the SRWR and it is corrected for the test equipment delays.



**Figure 12-3 – Delay in SRW connections**

## Definitions

- The mobile phone delay in the send (uplink) direction  $T_{SRW_s}$  is the delay between the first bit of a SRW speech frame at the SRW antenna to the last bit of the corresponding speech frame at the mobile network antenna.
- The mobile phone delay in the receive (downlink) direction  $T_{SRW_r}$  is the delay between the first bit of a speech frame at the mobile network antenna and the first bit of a SRW speech frame at the SRW antenna corresponding to that speech frame.

According to these definitions and for error-free radio conditions, the sum of send and receive delays  $T_{SRW_{rttd}} = T_{SRW_s} + T_{SRW_r}$  shall be less than 190 ms.

As a performance objective  $T_{SRW_{rttd}}$  should be less than 150 ms under these conditions.

NOTE 1 – The delay  $T_{SRW_{rttd}}$  should be minimized.

NOTE 2 – The system delay  $T_{system}$  depends on the transmission method used and the network simulator. The delay  $T_{system}$  must be known.

NOTE 3 – For 3GPP UMTS circuit-switched speech and 3GPP LTE MTSI-based speech, definitions, test methods, performance objectives and requirements are found in [b-3GPP TS 26.131] and [b-3GPP TS 26.132].

### 12.1.1.2 Test in send direction

The delay measured in the send direction is:

$$T_{SRW_s} + T_{system}$$

- 1) For the measurements, a 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 or equivalent. The test signal level is  $S_{SRW_{snd}}$ .

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

The test set-up 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.

### 12.1.1.3 Test receive direction

The delay measured in the receive direction is:

$$T_{SRW_r} + T_{system}$$

- 1) For the measurements, a 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 or equivalent. 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) shall be:

- $JLR_{SRW_{snd}} = 0 \pm 2$  dB
- $JLR_{SRW_{rcv}} = 0 \pm 2$  dB

No additional volume control shall be active.

NOTE – If it is possible to adjust the mobile phone volume control it should be verified that the measured  $JLR_{SRW_{rcv}}$  is not affected by different settings of this volume control.

### 12.2.2 Test SRW JLR in send direction

- 1) The test signal used for the measurements shall be the following two sentences (one male, one female voice) of clause 7.3.2 ([ITU T P.501]): "The last switch cannot be turned off" (sentence 1), "The hogs were fed chopped corn and garbage" (sentence 6). The test signal level is  $S_{SRW_{snd}}$ , 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 send sensitivity.

- 2) The test arrangement is according to clause 12.1. The SRW send sensitivity is calculated from each band of the 20 frequencies given in Table A.2 of [ITU-T P.79], bands 1-20.

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 JLR,  $JLR_{SRW_{snd}}$ , shall be calculated according to [ITU-T P.79], Formula A-23d, bands 1-20,  $M = 0.175$ , and the weighting factors  $W_J$  for JLR according to Table A.2 of [ITU-T P.79].
- 4) In case the SRW technology in the mobile phone supports command exchange according to table 12-0, the tests are repeated for case 1 and 2 with minimum and maximum volume control setting.

### 12.2.3 Test SRW JLR in receive direction

- 1) The test signal used for the measurements shall be the following two sentences (one male, one female voice) of clause 7.3.2 ([ITU T P.501]) "The last switch cannot be turned off" (sentence 1), "The hogs were fed chopped corn and garbage" (sentence 6). The test signal level 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 receive sensitivity is determined by the bands 1-20 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 JLR in receive  $JLR_{SRW_{rcv}}$  shall be calculated according to [ITU-T P.79], Formula A-23d, bands 1-20,  $M = 0.175$ , and the weighting factors  $W_J$  for JLR according to Table A.2 of [ITU-T P.79].
- 4) In case the SRW technology in the mobile phone supports command exchange according to Table 12-0, the tests are repeated for cases 1 and 2 with minimum and maximum volume control setting.

## 12.2.4 SRW linearity in send direction

### 12.2.4.1 Requirements

This test aims to detect any amplitude non-linearities, including AGC or companding. For acoustical signal level variation in the range of  $-30$  dB/0 dB from the nominal signal level  $S_{SRW_{snd}}$ , the measured  $JLR_{SRW_{snd}}$  shall not deviate more than  $\pm 2$  dB from the  $JLR_{SRW_{snd}}$  measured with the nominal signal level  $S_{SRW_{snd}}$ .

### 12.2.4.2 Test

- 1) The test signal used for the measurements shall be the following two sentences (one male, one female voice) of clause 7.3.2 ([ITU T P.501]) "The last switch cannot be turned off" (sentence 1), "The hogs were fed chopped corn and garbage" (sentence 6). The test signal levels are in the range of  $-30$  dBV to 0 dBV in steps of 10 dB relative to the nominal signal level  $S_{SRW_{snd}}$ , 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 send sensitivity.

- 2) The test arrangement is according to clause 12.1. The SRW send sensitivity is calculated from each band of the 20 frequencies given in Table A.2 of [ITU-T P.79], bands 1-20.

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 JLR,  $JLR_{SRW_{snd}}$  shall be calculated according to [ITU-T P.79], Formula A-23d, bands 1-20,  $M = 0.175$ , and the weighting factors in send direction according to Table A.2 of [ITU-T P.79].

## 12.2.5 SRW linearity in receive direction

### 12.2.5.1 Requirements

This test aims to detect any amplitude non-linearities, including AGC or companding. For network signal level variations in the range of  $-30$  dB to 0 dB relative to the nominal signal level, the measured  $JLR_{SRW_{rcv}}$  shall not deviate more than  $\pm 2$  dB from the  $JLR_{SRW_{rcv}}$  measured with nominal signal level.

### 12.2.5.2 Test

- 1) The test signal used for the measurements shall be the following two sentences (one male, one female voice) of clause 7.3.2 ([ITU T P.501]) "The last switch cannot be turned off" (sentence 1), "The hogs were fed chopped corn and garbage" (sentence 6). The test signal levels are in the range of  $-30$  dBV to 0 dBV in steps of 10 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 send sensitivity.

- 2) The test arrangement is according to clause 12.1. The SRW receive sensitivity is calculated from each band of the 20 frequencies given in Table A.2 of [ITU-T P.79], bands 1-20.

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 JLR,  $JLR_{SRW_{snd}}$  shall be calculated according to [ITU-T P.79], Formula A-23d, bands 1-20,  $M = 0.175$ , and the weighting factors in send direction according to Table A.2 of [ITU-T P.79].

## 12.3 SRW sensitivity frequency responses

### 12.3.1 SRW send sensitivity frequency response

#### 12.3.1.1 Requirements

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

The tolerance mask for the send 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 send sensitivity frequency response**

Frequency (Hz)	Upper limit	Lower limit
100	0	-3
2 000	0	-3
3 100	0	-4
6 300	0	-4
7 100	0	

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 the following two sentences (one male, one female voice) of clause 7.3.2 ([ITU T P.501]) "The last switch cannot be turned off" (sentence 1), "The hogs were fed chopped corn and garbage" (sentence 6). The test signal level is  $S_{SRW_{snd}}$ , 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 send sensitivity.

- 2) The test arrangement is according to clause 12.1. The SRW send sensitivity is determined in third octave intervals, as given by [IEC 61260] for frequencies of 100 Hz and 8 kHz, inclusive. In each third octave band, the level of the measured signal is referred to the level of the reference signal averaged over the complete test sequence length.
- 3) The sensitivity is determined in dBV/V.

### 12.3.2 SRW receive sensitivity frequency response

#### 12.3.2.1 Requirements

The receive 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 receive 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 receive sensitivity frequency response**

Frequency (Hz)	Upper limit	Lower limit
100	0	-3
2 000	0	-3
3 100	0	-4
6 300	0	-4
7 100	0	

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 the following two sentences (one male, one female voice) of clause 7.3.2 ([ITU T P.501]) "The last switch cannot be turned off" (sentence 1), "The hogs were fed chopped corn and garbage" (sentence 6). The test signal level is -16 dBm<sub>0</sub>, 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 receive sensitivity is determined in third octave intervals as given by [IEC 61260] for frequencies of 100 Hz and 8 kHz, inclusive. In each third octave band, the level of the measured signal is referred to the level of the reference signal, averaged over the complete test sequence length.
- 3) The sensitivity is determined in dBV/V.

## 12.4 SRW noise cancellation test in send direction

### 12.4.1 Requirements

This test aims to check whether any noise cancellation is active in the mobile phone. Acoustic echo control shall not be active in the mobile phone. The mobile phone noise cancellation is measured from SRWR to POI (reference speech codec of the system simulator, output).

The attenuation of the simulated background noise test signal shall not deviate more than  $\pm 4$  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 the left channel of the "Medium size car: Noise at constant driving conditions (100 km/h)" according to clause B.4.1 of [ITU-T P.501]. The first 20 seconds are used. The test signal is  $S_{SRW_{snd}}$ , the level is averaged over the test signal.
- 3) The level of the transmitted signal is measured at the electrical reference point. The A-weighted measured signal level is referred to the A-weighted test signal level and displayed vs. time. The levels are calculated from the time domain using an integration time of 250 ms.
- 4) The attenuation vs. time is determined over the entire test signal duration.

NOTE – If a non-linear or time variant behaviour of the phone is observed, the tests, as described in clause 11.13.5, can be applied to determine the behaviour of the phone in more detail. Instead of inserting the tests signals acoustically, they have to be inserted electrically.

## **12.5 SRW speech quality during single talk**

### **12.5.1 One-way listening speech quality in send direction**

#### **12.5.1.1 Requirement**

The nominal value for the listening speech quality measured from the SRW reference point to the electrical reference point (POI) shall be:

- MOS-LQOw  $\geq 3.8$

#### **12.5.1.2 Test**

The test uses [ITU-T P.863].

- 1) The test signals used are the English test sequences as specified in [ITU-T P.501] (two male speakers, two female speakers, two sentences each). The test signal level is  $S_{SRW_{snd}}$ , 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 per cent and 70 per cent.

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 12.1.  
The calculation is made using the signal recorded at the electrical interface.
- 3) The one-way speech quality is determined as MOS-LQOw.

#### **12.5.2 Listening speech quality stability in send direction**

This test is intended to detect speech quality instabilities due to unreliable SRW connection (speech frame loss). This test should be repeated in different representative positions of the phone within the vehicle, different SRW links active and different Wi-Fi links active. These tests are performed where the SRW link is installed in the target vehicle.

Listening quality stability during a call (if the position or transmission characteristics change during the call (or during several different calls) takes into account degradations generated on the signal by transmission link impairment and phone position.

In case of systems using SRW, this test aims to verify the integration of SRW radio network by evaluating the change in speech quality over time. This will help detect problems with RF coverage inside the car cabin and will verify the error concealment (packet loss, bit errors) caused by a weak RF link or interference with other radios.

The following steps provide guidance on how to proceed:

- 1) Check RF coverage from SRW unit to possible mobile positions (protocol analyser).
- 2) Identify weak or poor reception areas inside the vehicle (protocol analyser).  
Check bit error rate, rate of packet loss.
- 3) Use the speech quality measurement to rate the quality of error concealment of the weak areas identified in step 2 above and compare these with measurements from areas with good coverage.
- 4) Identify possible issues from interference with parallel SRW links or other networks. This will check how the SRW can handle interference and change to undisturbed channels. Speech quality measurements can be used to see the performance of handling these problems.

It is recommended to use a known phone as a reference.

### 12.5.2.1 Requirement

To provide a stable indication of listening speech quality, THRESHOLD1 is set to 0.1 and a linear weighting function is applied to express stability (ST-MOS) from a scale of 0 to 100. By definition, stability equals 100 when no variations occur, whereas stability, ST-MOS, equals 0 when MOS-LQOw variation is equal to, or more than 0.4.

- ST-MOS should be  $\geq 90$ .

This requirement applies only when nominal MOS-LQOw is fulfilled.

### 12.5.2.2 Test

Several measurements of MOS-LQOw score performed according to clause 12.5.1 are performed in series within the same call or for different calls, within the test arrangement defined below. Typically, for the same call, a measurement each 20 seconds is enough. The results are reported in terms of statistics.

The assessment of listening speech quality stability is performed in five steps:

- 1) To measure the MOS-LQOw periodically over the duration of one communication. N measurements provide N MOS-LQOw values.
- 2) For each MOS-LQOw (i) value i from 2 to N, MOS-LQOw\_GAP(i) is calculated as the absolute difference with the previous value MOS-LQOw (i-1):  
$$\text{MOS-LQOw\_GAP}(i) = |\text{MOS-LQOw}(i) - \text{MOS-LQOw}(i-1)|$$
- 3) In order to take into account the subjective perception and measurement accuracy, MOS-LQOw\_GAP (i) is set to 0 when the difference is equal or lower to THRESHOLD1:
  - if  $\text{MOS-LQOw\_GAP}(i) > (2 * \text{THRESHOLD1})$ , then  $\text{MOS-LQOw\_GAP}(i) = \text{MOS-LQOw\_GAP}(i)$
  - if  $\text{THRESHOLD1} < \text{MOS-LQOw\_GAP}(i) \leq 2 * \text{THRESHOLD1}$ , then  $\text{MOS-LQOw\_GAP}(i) = [\text{MOS-LQOw\_GAP}(i) * 2] - (2 * \text{THRESHOLD1})$
  - if  $\text{MOS-LQOw\_GAP}(i) \leq \text{THRESHOLD1}$ , then  $\text{MOS-LQOw\_GAP}(i) = 0$
- 4) The instability (INS\_MOS-LQOw) associated to the MOS-LQOw over the entire N measurements is defined by mean the value of MOS-LQOw\_GAP (i).  
$$\text{INS\_MOS-LQOw} = 1 / (N-1) \sum \text{MOS-LQOw\_GAP}(i) \text{ with } i = [2:N]$$
- 5) A linear weighting function is applied in order to express stability ST-MOS-LQOw on a 0 to 100 scale.

This formulation used to determine the listening quality stability (ST-MOS) as:

- $\text{ST-MOS} = 100 - (250 * \text{INS\_MOS})$
- and
- $\text{ST-MOS} = 0$  if  $[100 - (250 * \text{INS\_MOS})] < 0$

When ST-MOS is calculated within a single call, the call should be longer than three minutes (the recommended duration being between three and five minutes). The duration of each measurement depends on the length of the speech samples used for the test as described in clause 12.5.1.

(As an example, if a sample is 15 seconds and the analysis is done every 20 seconds, a minimum of 10 values will be measured.)

### **12.5.3 SRW one-way listening speech quality in receive direction**

#### **12.5.3.1 Requirement**

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

- MOS-LQO<sub>w</sub> ≥ 3.8

#### **12.5.3.2 Test**

The test uses [ITU-T P.863].

- 1) The test signals used are the English test sequences as specified in [ITU-T P.501] (two male speakers, two female speakers, two sentences each). The test signal is –16 dBm<sub>0</sub>, the test signal level is measured as "active speech level" according to [ITU-T P.56]. The speech activity should be between 30 per cent and 70 per cent.  
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 12.1. The signal measured at the SRWR is used for the calculation
- 3) The one-way speech quality is determined as MOS-LQO<sub>w</sub>.

### **12.5.4 Listening speech quality stability in receive direction**

This test is intended to detect speech quality instabilities due to unreliable SRW connections (speech frame loss). This test should be repeated in different representative positions of the phone within the vehicle and different combinations of SRW links active.

#### **12.5.4.1 Requirement**

To provide a stable indication of listening speech quality, THRESHOLD1 is set to 0.1 and a linear weighting function is applied to express stability (ST-MOS) within a scale of 0 to 100. By definition, stability equals 100 when no variations occur and stability ST-MOS equals 0 when MOS-LQO<sub>w</sub> variation is equal or more than 0.4.

- ST-MOS should be ≥ 90.

The requirement applies only when nominal MOS-LQO<sub>w</sub> is fulfilled.

#### **12.5.4.2 Test**

The test procedure is described in clause 12.5.2.2.

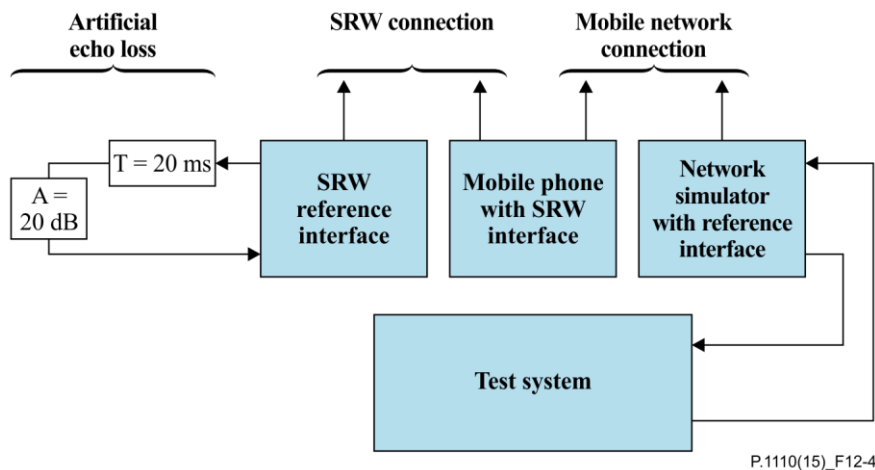
### **12.5.5 Verification of disabled echo control**

#### **12.5.5.1 Requirements**

Acoustic echo control shall not be active in the mobile phone. An artificial echo path consisting of an attenuation of 20 dB and a delay of 20 ms is inserted at the SRW reference interface. The echo loss measured shall be 20 dB ±2 dB when corrected by the measured JLR<sub>SRWsnd</sub> and JLR<sub>SRWrcv</sub>.

#### **12.5.5.2 Test**

- 1) For the test, an artificial echo path is inserted at the SRWR. The test set-up is shown in Figure 12-4:



**Figure 12-4 – Test set-up with artificial echo loss**

- 2) 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.
- 3) The test signal is the compressed real speech signal described in clause 7.3.3 of [ITU-T P.501]. The signal level shall be  $-16$  dBm<sub>0</sub>.
- 4) The first 17.0 seconds of the test signal (six sentences) are discarded from the analysis to allow for convergence of the acoustic echo canceller. The analysis is performed over the remaining length of the test sequence (last six sentences).
- 5) TCL is calculated as unweighted echo loss. 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.
- 6) The measured TCL is corrected by the measured  $JLR_{SRW_{snd}}$  and  $JLR_{SRW_{rcv}}$ .

### 13 Car-to-car communication

For further study, see also Annex E.

### 14 Guidance on subjective testing

Besides 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 the P.800-series of ITU-T Recommendations – 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.

For conducting the tests, the hands-free system under test has to be installed in the target car, which will be referenced here as near-end, as far-end either serves a landline phone (car-to-land test) or an observing car equipped also with the hands-free system under test (car-to-car test). 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 worse hands-free quality compared to landline connections.

The evaluation of the 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, a native language should be used. For the recordings, additional languages can be selected.



Since conversational tests are rather time-consuming, most of the 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 effective as possible, it is advantageous to use a tool providing test people on both sides of the telephone connection with the detailed test procedure and the possibility to easily do the rating.

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

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

For the evaluation, some ITU-T Recommendations can serve as a guideline, such as [ITU-T P.800], [ITU-T P.800.1] and others from the P.800-series of ITU-T Recommendations. The judgment 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 P.800-series of ITU-T Recommendations.

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

Tables 14-1 and 14-2 show possible test scenarios and rating categories.

**Table 14-1 – Overview of car-to-landline tests**

<b>Echo canceller</b>		<b>Rating</b>
Single talk	Three typical driving scenarios (Germany: e.g., 0 km/h, 100 km/h, 130 km/h, 160 km/h) Fan/defrost (off or "worst-case") Receive volume: nominal, maximum, varying Enclosure dislocation	Disturbance caused by echo Echo characteristics: <ul style="list-style-type: none"> <li>• Intensity</li> <li>• Duration</li> <li>• Frequency of occurrence</li> </ul> Background noise variation
Double talk	Three typical driving scenarios (Germany: e.g., 0 km/h, 100 km/h, 130 km/h, 160 km/h) Receive volume: nominal, maximum, varying Enclosure dislocation	Disturbance caused by echo Echo characteristics: <ul style="list-style-type: none"> <li>• Intensity</li> <li>• Duration</li> <li>• Frequency of occurrence</li> </ul> Speech level variation Speech intelligibility/listening effort

**Table 14-1 – Overview of car-to-landline tests**

<b>Echo canceller</b>		<b>Rating</b>
Conversation	0 km/h	Disturbance caused by echo Echo characteristics: <ul style="list-style-type: none"> <li>• Intensity</li> <li>• Duration</li> <li>• Frequency of occurrence</li> </ul> Speech level variation Speech intelligibility/listening effort
<b>Speech &amp; background noise quality (send direction)</b>		<b>Rating</b>
Stationary noise	Three typical driving scenarios (Germany: e.g., 0 km/h, 100 km/h, 130 km/h, 160 km/h) Driver's window closed/open 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 noises, wiper noise Pass by vehicles	Background noise quality Adaptation to background noise
<b>Speech quality (receive direction)</b>		<b>Rating</b>
Single talk far-end	Maximum receive volume New call	Speech level Speech sound quality Speech intelligibility/listening effort

**Table 14-2 – Overview of car-to-car tests**

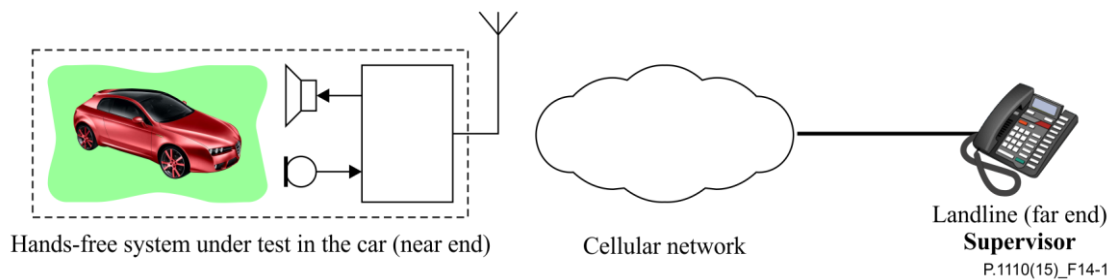
<b>Echo canceller</b>		<b>Rating</b>
Single talk	Three typical driving scenarios (Germany: e.g., 0 km/h, 100 km/h, 130 km/h, 160 km/h) Fan (off or "worst-case") Receive volume: nominal, maximum, varying Enclosure dislocation	Disturbance caused by echo Echo characteristics: <ul style="list-style-type: none"> <li>• Intensity</li> <li>• Duration</li> <li>• Frequency of occurrence</li> </ul> Background noise variation
Double talk	Three typical driving scenarios (Germany: e.g., 0 km/h, 100 km/h, 130 km/h, 160 km/h) Receive volume: nominal, maximum, varying Enclosure dislocation	Disturbance caused by echo Echo characteristics: <ul style="list-style-type: none"> <li>• Intensity</li> <li>• Duration</li> <li>• Frequency of occurrence</li> </ul> Speech level variation Speech intelligibility/listening effort

**Table 14-2 – Overview of car-to-car tests**

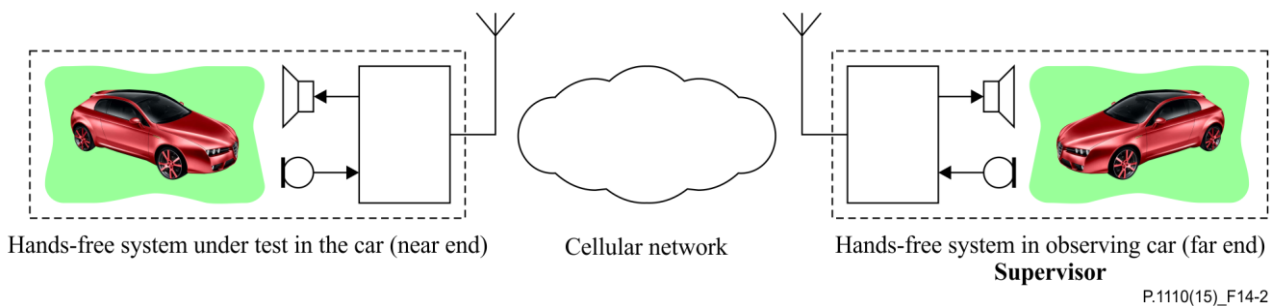
System stability (car-to-car)		Rating
Stability	Receive volume: nominal, maximum Test signal: speech or impulse-like excitation	Stability Echo characteristics: <ul style="list-style-type: none"> <li>• Intensity</li> <li>• Duration</li> <li>• Frequency of occurrence</li> </ul>
Speech and background noise quality (send direction)		Rating
Stationary noise	Three 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.		

**Test environment and equipment**

Figures 14-1 and 14-2 outline the test environment for the test scenarios of 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. He is guiding through the test procedure and doing most of the test evaluation.



**Figure 14-1 – Hands-free test car-to-landline**



**Figure 14-2 – 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 brand, model, model year, and hands-free system,

- appropriate recording equipment is needed at the landline station and the car for documentation purposes or 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 SRW connection between a phone and a hands-free system, the phone should be placed in an area with a good SRW and cellular coverage,
- some cell-phone networks have an influence on the hands-free performance (e.g., echo, AGC, noise reduction). 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 voices,
- for the duration of the tests, there should be no change in the driver, front-passenger, or landline speaker.

NOTE – The network quality of service (QoS) parameters can be monitored by specific systems during the tests (based on 3GPP).

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

- car brand, model, model year, tyres, type of road, engine type, interior trim (e.g., cloth or leather, sunroof),
- mobile telephone type and SW version,
- telematics microphone – description, location, orientation, distance to driver,
- network provider,
- HW and SW version of the hands-free DUT,
- HW and SW 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 for different countries depending, e.g., on the national speed limit for cars.

As an 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 limit, the maximum allowed speed might be used with additionally the climate control switched on to an appropriate level.

In addition to the driving scenarios, also some different settings of the fan/defrost/climate control/recirculation are useful background noise cases. As "worst-case" scenario, a setting of the fan might serve where the air stream directly flows to 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 tests is given in Annex C.

## **15 Corded headsets**

For further study.

### **15.1 Connector type**

### **15.2 Connector wiring and electrical specifications**

### **15.3 Headset receive characteristics**

### **15.4 Headset transmit characteristics**

### **15.5 Standard behaviour in the presence of corded headsets**

## **16 Wireless headsets**

For further study.

### **16.1 Wireless headset types**

### **16.2 Test methodology for verification of standard behaviour**

### **16.3 Standard behaviour in the presence of wireless headsets**

#### **16.3.1 Associate headset with vehicle**

#### **16.3.2 Enter the vehicle**

#### **16.3.3 Receive a call**

#### **16.3.4 Make a call**

**16.3.5 Terminate a call**

**16.3.6 Exit the vehicle**

**16.3.7 Listen to other audio source**

**16.3.8 Switch between other audio and telephony**

**16.3.9 Intra-vehicle communications**

## Annex A

### Speech quality measurements

(This annex forms an integral part of this Recommendation.)

In the following, a testing method for determining the speech quality in send and receive direction is described. The speech quality expressed in telecommunication objective speech quality assessment (TOSQA) mean opinion score (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 absolute category rating (ACR) quality scale, as it is defined in [ITU-T P.800].

Subjective tests carried out according to [ITU-T P.830], [ITU-T P.835], and [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.863] give results in terms of MOS-LQO.

#### 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 a more or less leaky condition 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 descriptor of the respective audio bandwidth, see the following provisional instructions:

- N for MOS scores obtained for narrowband (300-3400 Hz) speech relative to a narrowband high quality reference. This is applicable for instance to narrowband only subjective tests or to [ITU-T P.863] 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.863] scores.
- M for MOS scores obtained for narrowband 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.

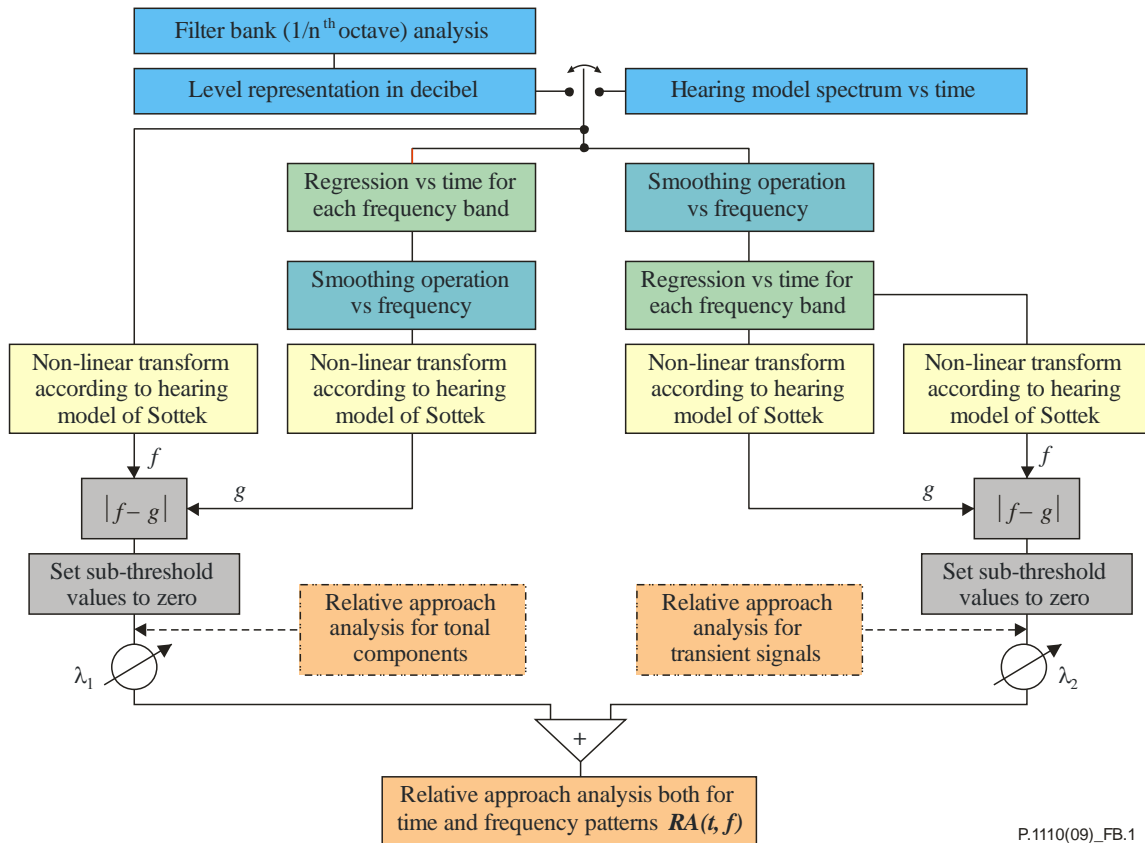
Further information can be found in [ITU-T P.800.1].

## Annex B

### Principles of Relative Approach

(This annex forms an integral part of this Recommendation.)

The **Relative Approach** [b-Sottek] 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.



**Figure B.1 – Block diagram of Relative Approach**

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 [b-Sottek].

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 pseudostationary 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.

Two different variants of the Relative Approach can be applied to the pre-processed signal. The first version 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 vs. frequency is performed. The next step is a non-linear transformation according to the hearing model of Sottek (see [b-Sottek]). This output is compared to the source signal which is also the 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 performs smoothing 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  $\lambda_1$  and  $\lambda_2$ , the weighting of 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 3D spectrograph displaying the deviation from the "close to the human expectation" between the estimated and the current signal is displayed versus time and frequency. The Relative Approach uses a time resolution of  $\Delta t = 6.66$  ms. The frequency range from 15 Hz to 24 kHz is divided into 128 frequency bands  $\Delta f_m$  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 for a questionnaire for subjective testing

(This annex forms an integral part of this Recommendation.)

#### C Performance rating

##### C.1 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 send direction (level, level variation, speech naturalness, etc.),
- speech quality in receive direction (level, level variation, speech naturalness),
- stability of the echo canceller for "closed loop" operation when doing car-to-car hands-free communication.

For the evaluation, some ITU-T Recommendations have served as guidelines, such as [ITU-T P.800], [ITU-T P.800.1] and others from the [ITU-T P.800] series Recommendations. 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 following items should be further considered for performance rating:

- 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 subcategory	Conversation type	Rating side	Rating type	Rating scales	Test condition/variation
Speech and background noise quality in send 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 startup, wiper, indicator
		Idle	Far-end	Adaptation to background noise	Adaptation to background noise	Transient background noise scenario: noise jump, e.g., fan/defrost startup
Speech quality in receive direction	Speech level	Single talk far-end	Near-end	Speech level (car, max. vol.)	Loudness preference (car) for max. vol.	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
Echo cancellation (EC)	Echo during single talk	Single talk far-end	Far-end	Echo	Disturbance caused by echo Echo characteristics (only to be rated if echo occurs): <ul style="list-style-type: none"> <li>• Intensity</li> <li>• Duration</li> <li>• Frequency of occurrence</li> <li>• Echo intelligibility</li> </ul>	Stationary background noise scenario: low/medium/high, fan/defrost; volume (car), movement of driver (enclosure dislocation)

**Table C.1 – Test categories and rating types overview**

Test category	Test subcategory	Conversation type	Rating side	Rating type	Rating scales	Test condition/ variation
	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): <ul style="list-style-type: none"> <li>• Intensity</li> <li>• Duration</li> <li>• Frequency of occurrence</li> <li>• Echo intelligibility</li> </ul>	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, max. volume

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

### **C.3 Speech and background noise quality in send direction**

#### **C.3.1 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 send 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

### C.3.2 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 send 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 rare 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

**C.3.3 Speech quality/speech naturalness**

Description:

The rating scale is applied in test cases which evaluate speech naturalness received on the far-end for different background noise scenarios on the near-end. This evaluation includes possible impairments caused by signal distortion and band limiting 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 send 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:
  - shrill, sharp, thin, tinny or muffled sounding speech,
  - 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 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

### C.3.4 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 send 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 is understood with no appreciable effort required.	4
Some words were hard to understand, moderate effort is required.	3
Many words were hard to understand, considerable effort required.	2
No meaning understood with any feasible effort.	1

### C.3.5 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 judgment, therefore, will get to a worse grade with the test scenario changing to a higher environmental noise. Therefore, it is problematic for this test category to be used as 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 send 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

**C.3.6 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 in clause C.3.7.

Test category:	Speech and background noise quality in send 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 it is 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 sounding unrelated to 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

**C.3.7 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 send 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 related to 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
Moderate unnatural sounding; Moderate distortion/synthetic sound	–	3
Clearly unnatural/distorted/synthetic sounding	–	2
Completely unnatural/distorted/synthetic sound	Unnatural	1

**C.3.8 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 on the far-end.

Test category:	Speech and background noise quality in send direction
Conversation type:	Idle
Rating side:	Rated at far-end
Test conditions:	Transient background noise: noise jump, e.g., fan/defrost startup
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 conducting such a test, the startup time of the defrost/fan 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 $\leq 1$ second		4
Adaptation time 2 ... 3 seconds		3
Adaptation time 3 ... 10 seconds		2
Adaptation time $\geq 10$ seconds	Very slow	1

#### **C.4 Speech quality in receive direction (in the car under test)**

##### **C.4.1 Speech sound quality/speech naturalness (receive)**

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 limiting 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 condition using nominal volume setting in the car.

Test category:	Speech quality in receive 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:
  - shrill, sharp, thin, tinny or muffled sounding speech,
  - emphasis on high frequencies or low frequencies.

For the rating scale, see Table C.4.

### C.4.2 Intelligibility/listening-effort (Receive)

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 nominal volume setting in the car.

Test category:	Speech quality in receive 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.

### C.4.3 Speech level (receive, 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 receive 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

### C.4.4 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) when engaged in a new call, after having set the volume to nominal in the prior call.

Test category:	Speech quality in receive 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**

For the evaluation of echo, as a first step, the disturbance caused by echo is judged. If echo is perceived, then an evaluation of additional echo characteristics is performed. The evaluation of these echo characteristics is 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 in the car is moving),
- 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, 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.

#### **C.5.1 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 [ITU-T P.831]

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

The rating scale is given in the following table.

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

**How would you judge the degradation/impairment/disturbance from 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

### C.5.2 Echo intensity

For diagnostic purposes only.

Description:

This rating is intended for diagnostic purposes and has only 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 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.13.

**Table C.13 – Echo intensity**

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

### C.5.3 Echo duration

For diagnostic purposes only.

Description:

This rating is intended for diagnostic purposes and has only 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 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.14.

**Table C.14 – Echo duration**

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

### C.5.4 Frequency of echo occurrence

For diagnostic purposes only.

Description:

This rating is intended for diagnostic purposes and has only 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 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.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	

### C.5.5 Echo intelligibility

For diagnostic purposes only.

Description:

This rating is intended for diagnostic purposes and has only 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	

### C.5.6 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 it is 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 sounding unrelated to the natural background will result in lower (worse) rating 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 sound, constant in sound and level, no artefacts	Natural	5
Slight difference between comfort noise and natural background noise perceivable; Slight distorted/synthetic sound, almost no artefacts, almost constant in sound and level	–	4
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

### C.5.7 Speech level variation during double talk

Description:

The scale is applicable to evaluate speech level variations during double-talk. The evaluation is done on both sides, 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: <ul style="list-style-type: none"> <li>• low background noise scenario</li> </ul> For rating at the far-end: <ul style="list-style-type: none"> <li>• stationary background noise scenarios of low/medium/high level;</li> <li>• additional background noise scenarios: different fan/defrost settings;</li> <li>• enclosure dislocation due to movement of the driver;</li> <li>• volume setting in the car: nominal, maximum.</li> </ul>
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 that are barely audible or that occur very rarely	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 side at all during double talk	1

### C.5.8 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 on both sides, 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: <ul style="list-style-type: none"> <li>• low background noise scenario</li> </ul> For rating at the far-end: <ul style="list-style-type: none"> <li>• stationary background noise scenarios of low/medium/high level;</li> <li>• additional background noise scenarios: different fan/defrost settings;</li> <li>• enclosure dislocation due to movement of the driver;</li> <li>• volume setting in the car: nominal, maximum.</li> </ul>
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 is understood during double talk with no appreciable effort required	4
Some words were hard to understand during double talk, moderate listening-effort is required	3
Many words were hard to understand during double talk, considerable listening-effort is 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 doing car-to-car hands-free communication. For the relevant tests, the hands-free system under test is installed in both cars. Moreover, both systems do not have 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 on 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 in a standstill position 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 adapt when the doors are closed afterwards). After that, the volume is set to maximum in both cars. For generating an impulse-like noise, for example, the doors of the cars could be slammed. Another possibility would be to quietly close the doors and generate impulse-like noise by clapping 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-filter: 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

## Annex D

### Standard set of user scenarios

(This annex forms an integral part of this Recommendation.)

**Table D.1 – Standard set of user scenarios used to collect noise recordings**

		Vehicle settings						Environmental conditions			
User scenario	Description	Vehicle speed	HVAC settings	Windows	Wipers	Turn signal	Back-ground talkers	Road surface (Note 3)	Wind speed	Precipitation	Temp.
1	Stationary vehicle	0 km/h (at idle)	FAN=Lowest setting	Up	Off	Off	None	N/A	< 5 m/s (12 mph)	None	> -20°C and < 40°C
2	City driving	60 km/h (37 mph)	FAN= Medium setting; AIRFLOW=Directed to windows	Up	Off	Off	None	Dry; rough road	< 5 m/s (12 mph)	None	> -20°C and < 40°C
3	Highway driving	120 km/h (75 mph)	FAN=Lowest setting	Up	Off	Off	None	Dry; rough road	< 5 m/s (12 mph)	None	> -20°C and < 40°C
4	Highway driving	120 km/h (75 mph)	FAN= Medium setting; AIRFLOW=Directed to windows	Up	Off	Off	None	Dry; rough road	< 5 m/s (12 mph)	None	> -20°C and < 40°C

**Table D.1 – Standard set of user scenarios used to collect noise recordings**

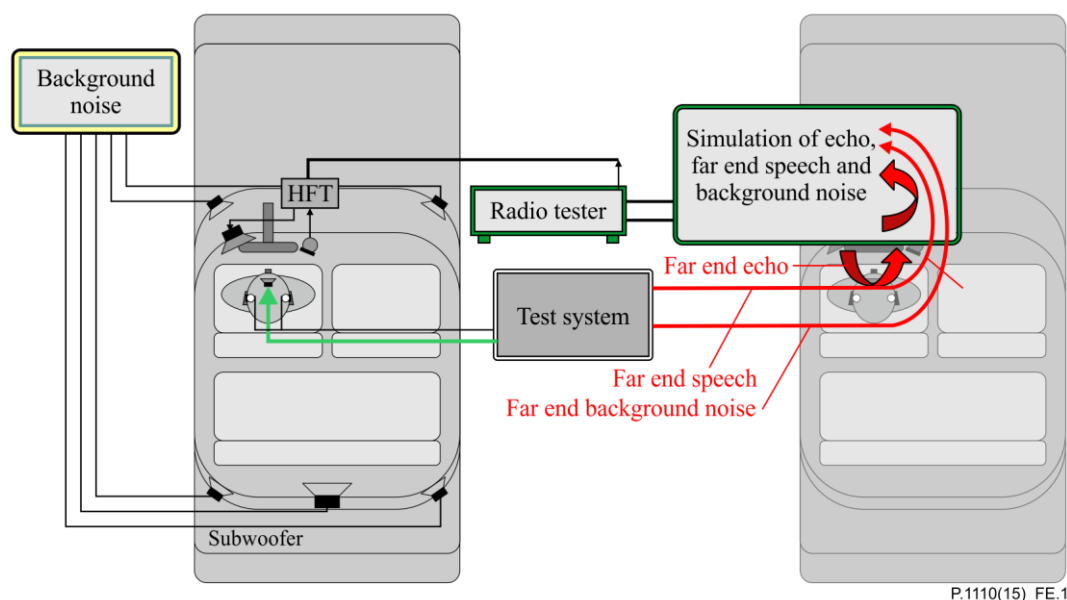
		Vehicle settings						Environmental conditions			
User scenario	Description	Vehicle speed	HVAC settings	Windows	Wipers	Turn signal	Back-ground talkers	Road surface (Note 3)	Wind speed	Precipitation	Temp.
5 (Note 1)	Highway driving	≥ 160 km/h (≥ 100 mph)	FAN=Lowest setting AIRFLOW=Directed to windows	Up	Off	Off	None	Dry; rough road	< 5 m/s (12 mph)	None	> -20°C and < 40°C
<p>NOTE 1 – Optional test: If the hands-free system is to be deployed in countries where higher driving speeds are allowed, then testing at 160 k.p.h. or higher should be performed, in addition to the standard set of user scenarios.</p> <p>NOTE 2 – Additional testing should be performed to verify that there are not any HVAC vent positions that result in severely degraded performance due to wind buffeting.</p> <p>NOTE 3 – Smooth road surfaces that generate very little tire noise shall not be used. Also, road surfaces with bumps that cause significant impulse noises shall not be used. If available, concrete surfaces are preferred because they often result in worst-case conditions that cause impairments not seen on other road surfaces.</p> <p>NOTE 4 – AIRFLOW refers to the HVAC mode settings related to how air is directed into the cabin. For example, in North American vehicles, there is typically a "Defrost" setting that will direct the flow of air toward the windows.</p> <p>NOTE 5 – The setting of the HVAC shall be documented. Some cars automatically adjust the HVAC settings during a call. In such cases the background noise recording shall be made with the HVAC setting active during the call.</p>											

## Annex E

### System stability with insufficient far-end echo loss

(This annex forms an integral part of this Recommendation.)

Car-to-car communication bears the risk of a closed loop feedback channel, especially when near-end and/or far-end echo loss is low. This may happen, e.g., at the beginning of a call, when ECs are not fully adapted or during echo path changes. In order to simulate this situation, the following laboratory test set-up is used:



Device under test (DUT, laboratory test set-up) Simulated vehicle at far end

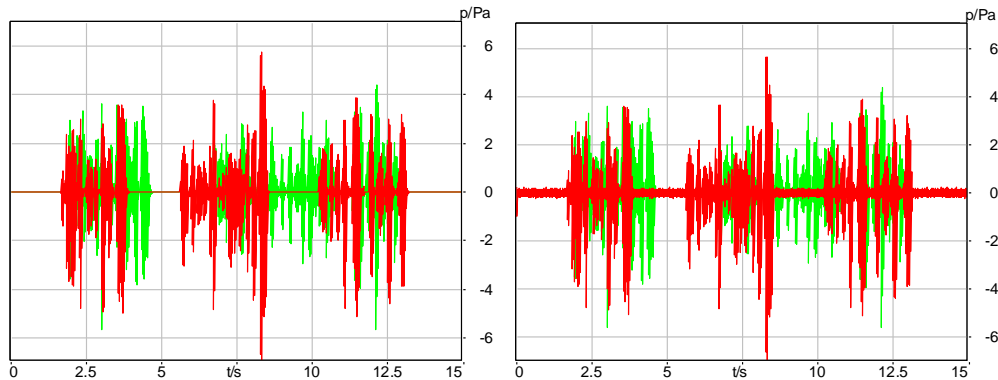
**Figure E.1 – Test set-up for simulating insufficient far-end echo loss**

The test set-up used to test car-to-car scenarios in a laboratory environment is shown in Figure E.1. The far-end side is simulated by an echo attenuation (echo return loss (ERL), echo path delay or measured impulse response in existing cars), the driver's voice (designated as "far end speech" and recorded from an HFT microphone in a test vehicle) and the background noise (recorded from an HFT microphone in a test vehicle).

Real speech, applied under single and double talk conditions for the DUT, is recommended in order to most realistically reproduce car-to-car communication. Two examples are shown above.

- The red coloured signal represents the simulated far-end signal, the green signal is applied via the artificial mouth in the test vehicle (DUT, laboratory set-up).
- The left-hand example represents the test signal without background noise in the far-end vehicle. The right-hand example shows a similar speech sequence but with added background noise at the far-end side.
- The Lombard effect should be considered on both sides in order to simulate realistic scenarios and realistic speech and noise levels. This can be achieved by:
  - i) using real Lombard speech recorded in the test lab; or
  - ii) choosing [ITU-T P.501] test sentences and considering the Lombard effect by the corresponding level adjustment. Guidelines for level adjustments as a function of background noise levels are given in clause 7.1.3;

- iii) the different driving conditions described in Annex D shall be considered including a no background noise condition.
- The sequences start with a double talk sequence, followed by a short single talk sequence. In the following graphs, the sending and receiving directions are represented by green and red, respectively.



The transmitted speech signals are recorded in receive direction of the DUT, i.e., in the vehicle in the laboratory set-up. No howling or feedback shall be detected.

### Requirements

The minimum far-end ERL still leading to a stable performance is documented.

This test shall be conducted for an echo path realization by a pure attenuation or by a simulated impulse response measured from a real car.

If different implementations are to be compared, it is recommended to simulate the far-end echo path by an attenuation.

### Test

- 1) The test arrangement is according to Figure E.1.
- 2) The tests are carried out under the following test conditions at the far-end side:
  - Test case 1 (mandatory): Simulation of ERL by pure attenuation, echo path delay 0 ms, variation of ERL from 50 dB to 0 dB (under study) in steps of 5 dB.
  - Test case 2 (customized): Simulation of ERL and echo path delay by measured impulse response in a real vehicle, variation of ERL from 50 dB to 0 dB (under study) in steps of 5 dB.
- 3) Each test shall be carried out with the EC coefficients initially cleared (reset, new call set-up) or starting from any artificial state (e.g., by converging EC with open drivers' door).
- 4) The recorded signals from the HATS (free-field or diffuse-field equalized (see clause 7.1.6.1)) are judged subjectively by the test engineers.

## Annex F

### Microphone related sound field reproduction technique – Multi-point noise simulation method (MPNS)

(This annex forms an integral part of this Recommendation.)

[b-ETSI TS 103 224] describes a background noise simulation method for handset, headset, and hands-free terminals, which are evaluated in an acoustically treated test room of a certain size (between 1,8 m × 2,4 m × 2,1 m to 8 m × 9 m × 4,5 m). It requires eight loudspeakers for the desired sound field reproduction, which are distributed around the HATS in a certain way to cover the most important directions of sound incidence in a typical background noise scenario. The necessary recordings require the same special microphone arrangement for all the use cases, where the eight microphones are arranged approximately in a circle with special spacings and vertical positions.

In motor vehicles, most of these requirements cannot be fulfilled: the vehicle interior is usually much smaller and the possible loudspeaker positions are much more limited. Furthermore, the required points of equalization are much more wide spread, which makes the circular microphone arrangement inapplicable. This also raises the need for another definition of a "fine-tuning set" of microphones which in [b-ETSI TS 103 224] are defined by a 10° rotation of the original ("calibration") set of microphones. The required adaptations of [b-ETSI TS 103 224] in order to use the multi-point noise simulation method (MPNS) successfully are described in the following:

#### F.1 Loudspeaker setup in the car

The multi-point noise simulation method requires at least 4 loudspeakers, which are arranged as described in [b-ETSI ES 202 396-1]. If more than 4 loudspeakers are used, the loudspeakers should not be placed close to each other but distributed in space as much as possible. The loudspeakers should not be placed in the "microphone space", i.e., between the microphones.

If the loudspeakers cannot reproduce frequencies down to 50 Hz, an additional subwoofer is needed. In this case, the weighted version of the Thikonov-equalization (Equation 9 in clause 6.2.4.1 of [b-ETSI TS 103 224]) is used during the inversion procedure, where the diagonal weighting matrix **W** accounts for the different frequency ranges of the loudspeakers.

Accordingly, the MPNS-Method is compatible with the loudspeaker arrangement of [b-ETSI ES 202 396-1] (four loudspeakers plus subwoofer).

#### F.2 Microphone setup in the car

No fixed microphone arrangement is required as it has to be arranged individually for each car's interior and for each application. Either the output signal of the microphones of the device under test (DuT) can be used directly for equalization or additional measurement microphones are placed in the vehicle interior close to the DuT microphones.

In any case, the same microphone arrangement is used for the recording of the background noise (see below) and the equalization of the system. The microphone characteristics like frequency or phase response and directionality shall be the same for both use cases.

[b-ETSI ES 202 396-1] differentiates between the "calibration set" of microphone positions which are used in all steps of the equalization procedure and the "fine-tuning set" of microphone positions which are only used in clauses 6.2.4.3.2 and 6.2.5. As the microphone arrangement is flexible the microphones have to be assigned to the calibration set and to the fine-tuning set. The fine-tuning set may be empty but, if "fine-tuning microphones" are used they shall be placed near the "calibration set" microphones, i.e., with a maximum distance of 5 cm.



If the DuT microphones can be used directly, the number of microphones in the "calibration set" must not exceed the number of loudspeakers in the setup excluding a potential subwoofer.

If it is not possible to use the DuT microphones directly, the number of loudspeakers shall be at least twice the number of microphones in the "calibration set". In general, it is recommended to place at least two measurement microphones near each location of DuT microphone(s) to create an "equalized zone" around the DuT microphone location. If it is not possible to use twice as many loudspeakers than microphones with this microphone arrangement, one of the two measurement microphones is assigned to the fine-tuning set of microphone positions.

### **F.3 Background noise recordings and reference noise**

Individual background noise recordings are made for each vehicle and microphone arrangement.

No "default" reference noise as in [b-ETSI TS 103 224] can be given for the proposed multi-point noise simulation method. Instead, one portion of the vehicle individual background noise recordings is selected as reference noise. The selected reference noise shall be 5 to 10 s long and representative for the background noise recordings in the individual car.

### **F.4 Equalization**

In general the equalization process is described in [b-ETSI TS 103 224]. The matrix size (used to identify the system and to calculate the matrix inversion for equalization) depends on the number of loudspeakers and microphones chosen for the background noise simulation. The following adaptations are required for the car:

In clause 6.2.3 of [b-ETSI TS 103 224], the impulse responses are pre-processed with a low-pass filter with a time-variant cut-off frequency to avoid a degradation of the quality of the inversion by high-frequency components of the tails of the impulse responses. Since the reverberation time in vehicle interiors is negligible compared to treated office rooms, the low-pass filtering is omitted for the proposed multi-point noise simulation method.

In clause 6.2.4.2 of [b-ETSI TS 103 224], different microphones are selected to calculate the inversion filters for each frequency band based on the corresponding wavelength. For the MPNS-Method, the microphone distances are unknown and therefore this type of optimization is not applied. All microphones are used across the whole frequency range to calculate the inversion filters.

In clause 6.2.4.3 [b-ETSI TS 103 224], the optimum regularization factor is searched subject to a limited loudspeaker output level to avoid distortion of the loudspeaker signals. "A *spectrum representative for the maximum level and spectral content*" ([b-ETSI TS 103 224], clause 6.2.4.3.1) can be calculated from the reference noise which was selected in F.3. For the MPNS-Method, the level of the filtered reference noise radiated by each loudspeaker to each microphone position shall not exceed the level which a pink noise of 0 dB [Pa] would generate in the same FFT-bin.

In clause 6.2.5 of [b-ETSI TS 103 224], a correction filter  $D(f)$  is applied to compensate, *inter alia*, for the arithmetic uncertainties (e.g., the regularization factor) of the inversion. This correction filter consists of a constant factor below 1.8 kHz and a minimum phase filter above. For the MPNS-Method, the microphone distances are unknown and therefore the lower cut-off frequency is reduced to 50 Hz. In order to retain the phase in the previously constant frequency range below 1.8 kHz, a zero-phase-filter is used instead of a minimum-phase filter for the frequency range above 50 Hz.

## F.5 Accuracy of the equalization

After a successful equalization the following criteria, which are comparable to those of [b-ETSI TS 103 224], shall be met:

1) Level accuracy

The level of the reproduced sound field at each microphone of the calibration set shall be accurate within  $\pm 1$  dB.

2) Magnitude and phase of the cross correlation between reference noise recording and simulated reference noise recording at the calibration position

The magnitude of the cross correlation between the reference noise recording (see clause F.3) and the reproduced signals at the calibration set positions averaged over the individual microphones shall fulfil the following requirements:

i) In the frequency range from  $100\text{ Hz}$  to  $1\text{ kHz}$  the magnitude of the complex coherence (normalized cross correlation spectrum) shall be larger than 0.9, measured in 1/3rd octaves.

ii) In the frequency range from  $100\text{ Hz}$  to  $1\text{ kHz}$  the phase of the complex coherence shall be accurate within  $\pm 10$  degrees and within  $\pm 30$  degrees in the range from  $1\text{ kHz}$  to  $1.5\text{ kHz}$ , both measured in 1/3rd octaves.

3) Spectrum reproduction accuracy

The difference between the amplitude spectrum of the original reference noise recording and the amplitude spectrum of the simulated reference noise recording (both measured in dB) for the individual microphones shall be within  $\pm 3$  dB, measured in 1/3rd octaves from 50 Hz to 10 kHz and  $\pm 6$  dB from 10 kHz to 16 kHz. The average spectrum accuracy, averaged over all microphones shall be within  $\pm 3$  dB from 50 Hz to 20 kHz.

## 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 below:

#### 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 for headset hands-free terminals, the test signal level is measured as "active speech level" according to [ITU-T P.56]. The speech activity should be between 30 per cent and 70 per cent. Finally, the level at the HATS-HFRP is adjusted to  $-25.7$  dBPa for speakerphone hands-free terminals.

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 "wideband" 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 per cent and 70 per cent.

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 right ear is used for the measurements. TMOS is determined using the settings "high quality handset" and "wideband" with TOSQA2001.

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

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

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

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

## Bibliography

- [b-ITU-T G.114] Recommendation ITU-T G.114 (2003), *One-way transmission time*.
- [b-ITU-T P.51] Recommendation ITU-T P.51 (1996), *Artificial mouth*.
- [b-ITU-T P.862] Recommendation ITU-T P.862 (2001), *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*.
- [b-ETSI EG 202 396-3] ETSI EG 202 396-3 V1.4.1 (2014), *Speech and multimedia Transmission Quality (STQ); Speech Quality performance in the presence of background noise; Part 3: Background noise transmission – Objective test methods*.
- [b-ETSI ES 202 396-1] ETSI ES 202 396-1 V1.5.1 (2014), *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*.
- [b-ETSI ES 202 739] ETSI ES 202 739 V1.3.1 (2009), *Speech and multimedia Transmission Quality (STQ); Transmission requirements for wideband VoIP terminals (handset and headset) from a QoS perspective as perceived by the user*.
- [b-ETSI ES 202 740] ETSI ES 202 740 V1.3.1 (2009), *Speech and multimedia Transmission Quality (STQ); Transmission requirements for wideband VoIP loudspeaking and handsfree terminals from a QoS perspective as perceived by the user*.
- [b-IEC 60268-4] IEC 60268-4 (2014), *Sound system equipment – Part 4: Microphones*.
- [b-3GPP TS 26.131] 3GPP TS 26.131 v12.3.0 (2014-12-22), *Terminal acoustic characteristics for telephony; Requirements*.
- [b-3GPP TS 26.132] 3GPP TS 26.132 v12.5.0 (2015-03-21), *Speech and video telephony terminal acoustic test specification*.
- [b-ETSI TS 103 224] ETSI TS 103 224 (2015), *Speech and multimedia Transmission Quality (STQ); A sound field reproduction method for terminal testing including a background noise database*.  
<<http://pda.etsi.org/pda/queryform.asp>>
- [b-Berger] Berger, J. (2001), *Results of objective speech quality assessment including receiving terminals using the advanced TOSQA2001*, ITU-T SG12, Contribution 20.
- [b-Fingscheidt] Fingscheidt, T., and Suhadi, S. (2007), *Quality assessment of speech enhancement systems by separation of enhanced speech, noise, and echo*, INTERSPEECH 2007, Antwerp, Belgium.
- [b-Kettler] Kettler, F., and Gierlich, H.W. (2008), *Evaluation of Hands-Free Terminals*, Topics in Speech and Audio Processing in Adverse Environments, edited by E. Hänsler, G. Schmidt, Springer, ISBN: 978-3-540-70601-4.
- [b-SAE J941] SAE J941 201003, *Motor Vehicle Drivers' Eye Locations*.  
<[http://standards.sae.org/j941\\_201003/](http://standards.sae.org/j941_201003/)>

- [b-Sottek] Sottek, R., Genuit, K. (2005), *Models of signal processing in human hearing*, International Journal of Electronics and Communications, pp. 157-165.
- [b-Steinert] Steinert, K., Suhadi, S., and Fingscheidt, T. (2009), *A Comparison of Instrumental Measures for Wideband Speech Quality Assessment of Hands-free Systems in Echoic Condition*, DAGA 2009, Rotterdam, The Netherlands.





## SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
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	Environment and ICTs, climate change, e-waste, energy efficiency; 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
<b>Series P</b>	<b>Telephone transmission quality, telephone installations, local line networks</b>
Series Q	Switching and signalling, and associated measurements and tests
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, next-generation networks, Internet of Things and smart cities
Series Z	Languages and general software aspects for telecommunication systems