Draft Recommendation P.emergency

Speech communication requirements for emergency calls originating from vehicles V0.4~~3~~

Summary

History

Keywords

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

FOREWORD

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

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

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

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

NOTE

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

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

INTELLECTUAL PROPERTY RIGHTS

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

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

Table of Contents

Recommendation P.emergency

Speech communication requirements for emergency calls originating from vehicles

1 Scope

This Recommendation defines use cases, requirements, and associated test methods for the speech communication for emergency call communication originating from vehicles using a dedicated emergency call system. This covers:

- built-in emergency call systems (manufacturer installed)
- after-market emergency call kits.

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

This Recommendation addresses the test of complete systems and covers the following use cases:

- The call is originated either automatically (or possibly manually) in the accident, in hands-free mode.
- The call is between the vehicle from where the emergency call is originated and the nearest PSAP.
- The requirements take into account talking and listening from all locations in the vehicle cabin.
- The requirements take into account "silent call", where information is obtained from background noise picked up by the emergency call system.
- The requirements focus primarily on achieving a sufficient level of intelligibility and communication quality.

The methods, the analysis and the performance parameters described in this Recommendation are based, where applicable on test signals and test procedures as defined in Recommendations ITU-T, ITU-T P.501, ITU-T P.502, ITU-T P.340 and ITU-T P.1100.

This Recommendation in principle covers speech communication requirements and tests for emergency call systems in narrowband and wideband mode. However, this version of the Recommendation covers narrowband only, wideband is for further study.

This Recommendation addresses the crash situation by simulation a post-crash situation as realistic as possible with respect to the impact of a post-crash situation on the acoustical environment. However, specific tests after a car crash are not in the scope of this Recommendation. It is assumed that the performance of an IVS system after a crash is adequately covered by simulating a post-crash situation in an un-crashed car.

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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Mis en forme : Français (Suisse)

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

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

[ITU-T P.863] Recommendation ITU-T P.862 (in force), *Perceptual objective listening quality assessment*
<http://www.itu.int/rec/T-REC-P.863>

[ITU-T P.863.1] Recommendation ITU-T P.863.1 (in force), *Application guide for recommendation P.863.1* .
<http://www.itu.int/rec/T-REC-P.863.1>

[ITU-T P.1100] Recommendation ITU-T P.1100 (in force), *Narrowband hands-free communication in motor vehicles.*
<http://www.itu.int/rec/T-REC-P.1100 >

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

[IEC 61260] IEC 61260 (1995), *Electroacoustics – Octave-band and fractional-octave-band filters*.
<http://webstore.iec.ch/webstore/webstore.nsf/artnum/019426>

3 Definitions

This Recommendation defines the following terms:

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

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

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

3.4 **diffuse-field equalization**: Equalization of the 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 [ITU-T P.58].

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

3.6 **free-field equalization**: The transfer characteristics of the artificial head are equalized in such a way that, for frontal sound incidence in anechoic conditions, the frequency response of the artificial head is flat. This equalization is specific to the HATS used.

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

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

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

3.10 head and torso simulator (HATS) for telephonometry: Manikin extending downward from the top of the head to the waist, designed to simulate the sound pick-up characteristics and the acoustic diffraction produced by a median human adult and to reproduce the acoustic field generated by the human mouth.

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

3.12 inboard ear: Ear closest to the centreline of the vehicle.

3.13 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.

3.14 mean opinion score – listening-only quality objective narrowband (MOS-LQOn): 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 (for further information see Annex A).

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

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

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

3.18 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 electromotive force to measured sound pressure).

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

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

4 Abbreviations and acronyms

The following abbreviations and acronyms are used:

ACR	Absolute Category Rating
A/D	Analogue/Digital
AGC	Automatic Gain Control
$A_{H,R}$	Attenuation Range in receive direction
$A_{H,R,dt}$	Attenuation Range in receive direction during double talk
$A_{H,S}$	Attenuation Range in send direction
$A_{H,S,dt}$	Attenuation Range in send direction during double talk
BGN	BackGround Noise

CSS	Composite Source Signal
D/A	Digital/Analogue
DI	Digital Interface
DRP	ear-Drum Reference Point
DTX	Discontinuous Transmission
EC	Echo Cancellation
ERL	Echo Return Loss
ERP	Ear Reference Point
FFT	Fast Fourier Transform
HATS	Head And Torso Simulator
HATS-HFRP	Head And Torso Simulator – Hands-Free Reference Point
HF	Hands-Free
HFT	Hands-Free Terminal
IVS	In-vehicle System
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
MRP	Mouth Reference Point
MSD	Minimum set of data
NR	Noise Reduction
PCM	Pulse Code Modulation
POI	Point Of Interconnection
PSAP	Public Safety Answering Point
QoS	Quality of Service
RF	Radio Frequency
RLR	Receive Loudness Rating
RLR_{AGC}	Minimum Receive Loudness Rating, triggered by AGC
SLR	Send Loudness Rating
S_{si}(diff)	Diffuse-field Sensitivity
S_{si}(direct)	Direct sound Sensitivity
S/N	Signal-to-Noise ratio
TCLw	weighted Terminal Coupling Loss
TMOS	TOSQA MOS
TOSQA	Telecommunications Objective Speech Quality Assessment
T_r	Receive Delay IVS

$T_{r,R}$	built-up time (receive direction)
$T_{r,S}$	built-up time (send direction)
T_{rtd}	Round Trip Delay IVS
T_s	Send Delay IVS

5 Conventions

dBm: Absolute power level relative to 1 milliwatt, expressed in dB.

dBm0: Absolute power level in dBm referred to a point of zero relative level (0 dBr point).

dBm0p: Weighted dBm0, according to [b-ITU-T O.41].

dBm0(C): C-weighted dBm0, according to [b-ISO 1999].

dBPa: Sound pressure level relative to 1 Pa, expressed in dB.

dBPa(A): A-weighted sound pressure level relative to 1 Pa, expressed in dB.

dBSPL: Sound pressure level relative to 20 µPa, expressed in dB; (94 dBSPL=0 dBPa).

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

dBr: Relative power level of a signal in a transmission path referred to the level at a reference point on the path (0 dBr point).

Vrms: Voltage – root mean square.

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

6 Test Setup and Preparation

6.1 Test arrangement

The acoustical interface for the in-vehicle system (IVS) is realized by using an artificial head (HATS – head and torso simulator) according to [ITU-T P.58]. The properties of the artificial head shall conform to [ITU-T P.58] for send as well as for receive acoustical signals.

All IVS emergency call implementations are connected to a system simulator conforming to the required transmission standard with implemented, calibrated audio interface.

For narrowband mode in GSM networks, the FR Codec or AMR codec can be used. If AMR codec is used, the bitrate of 12.2 kbit/s is used.

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

The test signals are fed electrically to the system simulator or acoustically to the artificial head. The test arrangement is shown in Figure 4.1

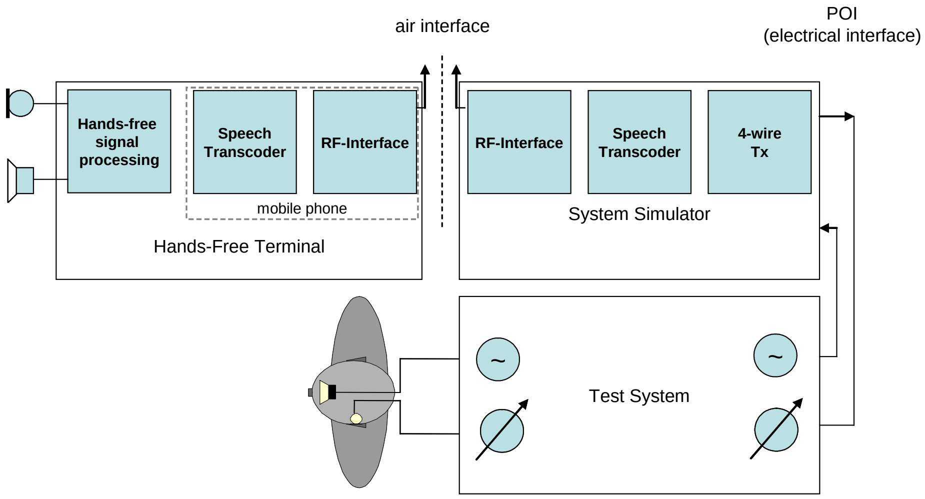


Figure 4.1 – Test arrangement for emergency call IVS (see[ITU-T P.1100])

6.2 Test arrangement in a car

The transmission performance of car hands-free terminals is measured in a car cabin. In order to simulate a realistic driving situation, background noise is inserted using a four-loudspeaker arrangement with subwoofer, while measurements with background noise are conducted. In Figure 4.1 the simulation arrangement is shown. More information on the test arrangement can be found in [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 ±1 dB. The third octave power density spectrum between 100 Hz and 10 kHz shall not deviate more than ±3 dB from the original spectrum. A detailed description of the equalization procedure as well as a database with background noises can be found in [b-ETSI ES 202 396-1].

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.

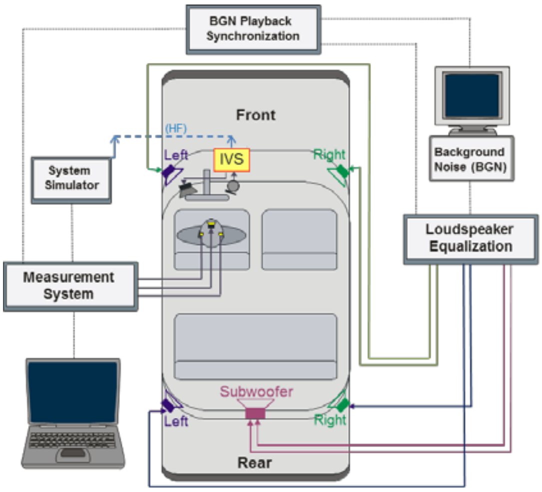


Figure 4.2 – Test arrangement with background noise simulation

6.3 Positioning of the emergency call IVS

The IVS, especially the acoustical interfaces (microphone/loudspeaker), are installed according to the requirements of the manufacturer. The positioning of the microphone/microphone array and loudspeaker are given by the manufacturer (e.g. also for aftermarket solutions) or defined by the installation in the vehicle. If no position requirements are given, the test laboratory will define the arrangements. Typically, the microphone is positioned close to the rear-view mirror, the loudspeaker can be positioned in the central console near the gear shift or below the dashboard on the drivers or front passengers side. In any case, the exact positioning has to be documented. IVS terminals installed by car manufacturers are measured in the original arrangement.

If not stated otherwise the measurements are conducted with the HATS placed at the drivers' position. For manually generated emergency calls it is assumed that the driver is communicating over the IVS with the public safety answering point (PSAP). Thus, a normal diver position is assumed and reproduced by the HATS during tests positioned in the driver's seat for the measurement, if not stated otherwise. For testing automatically generated emergency calls additional positions - the co-drivers' and the two outer passengers' back seat (2nd row), if available – are used as described in the individual test cases.

The drivers' position has to be in line with the average user's position; therefore, all positions and sizes of users have to be taken into account. Typically, all except the tallest 5% and the shortest 5% of the driving population have to be considered. The size of these persons can be derived, e.g., from the 'anthropometric data set' for the corresponding year (based on data used by the car manufacturers for example). The position of the HATS (mouth/ears) within the positioning arrangement is given individually by each car manufacturer. The position used has to be reported in detail in the test report. If no requirements for positioning are given, the distance from the microphone to the MRP is defined by the test laboratory.

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

It is recommended to verify some performance parameters especially for automatically generated emergency calls with the HATS positioned on the front passengers seat or on the passengers' seat in the 1st row behind the front passengers seat.

6.4 Operation Modes of emergency call IVS

An IVS can be used for manually or automatically generated emergency calls.

- For **manually generated emergency calls** a typical driving situation (constant speed), a parking vehicle or a vehicle involved in a minor accident (with no automatic emergency call generation) can be assumed.. It is further assumed that the driver or other vehicle occupant is still able to communicate with the PSAP in the same way as in a normal hands-free communication.
- For **automatically generated emergency calls** (detected and initiated by various sensors in the car) the vehicle is typically not moving, windows may be broken, a higher influence of ambient noise from outside the vehicle can be assumed (road noise, passing vehicles,…), passengers in the vehicle or even first-aiders may communicate with the PSAP, if the driver is unable to communicate.

Both scenarios are covered by this Recommendation.

6.5 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].

The sound pressure level is calibrated at the HATS-HFRP so that the average level at 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].

The Lombard effect refers to the change in speaking behaviour caused by acoustic noise. As no data are available to analyse the typical speech level in an emergency case and under emergency call specific noise scenarios, the output level of the mouth is increased to account for the "Lombard effect" in a non-emergency situation considering the known formulas [ITU-T P.1100]. The level is increased by 3 dB for every 10 dB that the long-term A-weighted noise level exceeds 50 dB(A). This relationship is shown in the following formula:

$$I(N)=\begin{cases} 0 & \textit{for} \quad N<50 \\ 0.3(N-50) & \textit{for} \quad 50\le N<77 \\ 8.0 & \textit{for} \quad N\ge 77 \end{cases}$$

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 send direction is taken into account independently (see section 9 on test signal levels).

6.6 Artificial ear

For IVS terminals the ear signals of both ears of the artificial head are used. The artificial head is diffuse-field equalized, more detailed information can be found in [ITU-T P.581].

6.7 Influence of the transmission system

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

6.8 Calibration and equalization

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

6.8.1 Calibration

- Acoustical calibration of the measurement microphones as well as of HATS microphone.
- Calibration and equalization of the artificial mouth at the MRP.
- HATS-HFRP calibration

Equalization

- Free-field equalization of the artificial head, in case of more than one loudspeaker diffuse-field equalization is used.

6.8.2 Reference measurement

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

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

6.9 System simulator settings

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

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

6.10 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.

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.

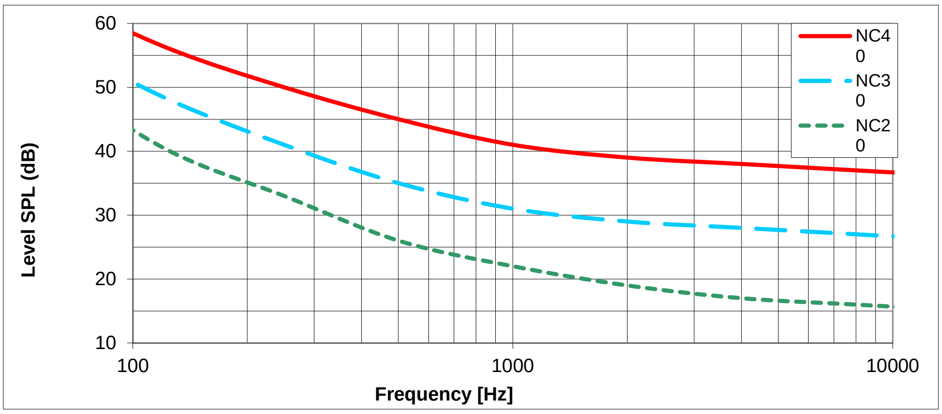


Figure 4.3 – NC-criteria for test environment

7 Test signals and test signal levels

7.1 Signals

Speech-like signals are used for the measurements which can be found in [ITU-T P.501]. Detailed information about the test signal used is to be 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: For single talk measurements, in cases where it can be shown that the IVS signal processing does not affect the measurement result 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 (1 male, 1 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).

For narrow-band IVS, all test signals – which are used in the receive direction – have to be band-limited. The band-limitation is achieved by bandpass filtering in the frequency range between 200 Hz and 4 kHz using bandpass filtering providing 24 dB/octave. In the 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 the receive direction, the band-limited test signal is measured; in the send direction no band-limitation is applied.

The average signal levels for the measurements are as follows:

– –16 dB_{m0} in the receive direction (typical signal level in networks).

– –1.7 dBPa in the send direction at the MRP (typical average speech levels,equivalent to –25.7 dBPa at the HATS-HFRP).

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

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

7.2 Background noise signals

For some measurements, typical background noise is inserted. This is described in the corresponding clauses. When playback of background noise is required, background noise conditions defined in this document . Other noise situations, may also be taken into account. In general, it is recommended to differentiate between the two operation modes for IVS and choose typical noise scenarions accordingly:

- **automatically generated emergency calls (A):**
  simulated emergency call call noise scenario (A1), e.g. stationary car (parking on a highway parking place), engine off, all 4 windows open, passing vehicles,
  simulated emergency call call noise scenario (A2), e.g. stationary car (parking on a highway parking place), engine off, all 4 windows open, passing vehicles, additional voice babble from outside of the vehicle
  spectrally adapted stationary noise to reproduce spectral content of scenario A1 (A3): white gaussian noise filtered by the average spectrum drived from scenario A1
- **manually generated emergency calls (B);**
  constant driving conditions simulating fixed driving speed (e.g., 130 km/h).

7.2.1 Recording of background noise

Background noise under constant driving condition (noise scenarios B) is recorded in the vehicle being tested. The measurement microphone is positioned close to the IVS microphone. Alternatively, the IVS microphone can be used for the recording of the background noise if the microphone is easily accessible.

NOTE – In case of microphone arrays the best simulation would be to record the electrical output signals of all microphones and insert them electrically as described below, since the 4-loudspeaker arrangement does not allow a real sound-field reproduction. With this methodology, structure-borne noise and wind noise coupled to the microphone can also be included.

Background noise in a simulated emergency scenario (A1, A2) can also be recorded with the vehicle being tested. However, the acoustic condition (stationary car, engine off, 4 windows open), may justify using a recorded noise scenario with another vehicle.

Commenté [Gie1]: To be discussed whether to include prerecorded noise scenarios in this Recommendation

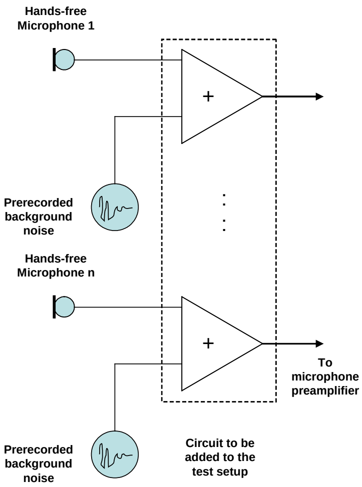
7.2.2 Playback of the recorded background noise

Two ways of background noise playback are recommended:

1) The test laboratory employs a 4-loudspeaker arrangement for acoustic background noise reproduction in the car cabin. Typically 2 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 IVS microphone and the artificial head at the driver's seat. Details can be found in [b-ES 202 396-1].

2) The background noise can be inserted electrically to the microphone signal and to the reference microphone positioned close to the IVS microphone. Therefore the background noise signals recorded at the electrical output of the IVS 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 has to be provided, see Figure xx. The test laboratory has to ensure the right calibration of the two signals.

NOTE – Both with analogue as well as digital electrical feedback of the noise signal structure-borne noise can be captured as well.



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

Figure 4.4 – Set-up for analogue electrical insertion of the pre-recorded background noise signal

8 Measurement parameters and requirements for IVS terminals

8.1 Preparation measurements

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

8.2 Delay

8.2.1 Requirements

In general the delay consists of an access specific delay and the hands-free implementation dependant delay.

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

The IVS roundtrip delay T_{rtd} consists of

- the hands-free signal processing in send and receive
- the access specific delay in send and receive
- the air-paths from mouth to microphone and from the loudspeakers to the ear):

T_{rtd} = Ts + Tr

(including the delay in send direction plus access delay in send plus the delay in receive direction plus access delay in receive).

T_{rtd} shall be less than T_{rtd} defined for mobile terminals in hand-held mode in those standards dealing with the same access technology. In case a delay performance objective is defined this performance objective shall be met.

Note 1 - For 3GPP UMTS circuit-switched speech and 3GPP LTE MTSI-based speech, definitions, test methods, performance objectives and requirements are found in 3GPP TS 26.131 [X] and TS 26.132 [Y].

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 G.114 [xx].

8.2.2 Delay in send direction

8.2.2.1 Test

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

$$T_s + \text{Tsystem}$$

NOTE 1 – The delay should be minimized.

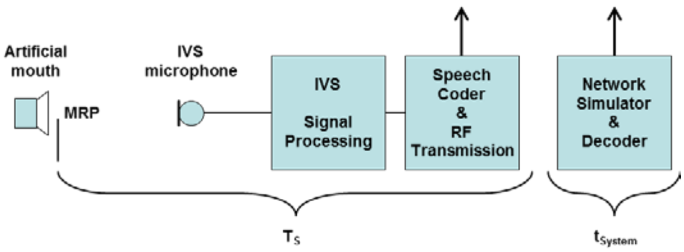


Figure 5.1 – Different blocks contributing to the delay in send direction

The system delay Tsystem is dependent on the transmission method used and the network simulator. The delay Tsystem must be known.

1) For the measurements, a composite source signal (CSS) according to [ITU-T P.501] is used. The pseudorandom noise (PN) part of the CSS has to be longer than the maximum expected delay. It is recommended to use a PN sequence of 16 k samples (with 48 kHz sampling rate). The test signal level is –1.7 dBPa at the MRP. 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 IVS terminal is in accordance with clause 64. The HATS is positioned on the drivers' seat.

2) The delay is determined by the 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.

8.2.3 Delay in receive direction

8.2.3.1 Test

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

$$T_r + Tsystem$$

NOTE 1 – The delay should be minimized.

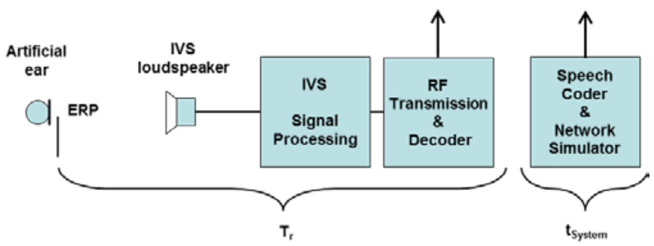


Figure 5.2 – Different blocks contributing to the delay in receive direction

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

1) For the measurements a composite source signal (CSS) in accordance with [ITU-T P.501] is used. The pseudorandom noise (PN) part of the CSS should be longer than the maximum expected delay. It is recommended to use a PN sequence of 16 k samples (with 48 kHz sampling rate). The test signal level is –16 dB_{m0} at the electrical interface (POI).

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

2) The test arrangement is in accordance with clause 4. The HATS is positioned on the drivers' seat. For the measurement the artificial head is free-field or diffuse-field equalized according to [ITU-T P.581]. The equalized output signal of the inboard ear is used for the measurement.

3) The delay is determined by cross-correlation analysis between the measured signal at the DRP and the original signal. The measurement is corrected by delays which are caused by the test equipment.

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

8.3 Loudness ratings

8.3.1 Requirements

The SLR between the MRP and the electrical reference point (POI) shall be:

$SLR \leq 2$2~~7~~ dB for all positions in the car (the HATS positioned either at the drivers' position, the co-drivers' and the two outer passengers' back seat (2nd row), if available);

A SLR = 13 dB +/– 4 dB for HATS positioned at the drivers' position is recommended.

The nominal (default) RLR between the POI and the artificial ear of the HATS shall be:

$RLR \leq 10$ dB for all positions in the car (the HATS positioned either at the drivers' position, the co-drivers' and the two outer passengers' back seat (2nd row), if available);

A RLR = -3 dB +/– 4 dB for HATS positioned on the drivers' seat is recommended.

Note 1: It is recommended to use background noise controlled AGC in receiving direction. The AGC should be designed to allow a SNR of ≥ 6 dB for all signal and noise conditions.

If a user-specific volume control is provided the requirement for RLR given above shall be measured at least for one setting of the volume control. This shall be the default setting. It is recommended to provide a volume control (manually or AGC controlled) which allows a loudness increase by at least 10 dB referred to the nominal value of RLR. The volume control range shall allow the setting of S/N ≥ 6 dB for all signal and noise conditions. This will allow sufficient loudness of the speech signal in the receive direction in the presence of high background noise. The minimum achievable RLR (maximum loudness) is called RLR_{AGC}.

8.3.2 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 –1.7 dBPa at the MRP. The test signal level is the average level of the complete test signal. 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 4. Tests are carried out with the HATS positioned on the drivers' seat and additionally carried out with the HATS on the back passengers seat (2nd row). The send sensitivity is calculated from each band of the 14 frequencies given in Table 1 of [ITU-T P.79], bands 4-17.

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

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

8.3.3 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 level is –16 dB_{m0}, measured at the electrical reference point and averaged over the complete test signal sequence.

2) The test arrangement is according to clause 4. Tests are carried out with the HATS positioned on the drivers' seat and additionally carried out with the HATS on the back passengers' seat (2^{nd} row). The artificial head is free-field or diffuse-field equalized 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 over the total time of analysis; the "right" and "left" signals are voltage-summed for each 1/3 octave band frequency band.

   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 [ITU-T P.79], Annex A without the LE factor.

4) The correction –8 dB in accordance with [ITU-T P.581] is used for the correction of the measurement results.

5) The test is repeated for maximum volume control setting in case a manually volume control is provided or the maximum gain setting (controlled by AGC in receiving direction) can be adjusted in test mode in order to verify the recommended control range.

8.4 Variation of RLR in presence of background noise

The intention of this test is the verification of the amplification range introduced by the AGC implementation in receiving direction. The RLR is determined in silent conditions (see test 5.3.3) and additionally in presence of background noise. The level of background noise needs to be sufficiently high (*to be defined*).

A stationary background noise scenario (A3) is used for testing in order to avoid additional AGC adjustment due to time variant level fluctuation in the background noise itself.

8.4.1 Requirements

The IVS shall cover a receiving loudness rating range from RLR -3 dB +/– 3 dB up to RLR_{AGC}. The RLR shall be automatically adjusted during background noise playback in the vehicle cabin. The adjustment shall be reached within 2s (*to be defined*).

8.4.2 Test

1) The RLR is determined as described in chapter 8.5 (Test loudness rating in receive direction).
2) In order to guarantee a stable and constant gain setting provided by the AGC the A3 background noise signal is used for this test instead of an emergency call specific noise scenario. The noise is played back in the vehicle and recorded with the HATS ("noise reference recording"). The background noise playback needs to be exactly synchronized to the recording process in order to guarantee reproducibility of the noise audio recording.
3) The background noise playback and recording is repeated coincident to the playback of the speech sequence used for RLR calculation. The HATS records background noise and speech sequence ("speech and noise recording"). The background noise playback needs to be exactly

synchronized to the recording process in order to guarantee reproducibility of the noise audio recording.

4) The “noise reference recording” is subtracted from the “speech and noise recording” in the time domain. This minimizes the level of recorded background noise, improves the signal to noise ratio and allows the accurate noise-free analysis of RLR_{AGC}.
5) The RLR_{AGC} is determined as described in chapter 8.3.3.

Further details and verification of the analysis method are under study.

8.5 Sensitivity frequency responses

8.5.1 Send sensitivity frequency response

8.5.1.1 Requirements

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.

Editors Note: The tolerance to be applied for IVS is under study. Such tolerance may be optimized for lower listening effort / higher speech intelligibility in the presence of background noise. Provisionally the mask as below is used.

Table 5-1 – Tolerance mask for the send sensitivity frequency response

Frequency [Hz]	Upper limit	Lower limit
200	4 dB	–∞ dB
250	4 dB	–∞ dB
315	4 dB	–∞ dB
400	4 dB	–∞ dB
500	4 dB	–8 dB
630	4 dB	–7 dB
800	4 dB	–6 dB
1 000	4 dB	–4 dB
1 300	6 dB	–4 dB
1 600	7 dB	–4 dB
2 000	8 dB	–4 dB
2 500	8 dB	–4 dB
3 100	8 dB	–4 dB
4 000	04dB	–∞ 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.

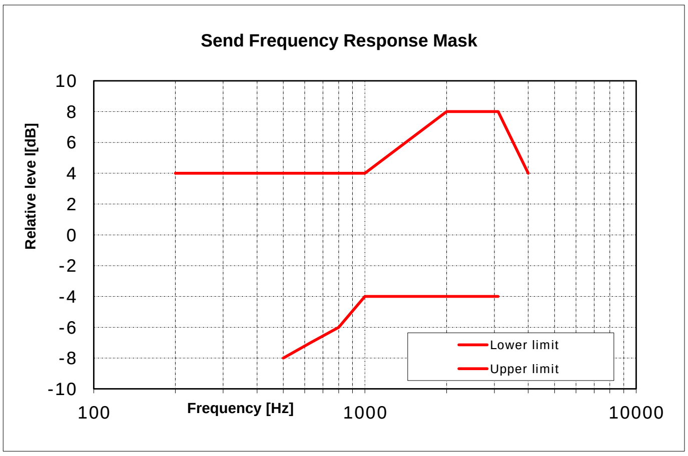


Figure 5-3 – Send frequency response mask (Fig. is informative)

8.5.1.2 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 –1.7 dBPa at the MRP. The test signal level is the average level of the complete test signal. 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. The sending frequency response is measured from the MRP to the output of the speech codec (POI).

2) The test arrangement is according to clause 6. The send sensitivity frequency response is determined in one-third octave intervals as given by [IEC 61260] for frequencies of 100 Hz to 4 kHz, inclusive. In each one-third octave band, the level of the measured signal is referred to the level of the reference signal averaged over the complete test sequence length.

3) The sensitivity is determined in dBV/Pa.

8.5.2 Receive sensitivity frequency response

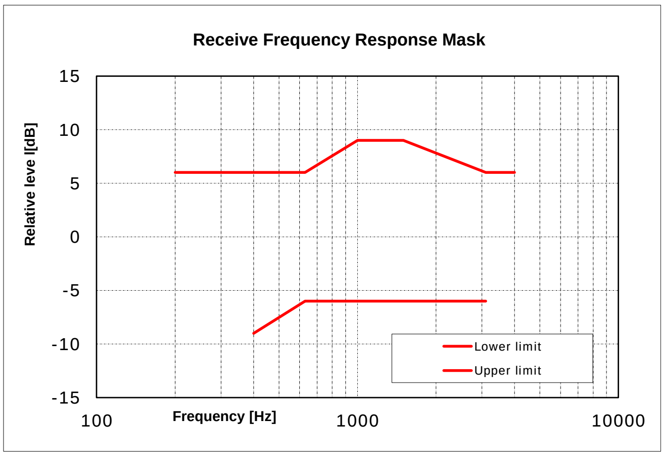
8.5.2.1 Requirements

The tolerance mask for the receive sensitivity frequency response is shown in Table 5.2, the mask is drawn by straight lines between the breaking points in Table 5.2 on a logarithmic (frequency) – linear (dB sensitivity) scale.

Editors Note: This tolerance mask to be applied for IVS is under study. Such tolerance may be optimized for lower listening effort / higher speech intelligibility in the presence of background noise. Provisionally the mask as below is used.

Table 5.2 – Tolerance mask for the receive sensitivity frequency response

Frequency [Hz]	Upper limit	Lower limit
200	6 dB	–∞ dB
250	6 dB	–∞ dB
315	6 dB	–∞ dB
400	6 dB	–9 dB
630	6 dB	-6 dB
1000	6 dB	-6 dB
1 500	9 dB	-6 dB
3 100	9 dB	–6 dB
4 000	6 dB	–∞ dB
NOTE – All sensitivity values are expressed in dB on an arbitrary scale.		



NOTE – This figure is informative.

Figure 5-4 – Receive frequency response mask

NOTE 1 – The limit curves shall be determined by straight lines joining successive co-ordinates 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 2 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, the present document uses the diffuse-field as reference. 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.

8.5.2.2 5.7.2.2 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 dBm0, measured at the electrical reference point and averaged over the complete test signal sequence.

2) The test arrangement is in accordance with clause 6. The receive sensitivity frequency response is measured from the electrical reference point (input of the system simulators, POI) to free-field or diffuse-field equalized HATS 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.

3) The sensitivity is determined in dBPa/V.

8.6 Speech quality during single talk

Commenté [gi2]: change to an intelligibility measurement if available

8.6.1 One-way speech quality in send

8.6.1.1 Requirement

8.6.1.2 ~~5~~ Test

8.6.2 One-way speech quality in receive

8.6.2.1 Requirement

8.6.2.2 Test

~~8.7~~

Mis en forme : Retrait : Gauche : 1,02 cm, Sans numérotation ni puces

~~8.8~~8.7 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) inside the car, a minimum background noise level of –64 dBPa(A) should not be exceeded.

~~8.8.1~~8.7.1 Idle channel noise in send direction

~~8.8.1.1~~8.7.1.1 Requirements

The maximum idle channel noise in the send direction, measured at the electrical reference point (POI) in quiet conditions shall be less than –64 dB_{m0}(P).

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

~~8.8.1.2~~8.7.1.2 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 4 composite source signals according to [ITU-T P.501]. The spectrum of the test signal at the MRP is equalized under free-field conditions. The level of the activation signal is –25.7 dBPa, measured at the HATS-HFRP.

2) The test arrangement is described in clause 6.

3) The idle channel noise is measured at the electrical reference point in the frequency range between 100 Hz and 4 kHz. The measurement requires a time window which starts exactly at the end of the activation signal. Any "ringing" of filters or receivers or reverberance

influence shall be taken into account, the time window must be shifted accordingly. The length for the time window is 1 second, which is the averaging time for the idle channel noise. The test laboratory 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.

4) The idle channel noise is determined by psophometric weighting.

5) Spectral peaks are measured in the frequency domain. The frequency spectrum of the psophometrically weighted idle channel noise is measured by a spectral analysis having a noise bandwidth of 8.79Hz (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 up to 3.4 kHz.

~~8.8.2~~8.7.2 Idle channel noise in receive direction

~~8.8.2.1~~8.7.2.1 Requirements

The requirements for the maximum noise produced by the IVS in case no signal is applied to the receive direction are as follows:

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

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

~~8.8.2.2~~8.7.2.2 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 4 composite source signals according to [ITU-T P.501]. The level of the activation level is adjusted to –16 dB_{m0}, 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 6. For the measurement of the IVS the artificial head is free-field equalized according to [ITU-T P.581]. The equalized output signal of the inboard ear is used for the measurement.

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 1 second, which is the averaging time for the idle channel noise.

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

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

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 8.79Hz (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 up to 3.4 kHz.

~~8.9~~8.8

~~8.10~~8.9 Echo performance without background noise

Due to the expected delay in networks, the echo loss presented at the electrical reference point (POI) should be at least 46 dB during single talk. This echo loss (TCL_W) should be achieved for a wide range of acoustical environments.

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

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

~~8.10.1~~8.9.1 Terminal coupling loss (TCLw)

~~8.10.1.1~~8.9.1.1 5.5.1.1 Requirements

The TCL_W in quiet environments should be at least 46 dB for nominal and for maximum setting of the volume control. The implemented echo control mechanism should provide sufficient echo loss for all typical environments and typical impulse responses.

NOTE: A TCLw of ≥ 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 send direction, it may not always be possible to measure an echo loss ≥ 50 dB.

When conducting the tests, it should be checked whether the signal measured is an echo signal and not comfort noise inserted in the send direction in order to mask an echo signal. This should be checked and verified during the tests, e.g., by comparing the analysis with the idle channel noise measurement results.

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

~~8.10.1.2~~8.9.1.2 5.5.1.2 Test

1) All tests are conducted in the car cabin; the test arrangement is described in clause 6xx. The noise level measured at the electrical access point (idle channel noise) shall be less than –63 dB_{m0}. 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 dB_{m0}.

3) The first 17.0 s of the test signal (6 sentences) are discarded from the analysis to allow for convergence of the acoustic echo canceller. The analysis is performed over the remaining length of the test sequence (last 6 sentences).

4) TCL_W is calculated according to clause B.4 of [ITU-T G.122], (trapezoidal rule). For the calculation, the average measured echo level at each frequency band is referred to the average level of the test signal measured in each frequency band. For the measurement, a time window has to be applied which is adapted to the duration of the actual test signal.

~~8.10.2~~8.9.2 Echo level versus time

~~8.10.2.1~~8.9.2.1 Requirements

This test is intended 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 should 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 should be less than 6 dB.

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

~~8.10.2.2~~8.9.2.2 Test

1) The test arrangement is in accordance with clause 6.

2) The test signal consists of a periodically repeated composite source signal according to [ITU-T P.501] with an average level of –5 dB_{m0} as well as an average level of –25 dB_{m0}. The echo signal is analysed during a period of at least 2.8 s, which represents 8 periods of the CS signal. The integration time for the level analysis shall be 35 ms, the analysis is referred to the level analysis of the reference signal. In addition, the test is repeated with 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 dB_{m0}. The echo signal is analysed during the complete test signal. The integration time for the level analysis shall be 35 ms.

3) When using the CS signal the measurement result is displayed as attenuation versus 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).

~~8.10.3~~8.9.3 Echo performance with time variant echo path and speech

~~8.10.3.1~~8.9.3.1 Requirements

To be discussed

~~8.10.3.2~~8.9.3.2 Test

To be discussed

~~8.11~~8.10 Switching characteristics

~~8.11.1~~8.10.1 Activation in send direction

The activation in the send direction is mainly determined by the built-up time $T_{r,S,min}$ and the minimum activation level ($L_{S,min}$). The minimum activation level is the level required to remove the inserted attenuation in the 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 below is always referred to the test signal level at the mouth reference point (MRP).

~~8.11.1.1~~8.10.1.1 Requirements

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

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

~~8.11.1.2~~8.10.1.2 Test

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

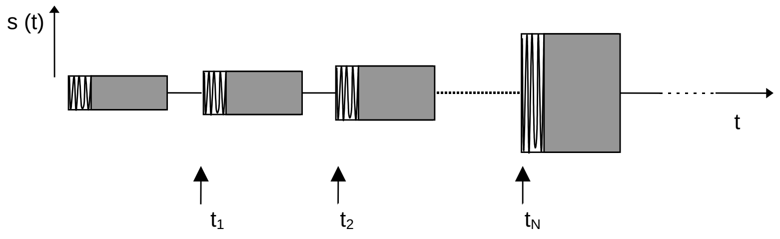


Figure xx – Test signal to determine the minimum activation level and the built-up time

The settings of the test signal are as follows.

Table xx – Settings of the CSS in send

	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.			

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 6.

2) The level of the transmitted signal is measured at the electrical reference point. The measured signal level is referred to the test signal level and displayed versus time. The levels are calculated from the time domain using an integration time of 5 ms.

3) The minimum activation level is determined from the CSS burst which indicates the first activation of the test object. The time between the beginning of the CSS burst and the complete activation of the test object is measured.

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

~~8.11.2~~8.10.2

~~8.12~~8.11 Double talk performance

NOTE – Before starting the double talk tests, the test laboratory 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.

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.

~~8.12.1~~8.11.1 Attenuation range in send direction during double talk: $A_{H,S,dt}$

~~8.12.1.1~~8.11.1.1 Requirements

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

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

Category	1	2a	2b	2c	3
	Full duplex capability	Partial duplex capability			No duplex capability
$A_{H,S,dt}$ [dB]	≤ 3	≤ 6	≤ 9	≤ 12	> 12

The IVS should provide a double talk capability of type 2b or better in sending direction (*preliminary*). The requirement apply for nominal (default) setting of the receive volume control.

Commenté [H3]: Class to be decide

Waiting for measurement results in January
-by Peiker
-by Novero
-by Gemalto
-by HEAD acoustics

Decision based on measurement results

It should be <=2c.

The requirements apply for nominal signal levels in the send and receive directions as well as for the level combinations nominal level in receive/–6 dB (re. nominal level) in send.

In general, Table xx 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.

~~8.12.1.2~~8.11.1.2 Test

The test signal to determine the attenuation range during double talk is shown in Figure 5.xx. 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.

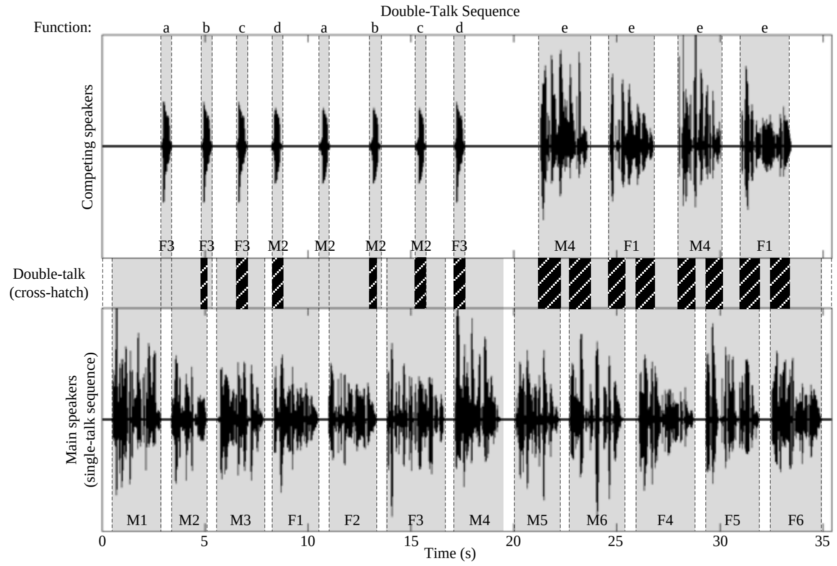


Figure xx – Double talk test sequence with overlapping speech sequences in send and receive direction

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

The settings for the test signals are as follows:

Table 5.9 – Timing of the double talk sequences

	Receive direction	**Send direction**
Average signal level (assuming an original pause length of 101.38 ms)	–16 dB_{m0}	–1.7 dBPa

1) The test arrangement is in accordance with clause 6. 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 dB_{m0} 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 Appendix III of [ITU-T P.502]. The double talk performance is analysed for each word and sentence produced by

the competing speaker. The requirement has to be met for each word and sentence produced by the competing speaker.

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

~~8.12.2~~8.11.2 Attenuation range in receive direction during double talk: $A_{H,R,dt}$

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

~~8.12.2.1~~8.11.2.1 Requirements

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

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

Category	1	2a	2b	2c	3
	Full duplex capability	Partial duplex capability			No duplex capability
$A_{H,R,dt}$ [dB]	≤ 3	≤ 5	≤ 8	≤ 10	> 10

The IVS should provide a double talk capability of type 1 in receiving direction. The requirements apply for nominal setting of the receive volume control.

Commenté [gi4]: 2b? based on Novero and Peiker contributipns?

In general, Table xx 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.

~~8.12.2.2~~8.11.2.2 Test

The test signal to determine the attenuation range during double talk is shown in Figure 5.11. 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 should be constant during the measurement.

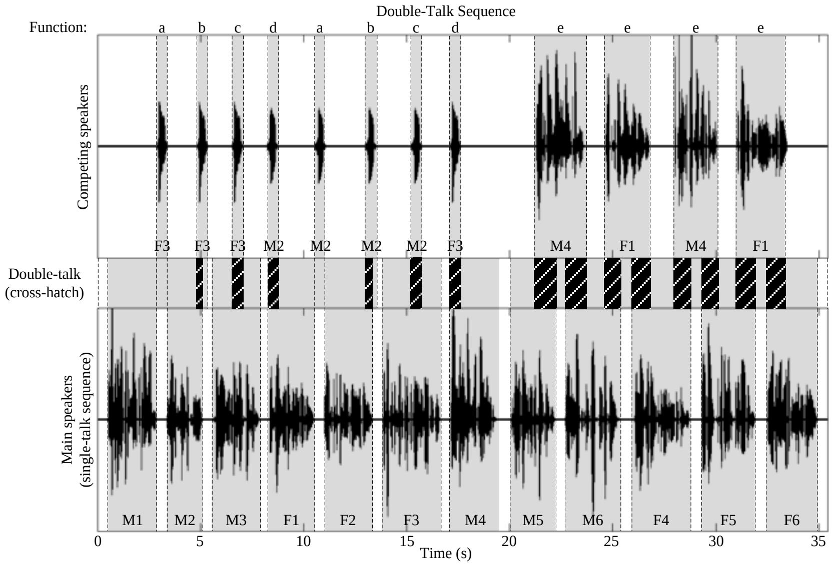


Figure 11 – Double talk test sequence with overlapping speech aequences in receive and send direction

The settings for the test signals are as follows:

Table 11 – Timing of the double talk sequences

	Receive direction	**Send direction**
Average signal level (assuming an original pause length of 101.38 ms)	–16 dB_{m0}	–1.7 dBPa

.

1) The test arrangement is in accordance with clause 6.

2) When determining the attenuation range in receive direction the signal measured at the loudspeaker of the IVS terminal is referred to the test signal inserted.

3) The attenuation range during double talk is determined as described in Appendix III of [ITU-T P.502]. The double talk performance is analysed for each word and sentence produced by the competing speaker. The requirement has to be met for each word and sentence produced by the competing speaker.

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

~~8.12.3~~8.11.3 Detection of echo components during double talk

~~8.12.3.1~~8.11.3.1 Requirements

The echo attenuation during double talk is based on the parameter talker echo loudness rating ($TELR_{dt}$). It is assumed that the terminal at the opposite end of the connection (PSAP side) provides nominal loudness rating (SLR + RLR = 10 dB). "Echo loss" is the echo suppression provided by the

IVS measured at the electrical reference point. Under these conditions, the requirements given in Table 12 are applicable (more information can be found in Annex A of [ITU-T P.340]).

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

Category	1	2a	2b	2c	3
	Full duplex capability	Partial duplex capability			No duplex capability
Echo loss [dB]	≥ 27	≥ 23	≥ 17	≥ 11	< 11

Mis en forme : Normal

The IVS should provide an echo-loss during double talk capability of type 2b. The requirements apply for nominal setting of the receive volume control.

Commenté [gi5]: 2b based on Novero and Peiker contributipns

~~8.12.3.2~~8.11.3.2 Test

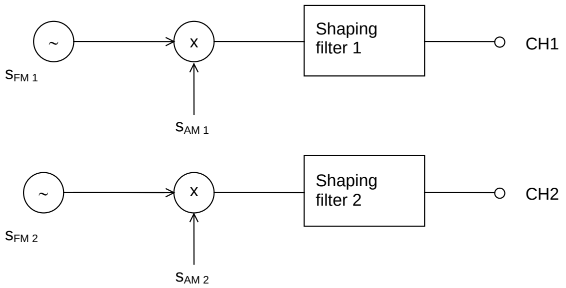
1) The test arrangement is in accordance with clause 6.

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

   The signals are fed simultaneously in the send and receive directions. The level in the send direction is –1.7 dBPa at the MRP (nominal level), the level in the receive direction is –16 dB_{m0} 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 echo signal is filtered by a comb filter using mid-frequencies and bandwidth according to the signal components of the signal in the 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 the 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. The echo attenuation is to be achieved for each individual frequency band from 200 Hz to 3 450 Hz according to the different categories.



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

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

The settings for the signals are as follows:

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

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

Figure 12 – Parameters of the two test signals for double talk measurement based on AM-FM modulated sine waves

~~8.12.4~~8.11.4 Robustness of double talk capability with far end PSAP noise

Commenté [gi6]: For future revisions

The intention of this test is to verify, that the implemented algorithms in the IVS do not erroneously hamper double talk capability, especially the transmission of near end speech, if the IVS receives ambient background noise from the PSAP side as downlink signal. This receive signal should not activate echo suppression unit in the IVS and hamper the transmission of near end voice.

~~8.12.4.1~~8.11.4.1 Requirements

for further study

~~8.12.4.2~~8.11.4.2 Test

for further study

~~8.13~~8.12 Background noise transmission

~~8.13.1~~8.12.1 Transparency of transmitted background noise after call setup (detection of “Silent Calls”)

The transmitted background noise in sending direction carries important acoustic information for the PSAP side (in addition to the transmitted MSD) to judge the relevance and importance of the automatically generated eCall, in particular if no person directly communicates from the vehicle with the PSAP operator. In this aspect the requirements for the transmission of background noise from the vehicle in uplink direction may significantly differ from the requirements for regular (convenience) hands-free telephony or manually generated eCalls.

.

Note: Background needs to be played back in the vehicle during this test in order to trigger the AGC control parameters accordingly.

~~8.13.1.1~~8.12.1.1 Requirements

for further study

~~8.13.1.2~~8.12.1.2 Test

for further study

~~8.13.2~~8.12.2 Speech quality in the presence of background noise

Commenté [gi7]: Change to intelligibility if procedure is available

~~8.13.2.1~~8.12.2.1 Requirements

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

Background noises scenario A1 :

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 scenario A2 :

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.

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.

~~8.13.2.2~~8.12.2.2 Test

1) The test arrangement is given in clause 6.

2)

~~8.13.3~~8.12.3 Silent call Performance

Annex A

Alternative method for determining the roundtrip delay

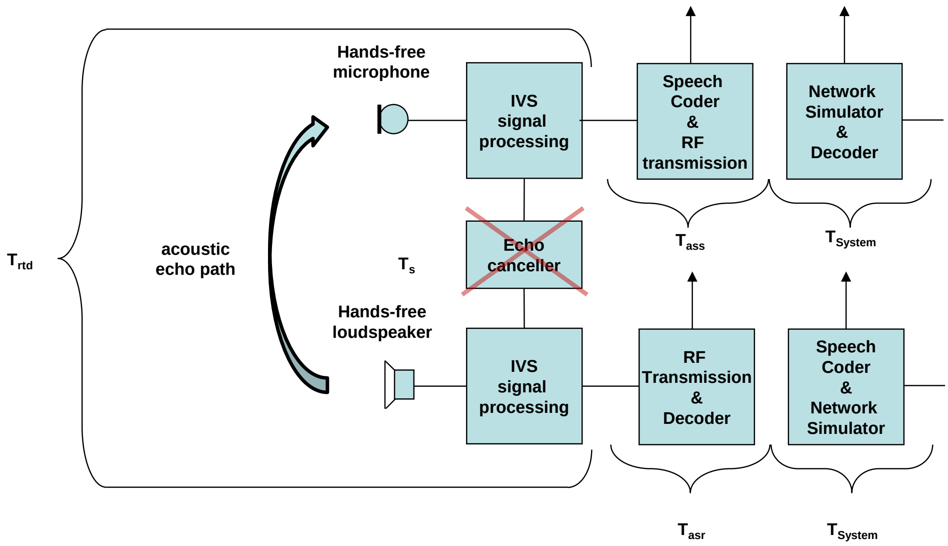
(This annex forms an integral part of this Recommendation.)

An alternative method to determine the roundtrip delay of an IVS is given below. Please not that this method can only be applied if the echo canceller of the IVS can be deactivated.

In case the IVS provides a test mode to disable echo cancellation and echo suppression signal processing, the round trip delay from POI (input of the reference speech coder of the system simulators) to the POI (output of the reference speech codec of the system simulator) can directly be measured:

$$T_{rtd} + t_{rtSystem}$$

NOTE 1 – The delay should be minimized.



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

1) For the measurement a composite source signal (CSS) in accordance with [ITU-T P.501] is used. The pseudorandom noise (PN) part of the CSS should be longer than the maximum expected delay. It is recommended to use a PN sequence of 16 k samples (with 48 kHz sampling rate). The test signal level is –16 dB_{m0} at the electrical interface (POI, input of system simulator).

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

2) The delay is determined by the cross-correlation analysis between the measured signal at the electrical access point (output of system simulator) 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.

SERIES OF ITU-T RECOMMENDATIONS

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

Printed in Switzerland
Geneva, 2011