International Telecommunication Union

# ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

# P.1140

(07/2022)

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

Communications involving vehicles

---

## Speech communication requirements for emergency calls originating from vehicles

Recommendation ITU-T P.1140

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 | P.10–P.19 |
| Voice terminal characteristics | P.30–P.39 |
| Reference systems | P.40–P.49 |
| Objective measuring apparatus | P.50–P.59 |
| Objective electro-acoustical measurements | P.60–P.69 |
| Measurements related to speech loudness | P.70–P.79 |
| Methods for objective and subjective assessment of speech quality | P.80–P.89 |
| Voice terminal characteristics | P.300–P.399 |
| Objective measuring apparatus | P.500–P.599 |
| Measurements related to speech loudness | P.700–P.709 |
| Methods for objective and subjective assessment of speech and video quality | P.800–P.899 |
| Audiovisual quality in multimedia services | P.900–P.999 |
| Transmission performance and QoS aspects of IP end-points | P.1000–P.1099 |
| **Communications involving vehicles** | **P.1100–P.1199** |
| Models and tools for quality assessment of streamed media | P.1200–P.1299 |
| Telemeeting assessment | P.1300–P.1399 |
| Statistical analysis, evaluation and reporting guidelines of quality measurements | P.1400–P.1499 |
| Methods for objective and subjective assessment of quality of services other than speech and video | P.1500–P.1599 |

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

## Recommendation ITU-T P.1140

# Speech communication requirements for emergency calls originating from vehicles


### Summary

Recommendation ITU-T P.1140 defines use cases, requirements and associated test methods for speech communication for emergency call communications originating from vehicles using a dedicated emergency call system covering built-in emergency call systems (manufacturer installed) as well as after-market emergency call kits.

This Recommendation contains an electronic attachment containing the set of freely-available test signals referred to within the Recommendation.



### History

| Edition | Recommendation | Approval | Study Group | Unique ID* |
|---|---|---|---|---|
| 1.0 | ITU-T P.1140 | 2015-06-29 | 12 | 11.1002/1000/12519 |
| 2.0 | ITU-T P.1140 | 2017-03-01 | 12 | 11.1002/1000/13177 |
| 3.0 | ITU-T P.1140 | 2022-07-29 | 12 | 11.1002/1000/15006 |

### Keywords

Emergency call, hands-free, motor vehicle.


* 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/11.1002/1000/11830-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/software copyrights, 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 appropriate ITU-T databases available via the ITU-T website at http://www.itu.int/ITU-T/ipr/.

## Table of Contents


| | | Page |
|---|---|---|
| 1 | Scope | 1 |
| 2 | References | 1 |
| 3 | Definitions | 3 |
| 3.1 | Terms defined elsewhere | 3 |
| 3.2 | Terms defined in this Recommendation | 3 |
| 4 | Abbreviations and acronyms | 4 |
| 5 | Conventions | 6 |
| 6 | Test set-up and preparation | 6 |
| 6.1 | Test arrangement | 6 |
| 6.2 | Background noise simulation in a car | 7 |
| 6.3 | Positioning of the emergency call IVS | 8 |
| 6.4 | Operation modes of emergency call IVS | 9 |
| 6.5 | Artificial mouth | 9 |
| 6.6 | Artificial ear | 9 |
| 6.7 | Influence of the transmission system | 9 |
| 6.8 | Calibration and equalization | 9 |
| 6.9 | System simulator settings | 10 |
| 6.10 | Environmental conditions | 10 |
| 7 | Test signals and test signal levels | 11 |
| 7.1 | Signals | 11 |
| 7.2 | Background noise signals | 12 |
| 8 | Measurement parameters and requirements for narrowband IVS terminals | 14 |
| 8.1 | Preparation measurements | 14 |
| 8.2 | Delay | 14 |
| 8.3 | Loudness ratings | 17 |
| 8.4 | Variation of RLR in the presence of background noise | 18 |
| 8.5 | Sensitivity frequency responses | 18 |
| 8.6 | Speech intelligibility during single talk | 21 |
| 8.7 | Idle channel noise | 21 |
| 8.8 | Echo performance | 23 |
| 8.9 | Switching characteristics | 25 |
| 8.10 | Double talk performance | 26 |
| 8.11 | Background noise transmission | 31 |
| 8.12 | Jitter buffer management | 32 |
| 9 | Measurement parameters and requirements for wideband IVS terminals | 35 |
| 9.1 | Preparation measurements | 35 |
| 9.2 | Delay | 35 |
| 9.3 | Loudness ratings | 38 |

9.4 Variation of RLR in the presence of background noise ................................ 39
9.5 Sensitivity frequency responses ...................................................................... 39
9.6 Speech intelligibility during single talk.......................................................... 42
9.7 Idle channel noise ........................................................................................... 42
9.8 Echo performance without background noise ............................................... 44
9.9 Switching characteristics ................................................................................. 47
9.10 Double talk performance ................................................................................. 48
9.11 Background noise transmission....................................................................... 53
9.12 Jitter buffer management ................................................................................. 53
Annex A – Alternative method for determining the roundtrip delay....................................... 55
Annex B – Void ....................................................................................................................... 56
Bibliography............................................................................................................................ 57

# Recommendation ITU-T P.1140

## Speech communication requirements for emergency calls originating from vehicles

### 1 Scope

This Recommendation[1] defines use cases, requirements, and associated test methods for the speech communication for emergency call communications originating from various types of vehicles (including e.g., busses, agricultural machines, construction vehicles, trucks) using a dedicated emergency call system installed in the cabin. These cover:

– built-in emergency call systems (manufacturer installed),

– after-market emergency call kits.

Test set-ups, measurement methods and requirements defined in this Recommendation are not applicable on powered two-wheelers without modifications, and are thus outside the scope of the Recommendation.

NOTE – An initial approach on modified test methods for powered two-wheelers was presented in [b-KettlerMotorcycles2017].

For test purposes, 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) during an 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 calls", 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 efficiency.

The methods, the analysis and the performance parameters described in this Recommendation are based, if applicable, on test signals and test procedures as defined in [ITU-T P.501], [ITU-T P.502], [ITU-T P.340], [ITU-T P.1100] and [ITU-T P.1110].

This Recommendation covers speech communication requirements and tests for emergency call systems in narrowband and wideband mode.

This Recommendation addresses the crash situation by simulating a post-crash situation as realistically as possible with respect to the impact of a post-crash situation on the acoustical environment. It is assumed that the performance of an in-vehicle system (IVS) after a crash is adequately covered by simulating a post-crash situation in an un-crashed car, provided that the system is still functional, and the acoustical elements (i.e., loudspeakers, microphones) remain at their original locations. Specific tests after a car crash are not within the scope of this Recommendation.

[1] This Recommendation contains an electronic attachment containing the set of freely-available test signals referred to within the Recommendation.

## 2 References

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


[ITU-T G.100.1] Recommendation ITU-T G.100.1 (2015), *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.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.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.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.570] Recommendation ITU-T P.570 (2018), *Artificial noise fields under laboratory conditions*. <http://www.itu.int/rec/T-REC-P.502>

[ITU-T P.581] Recommendation ITU-T P.581 (2022), *Use of head and torso simulator for hands-free and handset terminal testing*. <http://www.itu.int/rec/T-REC-P.581>

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

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

[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>

[ITU-T P.1110] Recommendation ITU-T P.1110 (in force), *Wideband hands-free communication in motor vehicles.*
<http://www.itu.int/rec/T-REC-P.1110>

[ETSI TS 103 224] ETSI TS 103 224 (2022), Speech and multimedia Transmission Quality (STQ); A sound field reproduction method for terminal testing including a background noise database.
<https://www.etsi.org/deliver/etsi_ts/103200_103299/103224/01.06.01_60/ts_103224v010601p.pdf>

[IEC 61260-1] IEC 61260-1:2014, *Electroacoustics – Octave-band and fractional-octave-band filters – Part 1: Specifications*.
<https://webstore.iec.ch/publication/5063>

# 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 communication efficiency**: Speed and accuracy of interactive communications between people over a telecommunication system.

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

**3.2.5 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 of [ITU-T P.58].

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

**3.2.7 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.

**3.2.8 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).

**3.2.9 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].

**3.2.10 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.

**3.2.11 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.

**3.2.12 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 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.

**3.2.15 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.

**3.2.16 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).

**3.2.17 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).

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

## 4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

| | |
|---|---|
| A/D | Analogue/Digital |
| AGC | Automatic Gain Control |
| AMR | Adaptive Multi-Rate |
| BGN | Background Noise |
| 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 |
| FR | Full Rate |
| HATS | Head And Torso Simulator |
| HATS-HFRP | Head And Torso Simulator – Hands-Free Reference Point |
| IMS | IP Multimedia Subsystem |
| IP | Internet Protocol |
| IVS | In-Vehicle System |
| $L_{S,min}$ | minimum activation level (send direction) |
| LTE | Long Term Evolution |
| MOS | Mean Opinion Score |
| MPNS | Multi-Point Noise Simulation |
| MRP | Mouth Reference Point |
| MSD | Minimum set of data |
| NR | Noise Reduction |
| PCM | Pulse Code Modulation |
| PN | Pseudorandom Noise |
| POI | Point Of Interconnection |
| PSAP | Public Safety Answering Point |
| QoS | Quality of Service |
| RF | Radio Frequency |
| RLR | Receive Loudness Rating |
| $RLR_{min}$ | minimum Receive Loudness Rating, achieved with background noise |
| SLR | Send Loudness Rating |
| S/N | Signal-to-Noise ratio |
| TCL | Terminal Coupling Loss |
| TCLw | weighted Terminal Coupling Loss |
| VoLTE | Voice over LTE |
| VoNR | Voice over New Radio |

## 5 Conventions

This Recommendation uses the following conventions:

| | |
|---|---|
| $A_{H,R}$ | 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

**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).

$T_{as}$ Access-specific delay (round trip)

$T_{asr}$ Access-specific delay in receive

$T_{ass}$ Access-specific delay in send

$T_{impr}$ Implementation dependent receive delay IVS

$T_{imps}$ Implementation dependent send delay IVS

$T_r$ Receive delay IVS

$T_{r,S}$ Built-up time (send direction)

$T_{rtd}$ Round trip delay IVS

$T_{rtdimp}$ Implementation dependent round trip delay IVS

$T_s$ Send delay IVS

**Vrms**: Voltage – root mean square.

## 6 Test set-up and preparation

### 6.1 Test arrangement

The acoustical interface for the in-vehicle system (IVS) is realized by using an artificial head or 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 IVS emergency call implementations are connected to a system simulator conforming to the required transmission standard with an implemented, calibrated audio interface.

The test signals are fed electrically to the system simulator or acoustically to the artificial head. The test arrangement is shown in Figure 6-1.

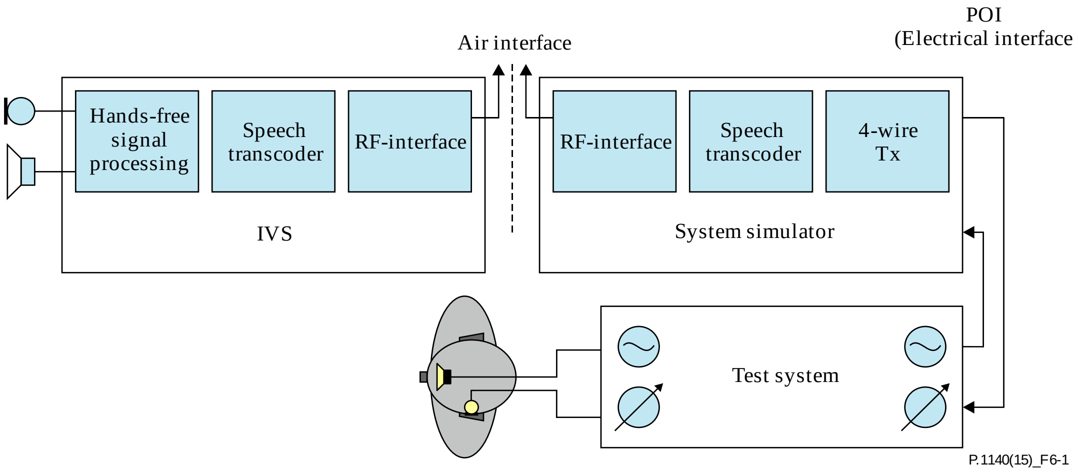



**Figure 6-1 – Test arrangement for emergency call IVS (see [ITU-T P.1100])**

## 6.2 Background noise simulation 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 6-2, the simulation arrangement is shown. More information on the test arrangement can be found in [ITU-T P.570] (equivalent to [b-ETSI ES 202 396-1]) as well as in clause 7.6 in [ETSI TS 103 224].

Some recordings for the background noise simulation shall be conducted in a real car as specified in clause 7.2. The noise signal is recorded by at least one measurement microphone positioned close to the hands-free microphone. If possible, the output signal of the hands-free microphones can be used directly. 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.

The loudspeaker arrangement is equalized and calibrated so that the power density spectra measured at the microphone positions are equal to the recorded ones. For the equalization procedure, the same measurement microphones or the hands-free microphones used for recording shall be 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.

For microphone arrays, distributed microphones or directional microphones and for noise types that are specifically recorded in a real car, the more sophisticated background noise simulation technology according to [ETSI TS 103 224] shall be used. Using this noise reproduction system, an accurate and automated equalization is performed and the generated sound-field is reproduced correctly in magnitude and phase up to a frequency of 2-3 kHz. A detailed description of the equalization procedure can be found in clause 7.6 of [ETSI TS 103 224], and shall be conducted between 50 Hz and 20 kHz.

For single channel microphone recordings, it is highly recommended to use the background noise simulation according to [ETSI TS 103 224] as well. For legacy reasons, also the manual equalization procedure according to [ITU-T P.570] can be used as an alternative noise simulation.

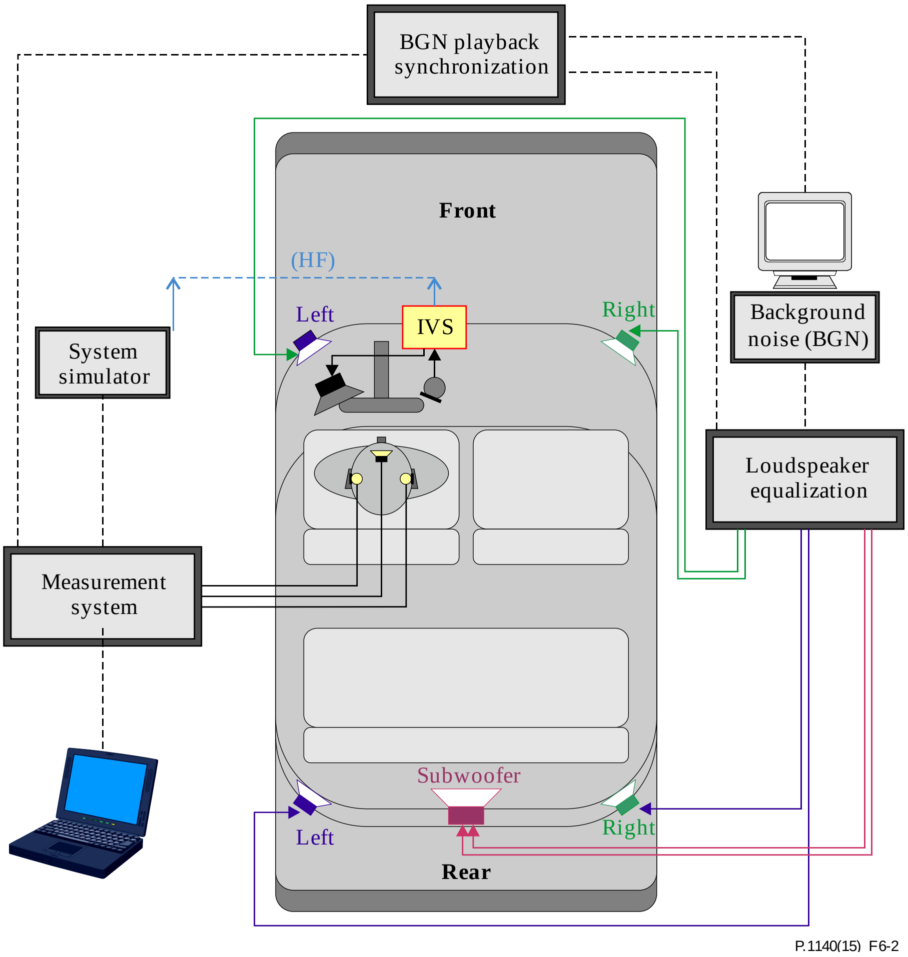



**Figure 6-2 – Test arrangement with background noise simulation**

## 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. 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 driver's 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-driver's and the two outer passengers' back seat (2nd row), if available, are used as described in the individual test cases.

The driver's 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. Detailed information can be found in [b-SAE J941]. 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.

### 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, etc.), 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. A detailed description for equalization at the MRP and level correction at the HATS-HFRP can be found in [ITU-T P.581].

### 6.6 Artificial ear

For measurements of IVS terminals in receive direction, 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]. The ear simulators according to [ITU-T P.57] shall be used and reported.

### 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 transmission system or the system simulator, reference measurements of the hands-free unit or measurements without acoustical components should be made to document this behaviour.

### 6.8 Calibration and equalization

The following preparation shall be completed before running the tests:

#### 6.8.1 Calibration

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

#### 6.8.2 Equalization

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

### 6.8.3 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

### 6.9.1 General

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

### 6.9.2 Circuit-switched networks

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

For wideband measurements according to the UMTS standard, the AMR-WB codec shall be used with a bitrate of 12.65 kbit/s.

### 6.9.3 Packet-switched networks

For measurements with AMR or AMR-WB codec via VoLTE or VoNR, the same bitrates as for circuit-switched networks apply.

For narrowband measurements with EVS-NB codec, a bitrate of 13.2 kbit/s shall be used.

For wideband measurements with EVS-WB codec, a bitrate of 13.2 kbit/s shall be used.

For wideband measurements with EVS-SWB codec, a bitrate of 24.4 kbit/s shall be used.

NOTE – For wideband measurements in packet-switched networks (VoLTE/VoNR), EVS-SWB codec can be used instead of AMR-WB or EVS-WB codec. The increased audio bandwidth can provide improved quality and intelligibility. However, the requirements for wideband terminals are defined independent of the codec, i.e., any increased complexity for supporting EVS-SWB and higher bandwidth in the IVS is not considered.

## 6.10 Environmental conditions

Unless specified otherwise, the idle noise level in the vehicle at all measurement locations shall be less than −54 dBPa(A) in conjunction with NC40 (see Figure 6-3).

For the specified tests, it is desirable to have an idle noise level of less than −74 dBPa(A) in conjunction with NC20.

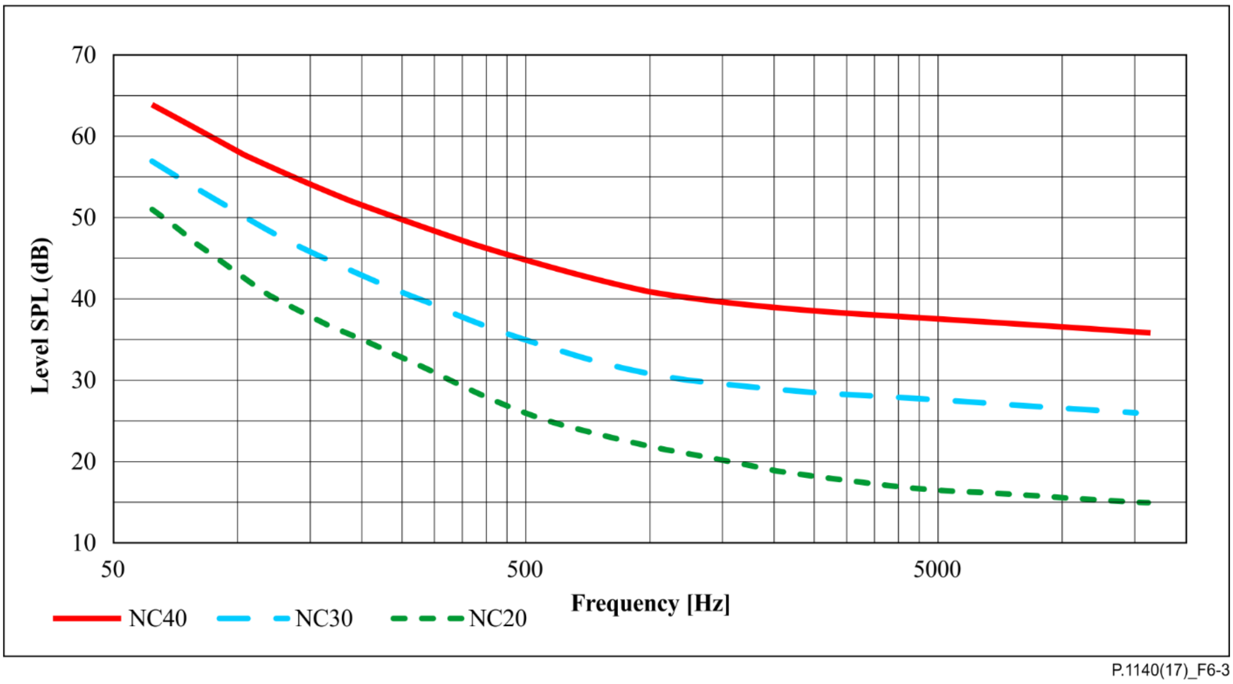



**Figure 6-3 – NC-criteria for a test environment**

### 6.10.1 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 or pressure-field measurement microphone is positioned in the test room inside 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.

## 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.

For wideband 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 50 Hz and 8 kHz using bandpass filtering providing 24 dB/octave. In the send direction, the test signals are used without band-limitation.

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

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 dBm0 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 1 – If different network signal levels are used in tests, this is stated in the individual description of the test.

NOTE 2 – In the present Recommendation, there are currently no test methods specified that are conducted with background noise in send direction, which would require an additional gain in speech level due to the so-called *Lombard effect*, as described in e.g., [ITU-T P.1100], [ITU-T P.1110], [b-ITU-T P.1150].

### 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 are defined in clause 6.2. 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 scenarios accordingly:

– **automatically generated emergency calls (A)**:

  **simulated emergency call noise scenario (A1)**, e.g., stationary car (parking on a highway parking place), engine off, all 4 windows open, passing vehicles, this file is annexed to the Recommendation;

  **simulated emergency 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, this file is annexed to the Recommendation;

  **spectrally adapted stationary noise to reproduce spectral content of scenario A1 (A3)**: white Gaussian noise filtered by the average spectrum derived from scenario A1, this file is annexed to the Recommendation;

  **simulated emergency call noise scenario (A4)**, e.g., stationary car (parking on a country road parking place), engine off, all 4 windows open, single passing vehicles, this file is annexed to the Recommendation;

  For all noise scenarios of type A: In case the IVS provides multiple microphones, the microphone relevant for the corresponding test position of HATS shall be used for equalization;

– **manually generated emergency calls (B):** Constant driving conditions simulating fixed driving speed (e.g., 130 km/h). For the reproduction of driving noise, recordings shall be conducted at all IVS microphone positions.

### 7.2.1 Recording of background noise

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

NOTE – 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 clause 7.6 of [ETSI TS 103 224] 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.

For automatically generated emergency calls the background noises included in this Recommendation shall be used.

The type of background noise used shall be noted in the test report.

### 7.2.2 Playback of the recorded background noise

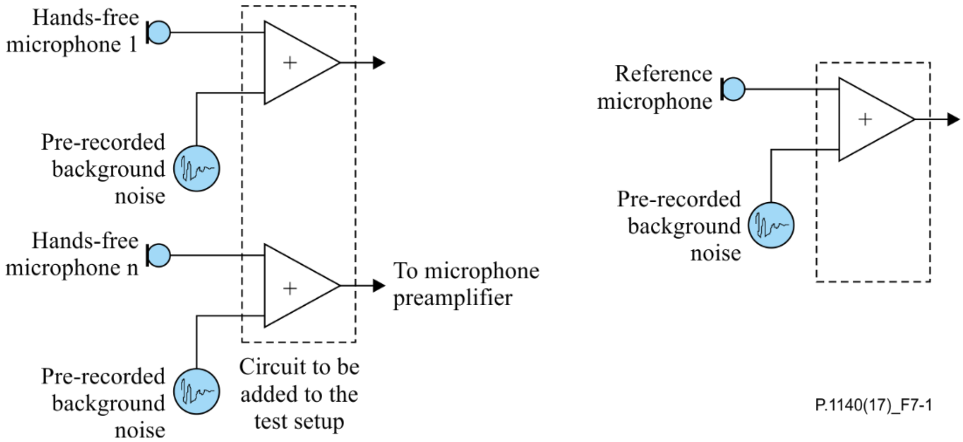
Two ways of background noise playback are recommended:

1) The test laboratory employs a 4-loudspeaker arrangement for acoustic background noise reproduction in the car cabin as described in clause 6.2. Typically, two loudspeakers each 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.

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 shall be provided, see Figure 7-1. The test laboratory shall ensure the right calibration of the two signals.

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

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 7-1 – Set-up for analogue electrical insertion of the pre-recorded background noise signal**

# 8 Measurement parameters and requirements for narrowband IVS terminals

As a general rule for all tests it should be noted that it is always the responsibility of the test lab to verify that measurement results are not influenced by second-order effects such as additional noise falsifying the result of the individual measurement, for example.

## 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 $T_{as}$ and the IVS implementation dependant 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 IVS implementation dependant delay $T_{rtdimp}$ consists of:

– The IVS signal processing in send and receive.

– The air-paths from mouth to microphone and from the loudspeakers to the ear.

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

– The IVS delay in the send (uplink) direction Ts 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 IVS delay in the receive (downlink) direction Tr 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.

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

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 Tas from these specifications, 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 [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 $T_{rtdimp}$ should not exceed 70 ms.

### 8.2.2 Delay in the send direction

#### 8.2.2.1 Test

The delay in the 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:

$$T_s + T_{System}$$

NOTE 1 – The delay should be minimized.

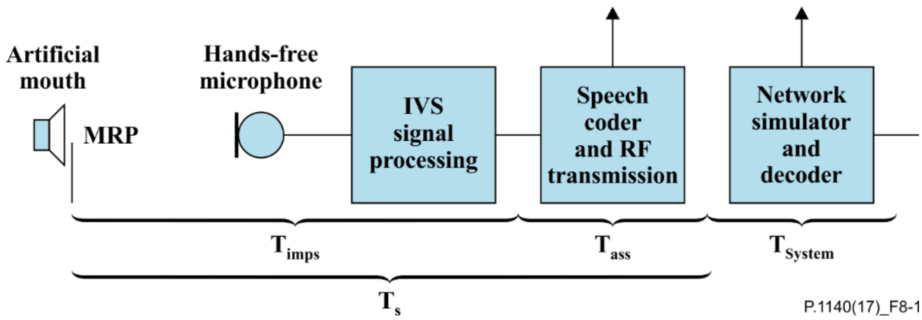



**Figure 8-1 – Different blocks contributing to the delay in the send direction**

The system delay $T_{System}$ is dependent on the transmission method used and the network simulator. This delay must be known and subtracted from the total measured delay.

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 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). 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 6. The HATS is positioned on the driver's 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 and access network.

3) The maximum of the cross-correlation function is used for the determination of the delay.

4) For packet-switched connections, the measurement shall be repeated at least for five individual calls. The maximum of the obtained delays is considered as the overall delay in send. The delay values of all individual calls shall be provided in the test report. In case the reference client of the test equipment does not compensate for the effect of speech frame synchronization, the maximum uncertainty quantity of 20 ms shall be subtracted from all

delays obtained in send direction (see Note 3). If applicable, the correction shall be documented in the test report.

NOTE 2 – For packet-switched connections via VoLTE or VoNR, a variability of up to 20 ms may be expected between different calls due to the synchronization between the speech frame processing in the IVS in send and the bits of the speech frames at the IVS antenna. This synchronization is attributed to the IVS send delay according to the definition of the IVS delay reference points. Hence, the maximum value of delays obtained from at least five individual calls is reported as the overall IVS delay in send direction.

NOTE 3 – A further variability of up to 20 ms may be expected between different calls due to the synchronization between the speech frames at the IVS antenna and the speech frame processing in the receiving reference client of the test system. Hence, if the reference client of the test equipment does not compensate for the effect of the speech frame synchronization (as specified by the manufacturer of the reference client), this maximum uncertainty is subtracted from the measured maximum delay to obtain the true IVS delay in send.

### 8.2.3 Delay in the receive direction

#### 8.2.3.1 Test

The delay in the 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:

$$T_r + T_{System}$$

NOTE 1 – The delay should be minimized.

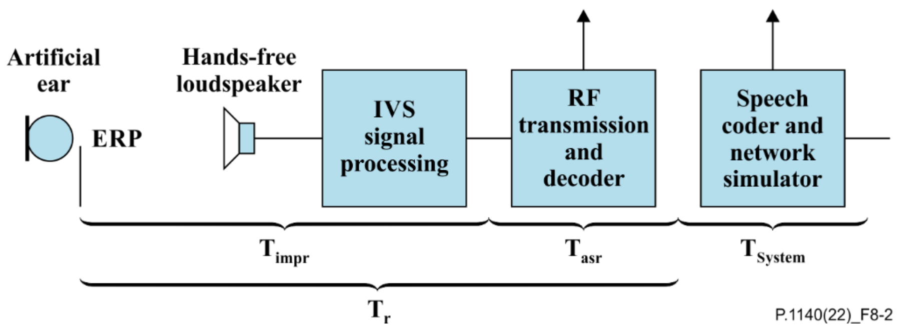



**Figure 8-2 – Different blocks contributing to the delay in the receive direction**

The system delay $T_{System}$ depends on the transmission system and on the network simulator used. This delay must be known and subtracted from the total measured delay.

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 shall be longer than the maximum expected delay. It is recommended to use a PN sequence of 16 k samples (with a 48 kHz sampling rate). The test signal level is −16 dBm0 at the electrical interface (POI).

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

2) The test arrangement is in accordance with clause 6. For packet-switched connections, no jitter or packet loss shall be introduced in the transmission. The HATS is positioned on the driver's seat. For the measurement the artificial head is 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 and access network.

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

5) For packet-switched connections, the measurement shall be repeated at least for five individual calls. The maximum of the obtained delays is considered as the overall delay in receive. The delay values of all individual calls shall be provided in the test report.

NOTE 2 – For packet-switched connections via VoLTE or VoNR, a variability of up to 20 ms may be expected between different calls due to the synchronization between the bits of the speech frames at the IVS antenna and the speech frame processing in the IVS. This synchronization is attributed to the IVS receive delay according to the definition of the IVS delay reference points. Hence, the maximum value of delays obtained from at least five individual calls is reported as the overall IVS delay in receive direction.

## 8.3 Loudness ratings

### 8.3.1 Requirements

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

– SLR ≤ 22 dB for all positions in the car (the HATS positioned either at the driver's position, the co-driver's and the two outer passengers' back seats (2nd row), if available);

– A SLR = 13 dB ± 4 dB for HATS positioned at the driver's position is recommended.

NOTE – In general it is recommended to maintain an SLR of 13 dB at all positions in the car to achieve a sufficient intelligibility under all conditions (soft or weak talkers, more than 2 seat rows). This can be achieved e.g., by distributed microphones or properly designed automatic gain control (AGC) in send.

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

– RLR ≤ 10 dB for all positions in the car (the HATS positioned either at the driver's position, the co-driver's and the two outer passengers' back seats (2nd row), if available);

– A RLR = −3 dB ± 4 dB for HATS positioned on the driver's seat is recommended.

It is required to use background noise controlled AGC in the receiving direction. Performance requirements for this feature are addressed in clause 8.4.

### 8.3.2 Test loudness rating in the 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 6. Tests are carried out with the HATS positioned on the driver's 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 the 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 dBm0, measured at the electrical reference point and averaged over the complete test signal sequence.

2) The test arrangement is according to clause 6. Tests are carried out with the HATS positioned on the driver's seat and additionally carried out with the HATS on the back passengers' seats ($2^{nd}$ row). The artificial head is 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.

## 8.4 Variation of RLR in the presence of background noise

The intention of this test is the verification of the amplification range introduced by the AGC implementation in the presence of noise in the receiving direction. The RLR is determined in silent conditions (see test 8.3.3) and additionally in presence of background noise. The level of background noise needs to be sufficiently high and representative for a scenario as defined in A1.

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 achieve a RLR ≤ −8 dB within 4 seconds after the start of playback of background noise scenario A3 in the vehicle cabin. The measured RLR is called $RLR_{A3}$.

NOTE – The AGC should be designed to allow a SNR of ≥ 6 dB for all signal and noise conditions. The AGC control range should allow $S/N \geq 6$ 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.

### 8.4.2 Test

1) The RLR is determined as described in clause 8.3.3 (Test loudness rating in the 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 the reproducibility of the noise audio recording.

3) The background noise playback and recording is started 4 seconds before 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 the 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 $RLR_{A3}$.

5) $RLR_{A3}$ is determined as described in clause 8.3.3.

## 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 8-1, the mask is drawn by straight lines between the breaking points in Table 8-1 on a logarithmic (frequency) – linear (dB sensitivity) scale.

**Table 8-1 – Tolerance mask for the narrowband 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 | 4 dB | −∞ 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. | | |

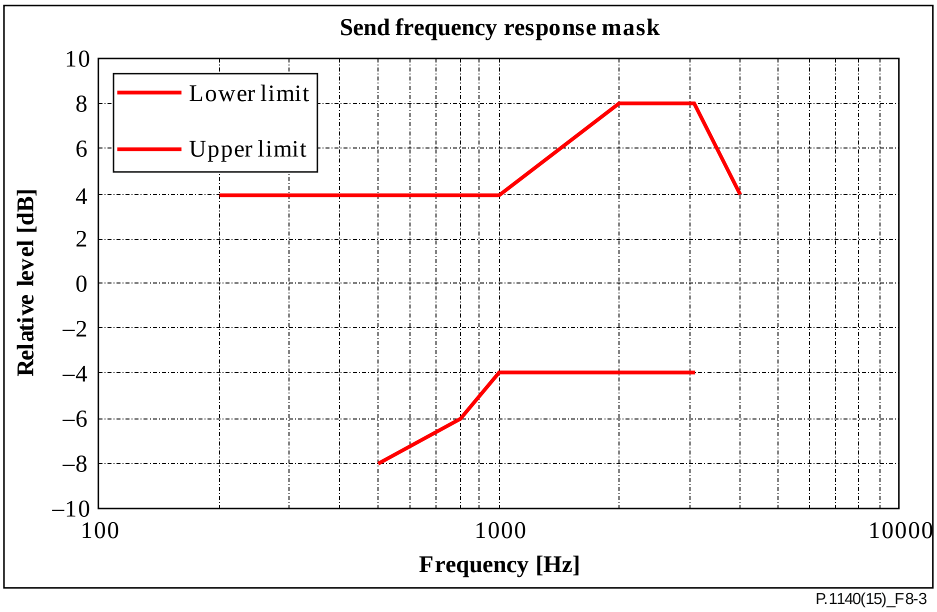



**Figure 8-3 – Narrowband send frequency response mask (Fig. is informative)**

### 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 61260-1] 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 8-2, the mask is drawn by straight lines between the breaking points in Table 8-2 on a logarithmic (frequency) – linear (dB sensitivity) scale.

**Table 8-2 – Tolerance mask for the narrowband receive sensitivity frequency response**

| Frequency [Hz] | Upper limit | Lower limit |
|---|---|---|
| 200 | 6 dB | −∞ dB |
| 250 | 6 dB | −∞ dB |
| 315 | 6 dB | −∞ dB |
| 400<br>630<br>1000<br>1 500 | 6 dB<br>6 dB<br>6 dB<br>9 dB | −9 dB<br>−6 dB<br>−6 dB<br>−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. | | |

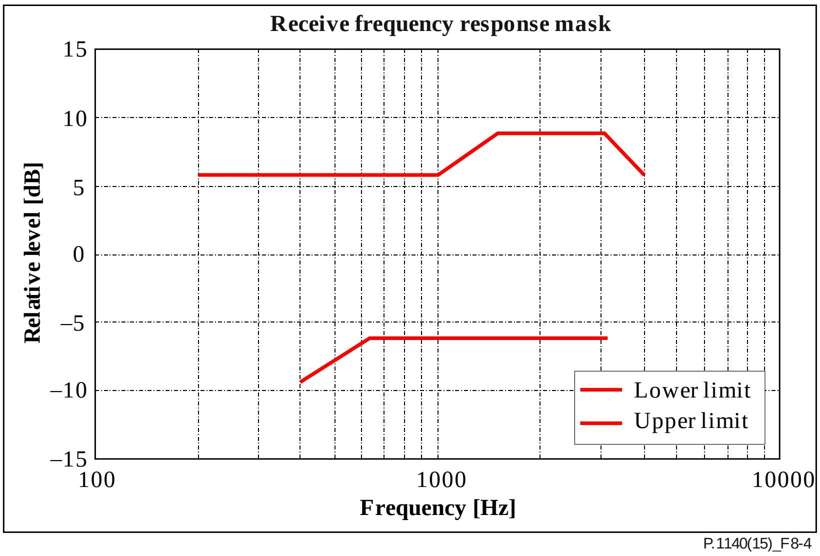



**Figure 8-4 – Narrowband receive frequency response mask (Fig. is informative)**

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 at a 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 Recommendation uses the diffuse-field as a reference.

#### 8.5.2.2 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) 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 intelligibility during single talk

### 8.6.1 One-way speech intelligibility in send

Speech intelligibility is one of the most important parameter for emergency call systems. The intention of this test is to verify, that a sufficient level of speech intelligibility can be achieved. Currently no objective method is available. As soon as a validated method is available applicable for emergency call systems this section should be added.

*For further study.*

### 8.6.2 Speech intelligibility in receive

Speech intelligibility is one of the most important parameter for emergency call systems. The intention of this test is to verify, that a sufficient level of speech intelligibility can be achieved. Currently no objective method is available. As soon as a validated method is available that is applicable for emergency call systems this section should be added.

*For further study.*

## 8.7 Idle channel noise

All tests are conducted with average RF-signal power settings. It is recommended in addition to check the requirement 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) shall not be exceeded.

### 8.7.1 Idle channel noise in the send direction

#### 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 dBm0(P).

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 these are considered to be inaudible due to the very low noise floor produced by the hands-free terminal on average, it is the responsibility of the test lab to demonstrate the desired performance subjectively.

#### 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 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 −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 8.79 Hz (determined using FFT 8 k samples/48 kHz sampling rate with a Hanning window or equivalent). The idle channel noise spectrum is stated in dB(P). 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.

### 8.7.2 Idle channel noise in the receive direction

#### 8.7.2.1 Requirements

The requirements for the maximum noise produced by the IVS in the case where 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 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, for example, 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 these are 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.

#### 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 four composite source signals according to [ITU-T P.501]. The level of the activation level is adjusted to −16 dBm0, measured at the electrical reference point. The level of the activation signal is averaged over the complete duration of the activation signal.

2) The test arrangement is according to clause 6. For the measurement of the IVS the artificial head is diffuse-field equalized according to [ITU-T P.581] (see also clause 6.8.2). 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 reverberation 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.

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 up to 3.4 kHz.

## 8.8 Echo performance

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 maximum background noise level of −64 dBPa(A) shall not be exceeded.

### 8.8.1 Terminal coupling loss (TCLw)

#### 8.8.1.1 Requirements

The TCLw in quiet environments shall be at least 46 dB for all settings of the volume control. The implemented echo control mechanism should provide sufficient echo loss for all typical environments and typical impulse responses.

NOTE 1 – A TCLw of ≥ 50 dB is recommended as a performance objective. 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 2 – 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.

#### 8.8.1.2 Test

1) All tests are conducted in the car cabin; the test arrangement is described in clause 6. 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 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) TCLw is calculated according to clause B.4 of [ITU-T G.122], (trapezoidal rule). For the calculation, the average test signal level at each frequency band is referred to the averaged measured echo signal 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.

### 8.8.2 Echo level versus time

#### 8.8.2.1 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 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 – 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.

#### 8.8.2.2 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 dBm0 as well as an average level of −25 dBm0. The echo signal is analysed during a period of at least 2.8 s, which represents 8 periods of the CS signal. The integration time for the level analysis shall be 35 ms, the analysis is referred to the level analysis of the reference signal. In addition, the test is repeated with 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 dBm0. 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).

### 8.8.3 Echo performance with time variant echo path and speech

#### 8.8.3.1 Requirements

This test is intended 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 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-drivers'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 that is quietly moving one arm close to the microphone during the measurement.

The type of echo path variation chosen shall be reported.

#### 8.8.3.2 Test

1) Before conducting the test, the echo canceller shall be fully converged.
2) The test arrangement is according to clause 6. 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 dBm0. 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 reference.
3) Now a second measurement is started introducing the time-variant echo path.
4) The test signal is 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 dBm0. 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 echo path is determined.
6) The measurement result is displayed as attenuation vs. time. The exact synchronization between the two measured signals has to be guaranteed.

## 8.9 Switching characteristics

### 8.9.1 Activation in the 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).

#### 8.9.1.1 Requirements

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

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

#### 8.9.1.2 Test

The structure of the test signal is shown in Figure 8-5. The test signal consists of CSS components according to [ITU-T P.501] with increasing level for each CSS burst.

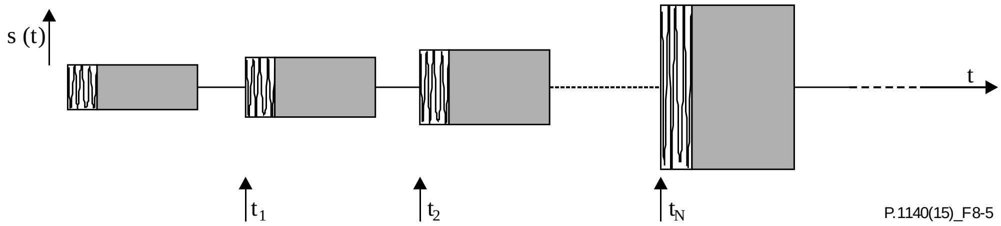



**Figure 8-5 – Test signal to determine the minimum activation level and the built-up time**

The settings of the test signal are as follows:

**Table 8-3 – 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 dB |
| NOTE – The level of the active signal part corresponds to an average level of −24.7 dBPa at the MRP for the CSS according to [ITU-T 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.

### 8.10 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 the send direction during double talk $A_{H,S,dt}$.

– Attenuation range in the receive direction during double talk $A_{H,R,dt}$.

– Echo attenuation during double talk.

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

#### 8.10.1.1 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 8-4.

**Table 8-4 – 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 shall provide a double talk capability of type 2b or better in the sending direction. The requirements apply for the nominal (default) setting of the receive volume control.

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 8-4 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.

#### 8.10.1.2 Test

The test signal to determine the attenuation range during double talk is shown in Figure 8-6. 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.

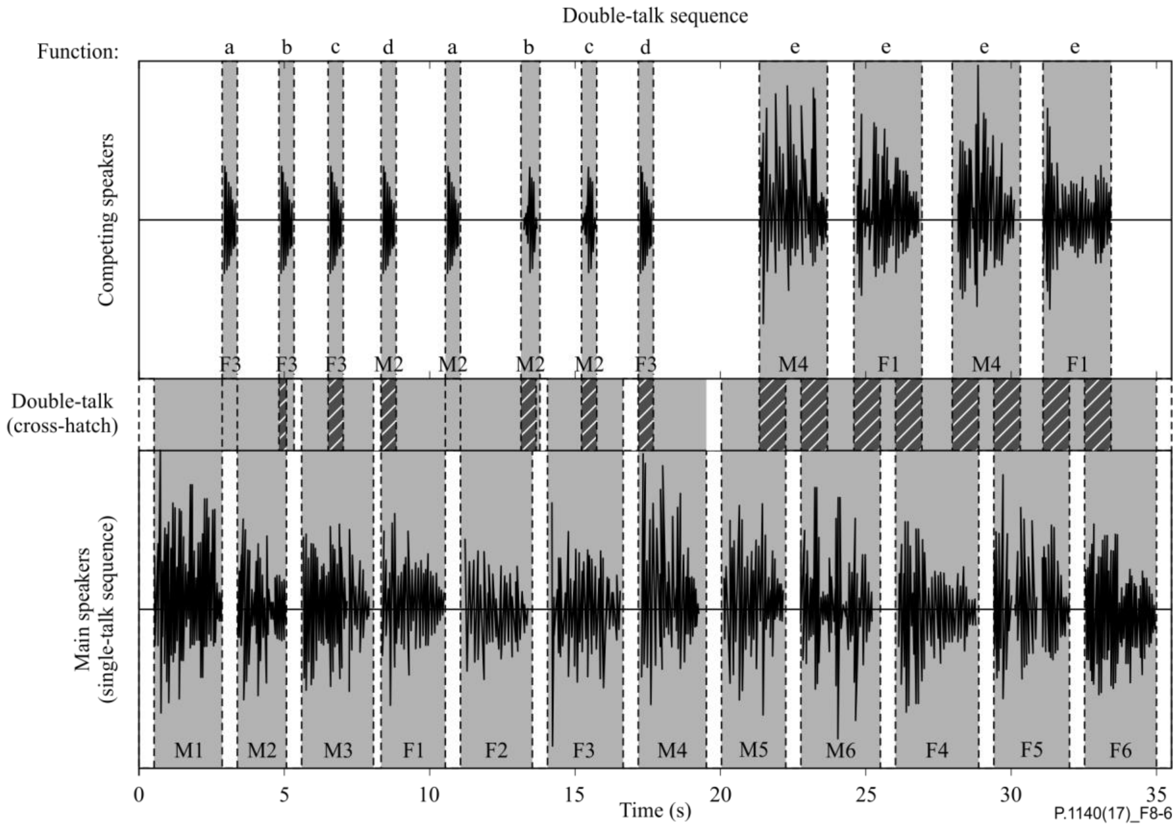



**Figure 8-6 – Double talk test sequence with overlapping speech sequences in the send and receive directions**

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 8-5 – Level of the double talk sequences**

| | **Receive direction** | **Send direction** |
|---|---|---|
| Average signal level | −16 dBm0 | −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 dBm0 is applied to the electrical reference point.

2) When determining the attenuation range in the 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 [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.

### 8.10.2 Attenuation range in the 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.

#### 8.10.2.1 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 8-6.

**Table 8-6 – 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 shall provide a double talk capability of type 2b or better in the receiving direction. The requirements apply for nominal setting of the receive volume control.

In general, Table 8-6 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.

#### 8.10.2.2 Test

The test signal to determine the attenuation range during double talk is shown in Figure 8-7. 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.

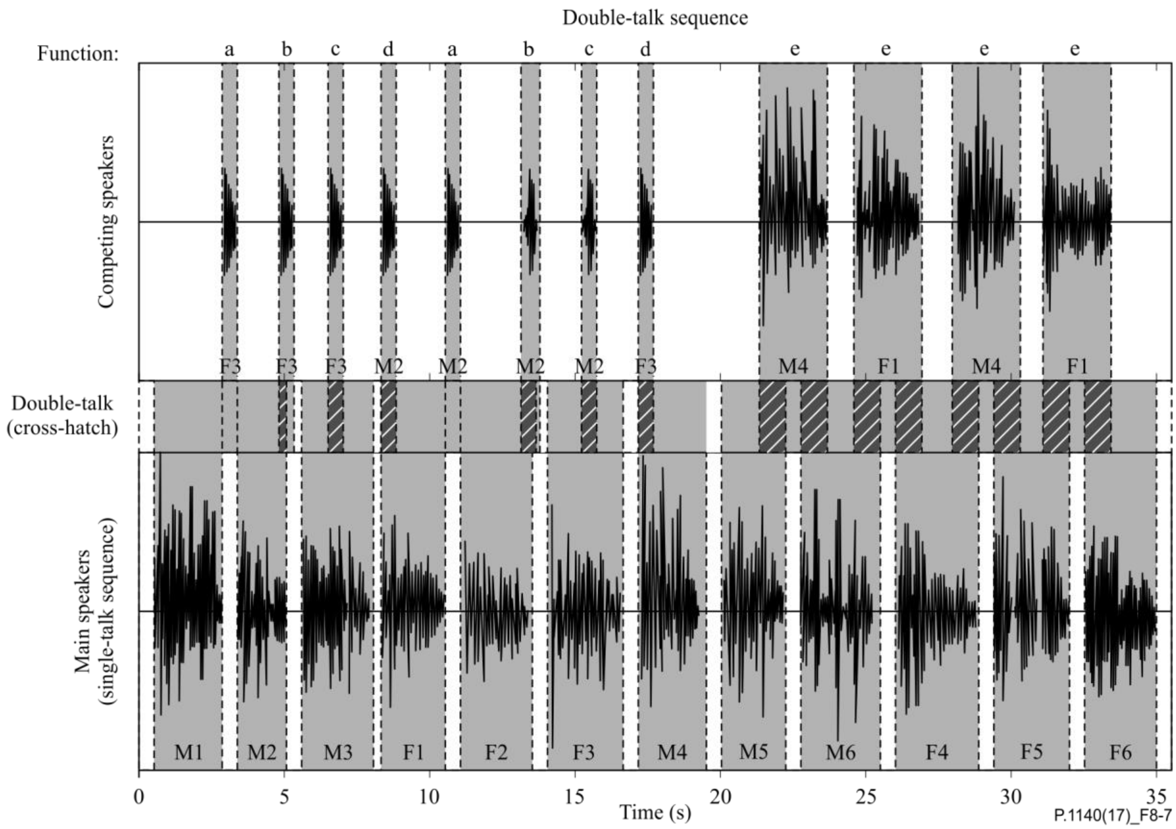



**Figure 8-7 – Double talk test sequence with overlapping speech sequences in the receive and send directions**

The settings for the test signals are as follows:

**Table 8-7 – Level of the double talk sequences**

| | **Receive direction** | **Send direction** |
|---|---|---|
| Average signal level | −16 dBm0 | −1.7 dBPa |

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

2) When determining the attenuation range in the 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.

### 8.10.3 Detection of echo components during double talk

#### 8.10.3.1 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 8-8 are applicable (more information can be found in Annex A of [ITU-T P.340]).

**Table 8-8 – 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 |

The IVS shall provide an echo-loss during double talk capability of type 2c or better. The requirements apply for a nominal setting of the receive volume control.

#### 8.10.3.2 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 signal is described in [ITU-T P.501]. The signal settings used are shown in Table 8-9. 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 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 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 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 8-8. The echo attenuation is to be achieved for each individual frequency band from 200 Hz to 3 450 Hz according to the different categories.

**Table 8-9 – Parameters of the two test signals for double talk measurement based on AM-FM modulated sine waves**

| Send direction | | Receive direction | |
|---|---|---|---|
| $f_0^{(1)}$ (Hz) | $\pm\Delta f^{(1)}$ (Hz) | $f_0^{(2)}$(Hz) | $\pm\Delta f^{(2)}$ (Hz) |
| 250 | ±5 | 270 | ±5 |
| 500 | ±10 | 540 | ±10 |
| 750 | ±15 | 810 | ±15 |
| 1000 | ±20 | 1080 | ±20 |
| 1250 | ±25 | 1350 | ±25 |
| 1500 | ±30 | 1620 | ±30 |
| 1750 | ±35 | 1890 | ±35 |
| 2000 | ±40 | 2160 | ±35 |
| 2250 | ±40 | 2400 | ±35 |
| 2500 | ±40 | 2650 | ±35 |
| 2750 | ±40 | 2900 | ±35 |
| 3000 | ±40 | 3150 | ±35 |
| 3250 | ±40 | 3400 | ±35 |
| 3500 | ±40 | 3650 | ±35 |
| 3750 | ±40 | 3900 | ±35 |
| | | | |

### 8.10.4 Robustness of double talk capability with far end background noise at the PSAP

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 the echo suppression unit in the IVS and hamper the transmission of the near end voice.

*For further study.*

## 8.11 Background noise transmission

### 8.11.1 Silent call performance

The intention of this test is to verify, that in the absence of speech the background noise captured by the IVS microphone is transmitted mostly transparently without artefacts to preserve the information contained in the background noise.

#### 8.11.1.1 Requirements

The level variation vs. time of the transmitted noise from the emergency call noise scenario (A4) shall be less than 10 dB when referred to a signal picked up close to the IVS microphone.

#### 8.11.1.2 Test

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

2) The test signal is the simulated emergency call noise scenario (A4) with a duration of 20 s. The test signal level is the original signal level.

3) The transmitted background noise signal picked up at the POI is analysed during between 11 s and 19 s as a level versus time analysis. The integration time for the level analysis shall be 125 ms. The analysis is referred to the level analysis of the reference signal. The reference signal is the background noise signal picked up by an omnidirectional measurement microphone positioned close to the IVS microphone. The reference signal is filtered using the modified IRS send filter as defined in [ITU-T P.830].

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

NOTE – More advanced techniques are under study.

### 8.11.2 Speech intelligibility in the presence of background noise

The intention of this test is to verify, that in the presence of background noise a sufficient level of speech intelligibility can be achieved. Currently no objective method is available. As soon as a validated method is available that is applicable for emergency call systems this section should be added.

*For further study.*

## 8.12 Jitter buffer management

### 8.12.1 Overview

For measurements conducted in packet-switched network connection with simulated packet arrival time variations and packet loss, the intention of the following test is to ensure that the jitter buffer of the IVS does not severely impact speech communication. To ensure a minimum conversational quality, the increase of delay due to higher and more variant packet arrival time should be minimized and the loudspeaker output in receive should still be comprehensible, i.e., clipping, attenuation or insufficient packet loss concealment of the received signal should be avoided.

### 8.12.2 Requirements

Requirements are not applicable for circuit-switched network connections.

For measurements conducted in packet-switched network connection with simulated packet arrival time variations and packet loss, the sum of the IVS delays in send and receive directions ($T_S + T_R$) shall be less than or equal to the delay requirements given in Table 8-10.

**Table 8-10 – IVS delay requirements for LTE/NR access**

| Test condition | Delay and loss profile | Requirement for maximum delay | Comment |
|---|---|---|---|
| 0 | Error and jitter free condition | $T_S + T_R \leq 170$ ms | Same as in clause 8.2 |
| 1 | dly_profile_20msDRX_10pct_BLER_e2e | $T_S + T_{R\text{-}CCVA} \leq 170$ ms | |
| 2 | dly_profile_40msDRX_10pct_BLER_e2e | $T_S + T_{R\text{-}CCVA} \leq 210$ ms | |
| 3 | dly_profile_40msDRX_22pct_BLER_e2e | $T_S + T_{R\text{-}CCVA} \leq 210$ ms | |

To ensure that the jitter buffer management does not significantly impact the loudspeaker output in receive direction, the loss in speech quality per profile shall not exceed the requirements given in Table 8-11.

NOTE – The purpose of this test is to provide a relative comparison of the objective speech quality between the reference and test conditions. This test is not to be construed as a method to evaluate the absolute objective speech quality of the device.

**Table 8-11 – IVS requirements on speech quality decrease for LTE/NR access**

| Test Condition | Delay and loss profile | Requirement for speech quality loss | Comment |
|---|---|---|---|
| 0 | Error and jitter free condition | - | Reference for actual test profiles |
| 1 | dly_profile_20msDRX_10pct_BLER_e2e | ≤ 0.3 | |
| 2 | dly_profile_40msDRX_10pct_BLER_e2e | ≤ 0.3 | |
| 3 | dly_profile_40msDRX_22pct_BLER_e2e | ≤ 1.0 | |

### 8.12.3 Test signal

The test signal consists of three repeats of the CSS according to [ITU-T P.501], followed by a longer speech sequence of 160 s. During the first two CSS bursts, the terminal can adapt its jitter buffer. The third CSS is used for measuring the delay in constant-delay condition, while the speech signal is used for delay and quality measurements under packet impairment conditions.

For all repeated CSS, the pseudo random noise (pn)-part shall be longer than the maximum expected delay. It is recommended to use a pn sequence of 32K samples (for sampling rate of 48 kHz). The test signal level is -16 dBm0 measured at the digital reference point or the equivalent analogue point.

For the speech signal, eight English test sentences according to clause C.2.3 of [ITU-T P.501], normalized to an active speech level of −16 dBm0, are used (two male, two female speakers). The sequence is composed in such a way that all sentences are centred within a 4.0 s time window, which results in a length of 32.0 s. The sequence is repeated five times, resulting in an overall length of 160.0 s.

### 8.12.4 Delay and loss profiles

The delay profiles according to Annex E of [3GPP TS 26.132] shall be used, which simulate the network behaviour of typical VoLTE/VoNR connections. The delay profiles for test condition 1 and 2 are theoretically constructed to simulate a semi-persistent scheduling transmission scheme with DRX enabled and initial BLER in sending and receiving directions of 10%, with ±3 ms of EPC jitter. The delay profile for test condition 3 is theoretically constructed to simulate a semi-persistent scheduling transmission scheme with DRX enabled and initial BLER in sending and receiving directions of 22%, with ±6 ms of EPC jitter.

Separate calls shall be established for each packet impairment condition. Delay profiles shall be applied between the reference client and system simulator.

Constant delay $T_c$, which corresponds to the minimum delay of each profile (i.e., the compensation value for the profile) shall be inserted at the beginning of the different delay/loss profiles, to avoid unnecessary delay jumps between the two measurement phases and realistic conditions for the second measurement test phase. The duration of constant delay corresponds to the duration of the CSS part of the test signal described in clause 8.12.3. Note that the profile files that are provided as an electronic attachment to [3GPP TS 26.132] do not include this initial constant delay condition.

### 8.12.5 Test

The test arrangement is in accordance with clause 6. An additional acoustic recording is conducted simultaneously with a measurement microphone positioned as close as possible to the loudspeaker of the IVS. Since the model according to [ITU-T P.863] is used for measuring the speech quality loss in receive direction, adequate signal-to-noise ratio and reduced impact of the acoustic environment are required for a sufficient prediction accuracy. Both aspects are considered by the additional measurement microphone recording.

The test signal according to clause 8.12.3 is played back in receive direction. The start of the delay profile shall be synchronized with the start of the downlink speech material reproduction to ensure a repeatable application of impairments relative to the test speech signal.

Measurements and analyses for delay and speech quality loss shall be repeated for all three delay and loss profiles described in clause 8.12.4, including error and jitter-free conditions, which is used for reference in the following analyses.

*Analysis for delay:*

1) The signal of the ear closest to the loudspeaker of the IVS shall be used for the analysis and noted in the test report.

2) In a first analysis step, the delay between the electrical access point of the test equipment and DRP in constant-delay condition, $T_{TEAP\text{-}DRP\text{-}constant}$, is calculated in the same way as described in clause 8.2.3.1 but using only the third CSS burst of the recording. The profile-dependent constant delay $T_c$ of is subtracted from $T_{TEAP\text{-}DRP\text{-}constant}$ to obtain $T_{R\text{-}constant}$.

3) The delay $T_{TEAP\text{-}DRP}(t) = T_{R\text{-}jitter}(t) + T_{System}$ is measured continuously for the speech signal during the inclusion of packet delay and loss profile. For the delay calculation, a cross-correlation with a rectangular window length of 4 s, centred at each sentence of the stimulus file, is used. The process is repeated for all 40 sentences. For each cross-correlation analysis, the maximum of the envelope is obtained producing one delay value per sentence.

4) The IVS delay in the receive direction, $T_{R\text{-}jitter}(t)$, is obtained by subtracting the delay introduced by the test equipment ($T_{System}$) from $T_{TEAP\text{-}DRP}(t)$.

5) The difference $D_T$ between maximum receive delay obtained with at least five individual calls (see clause 8.2.3) and the delay $T_{R\text{-}constant}$ measured for the CSS burst (see step 2) in constant delay condition is calculated. The quantity CCVA = max(0, $D_T$) ("Call-to-Call Variability Adjustment") shall be added to $T_{R\text{-}jitter}(t)$ to obtain $T_{R\text{-}CCVA}(t) = T_{R\text{-}jitter}(t) + CCVA$.

6) The first two sentences are used for convergence of the jitter buffer management and are discarded from $T_{R\text{-}CCVA}(t)$ for the further analysis. Then, the two largest delay values of the remaining 38 sentences are excluded as well. The CCVA-adjusted IVS delay in receive, $T_{R\text{-}CCVA}$, shall be reported as the maximum of the remaining 36 sentences. All 40 delay per- sentence values of $T_{R\text{-}CCVA}(t)$ shall be provided in the test report.

NOTE – For packet-switched connections via VoLTE or VoNR, a variability of up to 20 ms may be expected between different calls due to the synchronization between the bits of the speech frames at the IVS antenna and the speech frame processing in the IVS (see clause 8.2.3). The effect of this possible call-to-call variation is taken into account with the quantity CCVA = max(0,$D_T$).

*Analysis for speech quality loss:*

1) In a first step, the initial CSS bursts and the first two sentences (used for convergence of the jitter buffer) are discarded from the analysis of the measurement microphone recording. The overall delay included in the recording shall be compensated by the test system with the previously determined delay.

2) For each sentence pair (8.0 s of signal) of the remaining recording, the speech quality shall be determined according to the prediction model [ITU-T P.863] in fullband mode. Since the IVS is operating in narrowband mode, the method according to Appendix III of [ITU-T P.863] shall be applied, including the active speech level calibration of each recorded sentence pair to −21 dBPa.

3) An average $MOS\text{-}LQO_f$ is calculated for each delay and loss profile as the arithmetic average across the 19 sentence pairs.

4) The difference between average $MOS\text{-}LQO_f$ of reference and test condition is evaluated as the speech quality loss.

# 9 Measurement parameters and requirements for wideband IVS terminals

As a general rule for all tests, it should be noted that it is always the responsibility of the test lab to verify that measurement results are not influenced by second-order effects such as additional noise falsifying the result of the individual measurement, for example.

## 9.1 Preparation measurements

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

## 9.2 Delay

### 9.2.1 Requirements

In general, the delay consists of an access specific delay $T_{as}$ and the IVS implementation dependant 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 IVS implementation dependant delay $T_{rtdimp}$ consists of:

– The IVS signal processing in send and receive.

– The air-paths from mouth to microphone and from the loudspeakers to the ear.

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

– The IVS delay in the send (uplink) direction Ts 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 IVS delay in the receive (downlink) direction Tr 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.

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

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 Tas from these specifications, 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 $T_{rtdimp}$ should not exceed 70 ms.

### 9.2.2 Delay in the send direction

#### 9.2.2.1 Test

The delay in the 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 (see Figure 9-1) is:

$$T_s + T_{System}$$

NOTE 1 – The delay should be minimized.

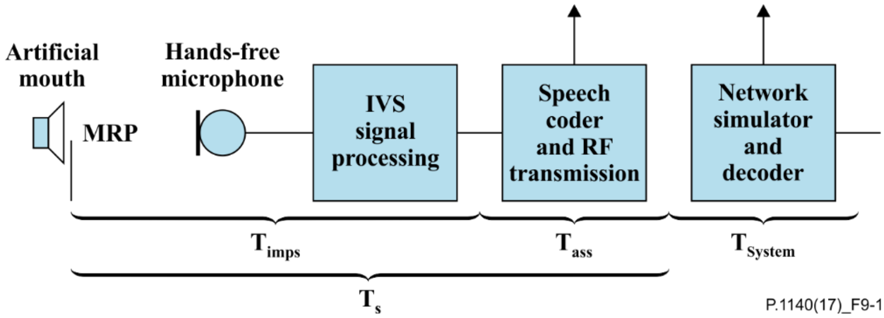



**Figure 9-1 – Different blocks contributing to the delay in the send direction**

The system delay $T_{System}$ is dependent on the transmission method used and the network simulator. This delay must be known and subtracted from the total measured delay.

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 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). 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 6. The HATS is positioned on the driver's 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 and access network.

3) The maximum of the cross-correlation function is used for the determination of the delay.

4) For packet-switched connections, the measurement shall be repeated at least for five individual calls. The maximum of the obtained delays is considered as the overall delay in send. The delay values of all individual calls shall be provided in the test report. In case the reference client of the test equipment does not compensate for the effect of speech frame synchronization, the maximum uncertainty quantity of 20 ms shall be subtracted from all delays obtained in send direction (see Note 3). If applicable, the correction shall be documented in the test report.

NOTE 2 – For packet-switched connections via VoLTE or VoNR, a variability of up to 20 ms may be expected between different calls due to the synchronization between the speech frame processing in the IVS in send and the bits of the speech frames at the IVS antenna. This synchronization is attributed to the IVS send delay according to the definition of the IVS delay reference points. Hence, the maximum value of delays obtained from at least five individual calls is reported as the overall IVS delay in send direction.

NOTE 3 – A further variability of up to 20 ms may be expected between different calls due to the synchronization between the speech frames at the IVS antenna and the speech frame processing in the receiving reference client of the test system. Hence, if the reference client of the test equipment does not compensate for the effect of the speech frame synchronization (as specified by the manufacturer of the reference client), this maximum uncertainty is subtracted from the measured maximum delay to obtain the true IVS delay in send.

### 9.2.3 Delay in the receive direction

#### 9.2.3.1 Test

The delay in the 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 (see Figure 9-2) is:

$T_r + T_{System}$

NOTE 1 – The delay should be minimized.

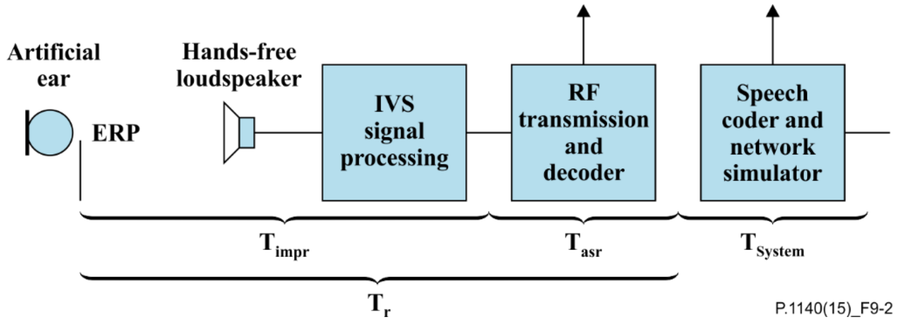



**Figure 9-2 – Different blocks contributing to the delay in the receive direction**

The system delay $T_{System}$ depends on the transmission system and on the network simulator used. This delay must be known and subtracted from the total measured delay.

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 shall be longer than the maximum expected delay. It is recommended to use a PN sequence of 16 k samples (with a 48 kHz sampling rate). The test signal level is −16 dBm0 at the electrical interface (POI).

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

2) The test arrangement is in accordance with clause 6. For packet-switched connections, no jitter or packet loss shall be introduced in the transmission. The HATS is positioned on the driver's seat. For the measurement the artificial head is 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 and access network.

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

5) For packet-switched connections, the measurement shall be repeated at least for five individual calls. The maximum of the obtained delays is considered as the overall delay in receive. The delay values of all individual calls shall be provided in the test report.

NOTE 2 – For packet-switched connections via VoLTE or VoNR, a variability of up to 20 ms may be expected between different calls due to the synchronization between the bits of the speech frames at the IVS antenna and the speech frame processing in the IVS. This synchronization is attributed to the IVS receive delay according to the definition of the IVS delay reference points. Hence, the maximum value of delays obtained from at least five individual calls is reported as the overall IVS delay in receive direction.

### 9.3 Loudness ratings

#### 9.3.1 Requirements

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

– SLR ≤ 22 dB for all positions in the car (the HATS positioned either at the driver's position, the co-driver's and the two outer passengers' back seats (2nd row), if available);

– A SLR = 13 dB ± 4 dB for HATS positioned at the driver's position is recommended.

NOTE – In general it is recommended to maintain an SLR of 13 dB at all positions in the car to achieve a sufficient intelligibility under all conditions (soft or weak talkers, more than 2 seat rows). This can be achieved, for example, by distributed microphones or properly designed AGC in send.

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

– RLR ≤ 10 dB for all positions in the car (the HATS positioned either at the driver's position, the co-driver's and the two outer passengers' back seats (2nd row), if available);

– A RLR = −3 dB ± 4 dB for HATS positioned on the driver's seat is recommended.

It is required to use background noise controlled AGC in the receiving direction. Performance requirements for this feature are addressed in clause 9.4.

#### 9.3.2 Test loudness rating in the 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 6. Tests are carried out with the HATS positioned on the driver's 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 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].

#### 9.3.3 Test loudness rating in the 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 dBm0, measured at the electrical reference point and averaged over the complete test signal sequence.

2) The test arrangement is according to clause 6. Tests are carried out with the HATS positioned on the driver's seat and additionally carried out with the HATS on the back passengers' seats (2nd row). The artificial head is 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.

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

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.

### 9.4 Variation of RLR in the presence of background noise

The intention of this test is the verification of the amplification range introduced by the AGC implementation in the presence of noise in the receiving direction. The RLR is determined in silent conditions (see test 9.3.3) and additionally in presence of background noise. The level of background noise needs to be sufficiently high and representative for a scenario as defined in A1.

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.

#### 9.4.1 Requirements

The IVS shall achieve a RLR ≤ −8 dB within 4 seconds after the start of playback of background noise scenario A3 in the vehicle cabin. The measured RLR is called $RLR_{A3}$.

NOTE – The AGC should be designed to allow a SNR of ≥ 6 dB for all signal and noise conditions. The AGC control range should allow $S/N \geq 6$ 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.

#### 9.4.2 Test

1) The RLR is determined as described in clause 9.3.3 (Test loudness rating in the 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 the reproducibility of the noise audio recording.

3) The background noise playback and recording is started 4 seconds before 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 the 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 $RLR_{A3}$.

5) $RLR_{A3}$ is determined as described in clause 9.3.3.

### 9.5 Sensitivity frequency responses

#### 9.5.1 Send sensitivity frequency response

##### 9.5.1.1 Requirements

The tolerance mask for the send sensitivity frequency response is shown in Table 9-1 and Figure 9-3, the mask is drawn by straight lines between the breaking points in Table 9-1 on a logarithmic (frequency) – linear (dB sensitivity) scale.

**Table 9-1 – Tolerance mask for the wideband send sensitivity frequency response**

| Frequency [Hz] | Upper limit | Lower limit |
|---|---|---|
| 100 Hz | 4 dB | −∞ dB |
| 200 Hz | 4 dB | −10 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 | −∞ 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. | | |

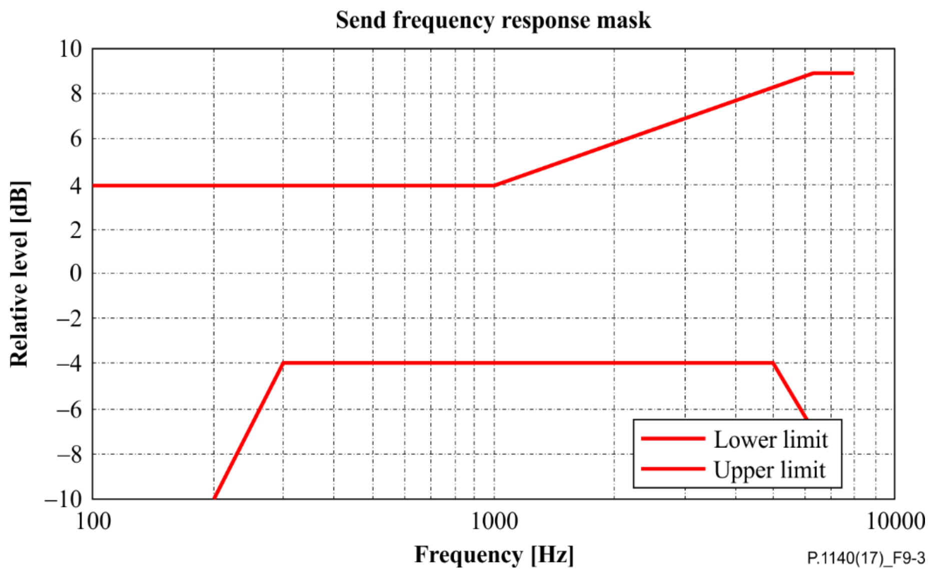



**Figure 9-3 – Wideband send frequency response mask (Fig. is informative)**

### 9.5.1.2 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 61260-1] for frequencies of 100 Hz to 8 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.

### 9.5.2 Receive sensitivity frequency response

#### 9.5.2.1 Requirements

The tolerance mask for the receive sensitivity frequency response is shown in Table 9-2 and Figure 9-4, the mask is drawn by straight lines between the breaking points in Table 9-2 on a logarithmic (frequency) – linear (dB sensitivity) scale.

**Table 9-2 – Tolerance mask for the wideband receive sensitivity frequency response**

| Frequency [Hz] | Upper limit | Lower limit |
|---|---|---|
| 125 | 6 dB | −∞ dB |
| 200 | 6 dB | −∞ dB |
| 315 | 6 dB | −∞ dB |
| 400<br>630<br>6300<br>8000 | 6 dB<br>6 dB<br>6 dB<br>6 dB | −9 dB<br>−6 dB<br>−6 dB<br>−∞ dB |
| NOTE – All sensitivity values are expressed in dB on an arbitrary scale. | | |

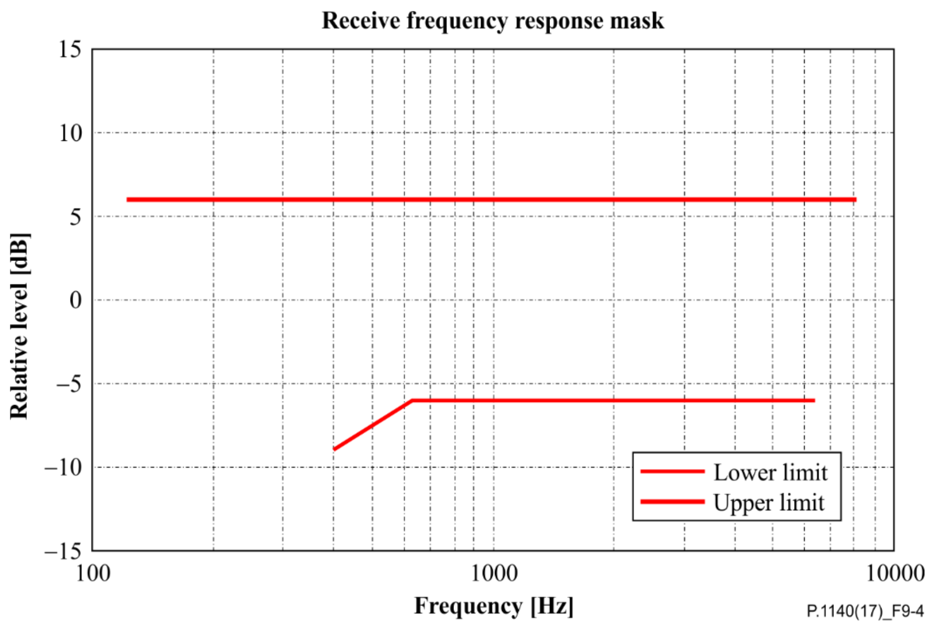



**Figure 9-4 – Receive frequency response mask (Fig. is informative)**

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 at a 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 Recommendation uses the diffuse-field as a reference.

#### 9.5.2.2 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 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 from 100 Hz to 8 kHz.

3) The sensitivity is determined in dBPa/V.

## 9.6 Speech intelligibility during single talk

### 9.6.1 One-way speech intelligibility in send

Speech intelligibility is one of the most important parameter for emergency call systems. The intention of this test is to verify, that a sufficient level of speech intelligibility can be achieved. Currently no objective method is available. As soon as a validated method is available that is applicable for emergency call systems this section should be added.

*For further study.*

### 9.6.2 Speech intelligibility in receive

Speech intelligibility is one of the most important parameter for emergency call systems. The intention of this test is to verify that a sufficient level of speech intelligibility can be achieved. Currently no objective method is available. As soon as a validated method is available that is applicable for emergency call systems this section should be added.

*For further study.*

## 9.7 Idle channel noise

All tests are conducted with average RF-signal power settings. It is recommended, in addition, to check the requirement 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) shall not be exceeded.

### 9.7.1 Idle channel noise in the send direction

#### 9.7.1.1 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 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 which are 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.

#### 9.7.1.2 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 −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 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 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 A-weighting.

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 a 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 up to 7 kHz.

### 9.7.2 Idle channel noise in the receive direction

#### 9.7.2.1 Requirements

The requirements for the maximum noise produced by the IVS in the case where 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 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 which are considered to be inaudible due to the very low noise floor produced by the hands-free terminal on average, it is the responsibility of the test lab to demonstrate the desired performance subjectively.

#### 9.7.2.2 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 dBm0, measured at the electrical reference point. The level of the activation signal is averaged over the complete duration of the activation signal.

2) The test arrangement is according to clause 6. For the measurement of the IVS the artificial head is diffuse-field equalized (see clause 6.8.2), 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 reverberation 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.

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 up to 7 kHz.

## 9.8 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) 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 maximum background noise level of −64 dBPa(A) shall not be exceeded.

### 9.8.1 Terminal coupling loss

#### 9.8.1.1 Requirements

The terminal coupling loss (TCL) in quiet environments shall be at least 46 dB for all settings of the volume control. The implemented echo control mechanism should provide sufficient echo loss for all typical environments and typical impulse responses.

NOTE 1 – A TCL of ≥ 50 dB is recommended as a performance objective. 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, for example, by comparing the analysis with the idle channel noise measurement results.

NOTE 2 – There may be implementations where echo problems are 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.

#### 9.8.1.2 Test

1) All tests are conducted in the car cabin; the test arrangement is described in clause 6. 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 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 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 of [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 $f_N$.

### 9.8.2 Echo level versus time

#### 9.8.2.1 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 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 – 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.

#### 9.8.2.2 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 dBm0 as well as an average level of −25 dBm0. The echo signal is analysed during a period of at least 2.8 s, which represents 8 periods of the CS signal. The integration time for the level analysis shall be 35 ms, the analysis is referred to the level analysis of the reference signal. In addition, the test is repeated with 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 dBm0. 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).

### 9.8.3 Echo performance with time variant echo path and speech

#### 9.8.3.1 Requirements

This test is intended 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 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 that is quietly moving one arm close to the microphone during the measurement.

The type of echo path variation chosen shall be reported.

#### 9.8.3.2 Test

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

2) The test arrangement is according to clause 6. 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 dBm0. 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 reference.

3) Now a second measurement is started introducing the time-variant echo path.

4) The test signal is 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 dBm0. 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 echo-path is determined.

6) The measurement result is displayed as attenuation vs. time. The exact synchronization between the two measured signals has to be guaranteed.

### 9.9 Switching characteristics

#### 9.9.1 Activation in the 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).

##### 9.9.1.1 Requirements

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

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

##### 9.9.1.2 Test

The structure of the test signal is shown in Figure 9-5. The test signal consists of CSS components according to [ITU-T P.501] with increasing level for each CSS burst.

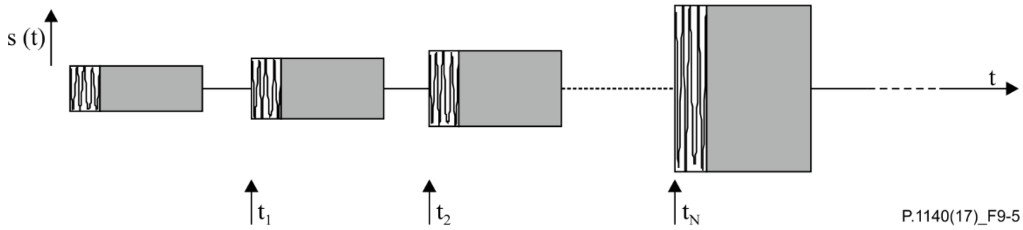



**Figure 9-5 – Test signal to determine the minimum activation level and the built-up time**

The settings of the test signal are shown in Table 9-3.

**Table 9-3 – 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 dB |
| NOTE – The level of the active signal part corresponds to an average level of −24.7 dBPa at the MRP for the CSS according to [ITU-T 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 corresponding CSS burst.

### 9.10 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 the send direction during double talk $A_{H,S,dt}$.

– Attenuation range in the receive direction during double talk $A_{H,R,dt}$.

– Echo attenuation during double talk.

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

##### 9.10.1.1 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 9-4.

**Table 9-4 – 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 shall provide a double talk capability of type 2b or better in the sending direction. The requirements apply for the nominal (default) setting of the receive volume control.

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 9-4 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.

##### 9.10.1.2 Test

The test signal to determine the attenuation range during double talk is shown in Figure 9-6. 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.

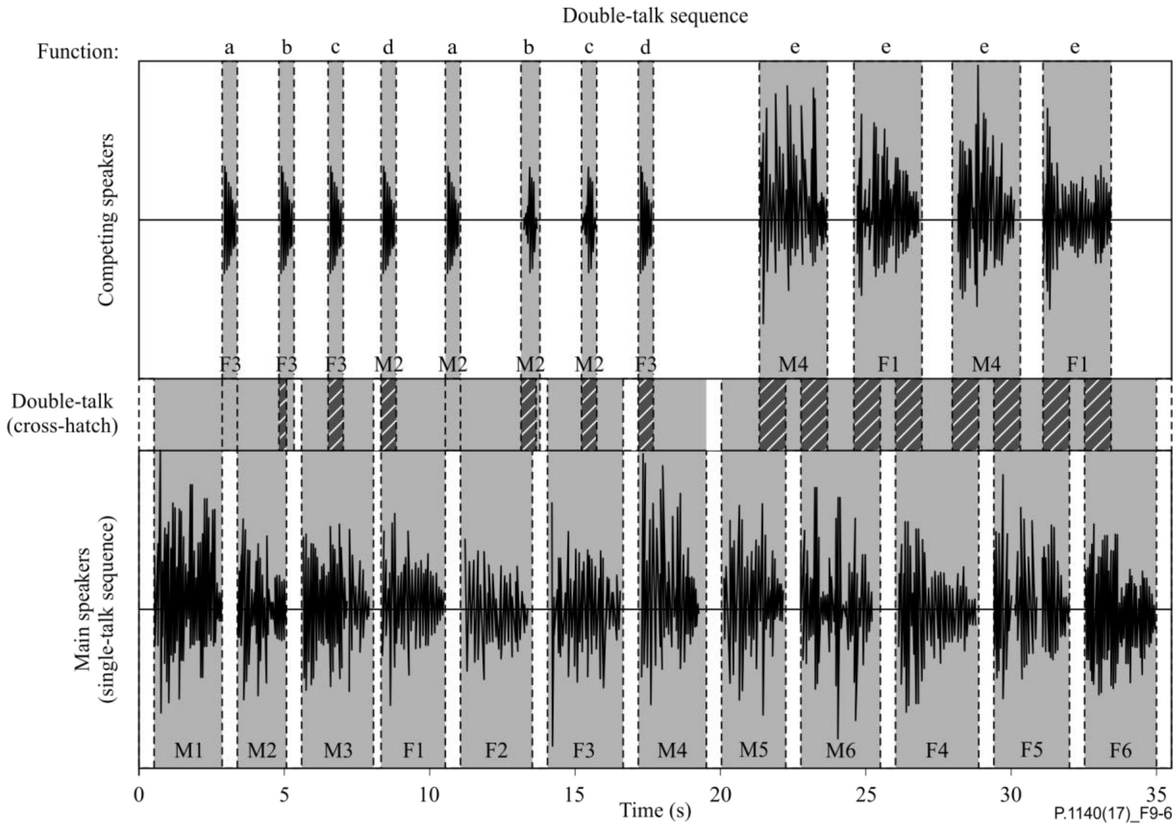



**Figure 9-6 – Double talk test sequence with overlapping speech sequences in the send and receive directions**

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 9-5 – Level of the double talk sequences**

| | Receive direction | Send direction |
|---|---|---|
| Average signal level | −16 dBm0 | −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 dBm0 is applied to the electrical reference point.

2) When determining the attenuation range in the 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.

### 9.10.2 Attenuation range in the 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.

#### 9.10.2.1 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 9-6.

**Table 9-6 – 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 shall provide a double talk capability of type 2b or better in the receiving direction. The requirements apply for nominal setting of the receive volume control.

In general, Table 9-6 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.

#### 9.10.2.2 Test

The test signal to determine the attenuation range during double talk is shown in Figure 9-7. 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.

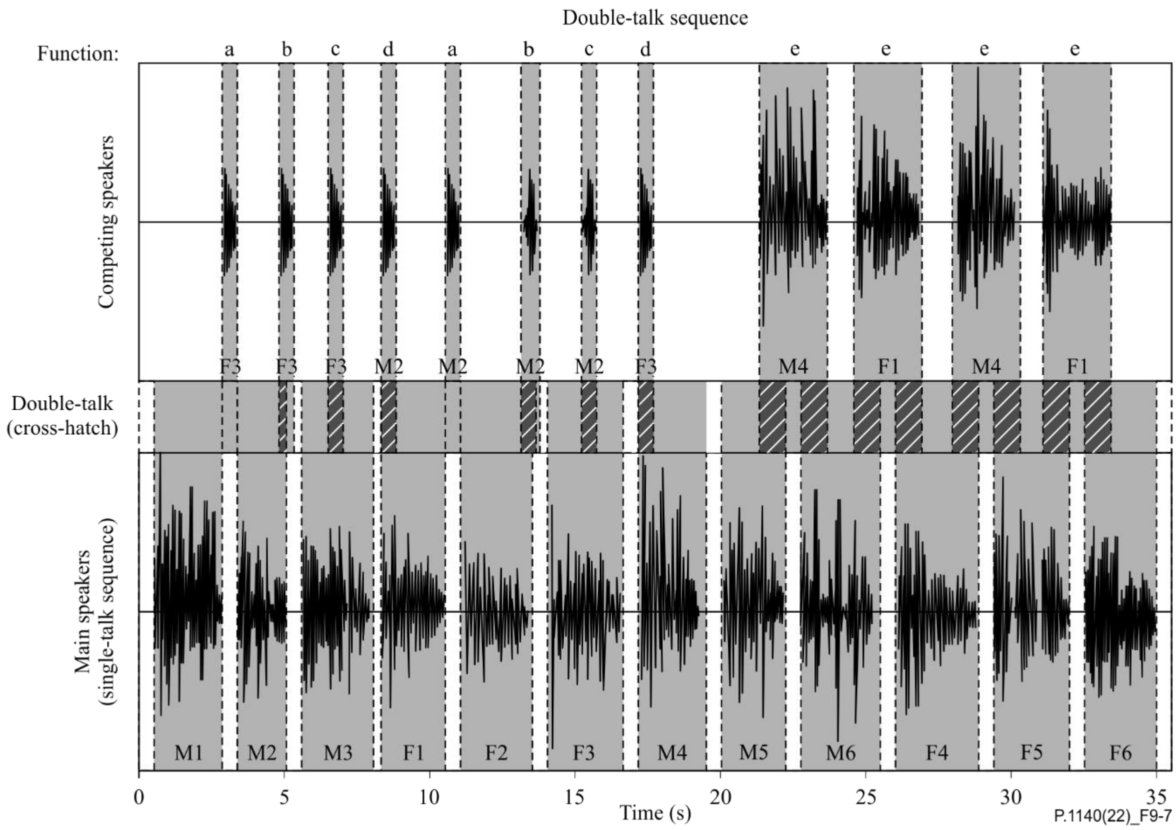



**Figure 9-7 – Double talk test sequence with overlapping speech sequences in the receive and send directions**

The settings for the test signals are given in Table 9-7.

**Table 9-7 – Level of the double talk sequences**

| | Receive direction | Send direction |
|---|---|---|
| Average signal level | −16 dBm0 | −1.7 dBPa |

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

2) When determining the attenuation range in the 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.

### 9.10.3 Detection of echo components during double talk

#### 9.10.3.1 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 9-8 are applicable (more information can be found in Annex A of [ITU-T P.340]).

**Table 9-8 – 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 |

The IVS shall provide an echo-loss during double talk capability of type 2c or better. The requirements apply for a nominal setting of the receive volume control.

#### 9.10.3.2 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 signal is described in [ITU-T P.501]. The signal settings used are shown in Table 9-9. 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 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 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 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 9-8. The echo attenuation is to be achieved for each individual frequency band from 200 Hz to 6 950 Hz according to the different categories.

**Table 9-9 – Parameters of the two test signals for double talk measurement based on AM-FM modulated sine waves**

| **Send direction** | | **Receive direction** | |
|---|---|---|---|
| $\boldsymbol{f_0^{(1)}}$ **(Hz)** | $\pm\Delta \boldsymbol{f^{(1)}}$ **(Hz)** | $\boldsymbol{f_0^{(2)}}$**(Hz)** | $\pm\Delta \boldsymbol{f^{(2)}}$ **(Hz)** |
| 125 | ±2.5 | 180 | ±2.5 |
| 250 | ±5 | 270 | ±5 |
| 500 | ±10 | 540 | ±10 |
| 750 | ±15 | 810 | ±15 |
| 1 000 | ±20 | 1 080 | ±20 |
| 1 250 | ±25 | 1 350 | ±25 |
| 1 500 | ±30 | 1 620 | ±30 |
| 1 750 | ±35 | 1 890 | ±35 |
| 2 000 | ±40 | 2 160 | ±35 |
| 2 250 | ±40 | 2 400 | ±35 |
| 2 500 | ±40 | 2 650 | ±35 |
| 2 750 | ±40 | 2 900 | ±35 |
| 3 000 | ±40 | 3 150 | ±35 |
| 3 250 | ±40 | 3 400 | ±35 |
| 3 500 | ±40 | 3 650 | ±35 |
| 3 750 | ±40 | 3 900 | ±35 |
| 4 000 | ±40 | 4 150 | ±35 |
| 4 250 | ±40 | 4 400 | ±35 |
| 4 500 | ±40 | 4 650 | ±35 |
| 4 750 | ±40 | 4 900 | ±35 |
| 5 000 | ±40 | 5 150 | ±35 |
| 5 250 | ±40 | 5 400 | ±35 |
| 5 500 | ±40 | 5 650 | ±35 |
| 5 750 | ±40 | 5 900 | ±35 |
| 6 000 | ±40 | 6 150 | ±35 |
| 6 250 | ±40 | 6 400 | ±35 |
| 6 500 | ±40 | 6 650 | ±35 |
| 6 750 | ±40 | 6 900 | ±35 |
| 7 000 | ±40 | | |

### 9.10.4 Robustness of double talk capability with far end background noise at the PSAP

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 the echo suppression unit in the IVS and hamper the transmission of the near end voice.

*For further study.*

## 9.11 Background noise transmission

### 9.11.1 Silent call performance

The intention of this test is to verify that in the absence of speech the background noise captured by the IVS microphone is transmitted mostly transparently without artefacts to preserve the information contained in the background noise.

#### 9.11.1.1 Requirements

The level variation vs. time of the transmitted noise from the emergency call noise scenario (A4) shall be less than 10 dB when referred to a signal picked up close to the IVS microphone.

#### 9.11.1.2 Test

1) The test arrangement is in accordance with clause 6.
2) The test signal is the simulated emergency call noise scenario (A4) with a duration of 20 s. The test signal level is the original signal level.
3) The transmitted background noise signal picked up at the POI is analysed during between 11 s and 19 s as a level versus time analysis. The integration time for the level analysis shall be 125 ms. The analysis is referred to the level analysis of the reference signal. The reference signal is the background noise signal picked up by an omnidirectional measurement microphone positioned close to the IVS microphone. The reference signal is filtered using the modified IRS send filter as defined in [ITU-T P.830].
4) The measurement result is displayed as attenuation versus time. The exact synchronization between the input and output signals has to be guaranteed.

NOTE – More advanced techniques are under study.

### 9.11.2 Speech intelligibility in the presence of background noise

The intention of this test is to verify, that in the presence of background noise a sufficient level of speech intelligibility can be achieved. Currently no objective method is available. As soon as a validated method is available which is applicable for emergency call systems this section should be added.

*For further study.*

## 9.12 Jitter buffer management

### 9.12.1 Overview

For measurements conducted in packet-switched network connection with simulated packet arrival time variations and packet loss, the intention of the following test is to ensure that the jitter buffer of the IVS does not severely impact speech communication. To ensure a minimum conversational quality, the increase of delay due to higher and more variant packet arrival time should be minimized and the loudspeaker output in receive should still be comprehensible, i.e., clipping, attenuation or insufficient packet loss concealment of the received signal should be avoided.

### 9.12.2 Requirements

Requirements are not applicable for circuit-switched network connections.

For measurements conducted in packet-switched network connection with simulated packet arrival time variations and packet loss, the sum of the IVS delays in send and receive directions ($T_S + T_R$) shall be less than or equal to the delay requirements given in Table 9-10.

**Table 9-10 – IVS delay requirements for LTE/NR access**

| Test condition | Delay and loss profile | Requirement for maximum delay | Comment |
|---|---|---|---|
| 0 | Error and jitter free condition | $T_S + T_R \leq 170$ ms | Same as in clause 9.2 |
| 1 | dly_profile_20msDRX_10pct_BLER_e2e | $T_S + T_R \leq 170$ ms | |
| 2 | dly_profile_40msDRX_10pct_BLER_e2e | $T_S + T_R \leq 210$ ms | |
| 3 | dly_profile_40msDRX_22pct_BLER_e2e | $T_S + T_R \leq 210$ ms | |

To ensure that the jitter buffer management does not significantly impact the loudspeaker output in receive direction, the loss in speech quality per profile shall not exceed the requirements given in Table 9-11.

NOTE – The purpose of this test is to provide a relative comparison of the objective speech quality between the reference and test conditions. This test is not to be construed as a method to evaluate the absolute objective speech quality of the device.

**Table 9-11 – IVS requirements on speech quality decrease for LTE/NR access**

| Test Condition | Delay and loss profile | Requirement for speech quality loss | Comment |
|---|---|---|---|
| 0 | Error and jitter free condition | - | Reference for actual test profiles |
| 1 | dly_profile_20msDRX_10pct_BLER_e2e | ≤ 0.4 | |
| 2 | dly_profile_40msDRX_10pct_BLER_e2e | ≤ 0.4 | |
| 3 | dly_profile_40msDRX_22pct_BLER_e2e | ≤ 1.0 | |

### 9.12.3 Test signal

The same test signal as described in clause 8.12.3 shall be used.

### 9.12.4 Delay and loss profiles

The same delay and loss profiles as described in clause 8.12.4 shall be used.

### 9.12.5 Test

The same test arrangement and procedure as described in clause 8.12.5 shall be used. For the determination of speech quality loss, Appendix III of [ITU-T P.863] shall not be applied, whereas the active speech level calibration of each recorded sentence pair to −21 dBPa shall still be applied.

# Annex A

## Alternative method for determining the roundtrip delay

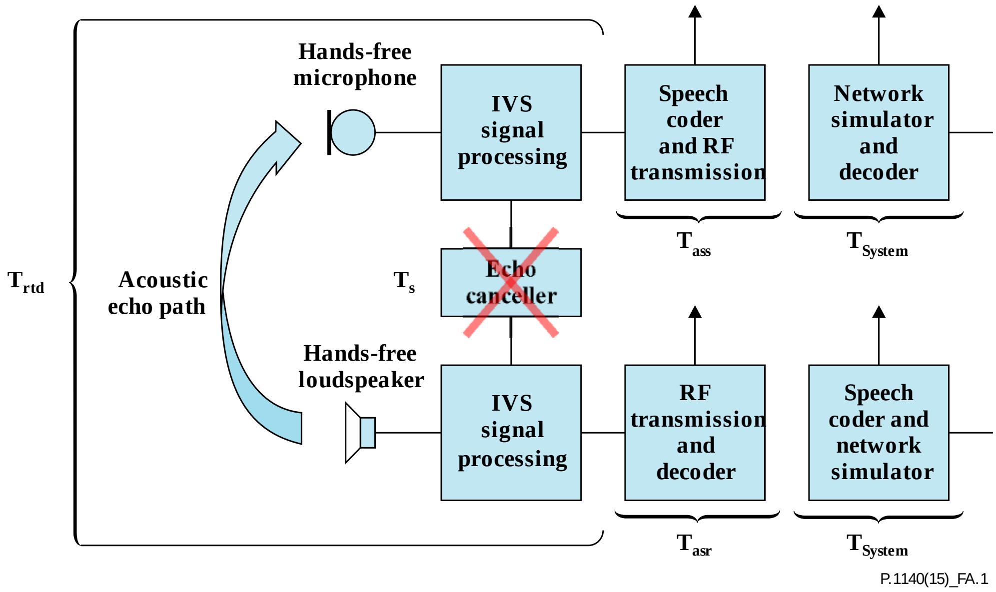
(This annex forms an integral part of this Recommendation.)

An alternative method to determine the roundtrip delay of an IVS is given below. Please note that this method can only be applied if the echo canceller of the IVS can be deactivated.

In the case where the IVS provides a test mode to disable echo cancellation and echo suppression signal processing, the roundtrip 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 be measured directly:

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

NOTE – The delay should be minimized.




The system delay $t_{rtSystem}$ depends 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 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). The test signal level is −16 dBm0 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.

# Annex B

## Void

# Bibliography


[b-ITU-T G.114] Recommendation ITU-T G.114 (in force), *One-way transmission time*.
<http://www.itu.int/ITU-T/recommendations/rec.aspx?rec=6254>

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

[b-ITU-T P.1150] Recommendation ITU-T P.1150 (in force), *In-car communication audio specification*.
<https://www.itu.int/rec/T-REC-P.1150>

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

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

[b-KettlerMotorcycles2017] Frank Kettler, Silvia Poschen, Marc Lepage, Radi Serafimov, *Speech Communication in Emergency Call Scenarios for Motorcycles*, Proceedings of DAGA Conference 2017, Kiel.

[b-SAE J941] [b-SAE J941] SAE J941 *201003: Surface vehicle recommended practice*.

[b-3GPP TS 26.131] 3GPP TS 26.131 (in force), *Terminal acoustic characteristics for telephony; Requirements*.
<http://www.3gpp.org/DynaReport/26131.htm>

[b-3GPP TS 26.132] 3GPP TS 26.132 (in force), *Speech and video telephony terminal acoustic test specification*.
<http://www.3gpp.org/DynaReport/26132.htm>

## 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 |
| **Series P** | **Telephone transmission quality, telephone installations, local line networks** |
| 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 |

Printed in Switzerland
Geneva, 2022