

International Telecommunication Union

**ITU-T**

TELECOMMUNICATION  
STANDARDIZATION SECTOR  
OF ITU

**P.381**

(07/2016)

SERIES P: TERMINALS AND SUBJECTIVE AND  
OBJECTIVE ASSESSMENT METHODS

Voice terminal characteristics

---

**Technical requirements and test methods for  
the universal wired headset or headphone  
interface of digital mobile terminals**

Recommendation ITU-T P.381

ITU-T



ITU-T P-SERIES RECOMMENDATIONS  
**TERMINALS AND SUBJECTIVE AND OBJECTIVE ASSESSMENT METHODS**

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	Series	P.10
<b>Voice terminal characteristics</b>	<b>Series</b>	<b>P.30</b>
		<b>P.300</b>
Reference systems	Series	P.40
Objective measuring apparatus	Series	P.50
		P.500
Objective electro-acoustical measurements	Series	P.60
Measurements related to speech loudness	Series	P.70
Methods for objective and subjective assessment of speech quality	Series	P.80
Methods for objective and subjective assessment of speech and video quality	Series	P.800
Audiovisual quality in multimedia services	Series	P.900
Transmission performance and QoS aspects of IP end-points	Series	P.1000
Communications involving vehicles	Series	P.1100
Models and tools for quality assessment of streamed media	Series	P.1200
Telemeeting assessment	Series	P.1300
Statistical analysis, evaluation and reporting guidelines of quality measurements	Series	P.1400
Methods for objective and subjective assessment of quality of services other than speech and video	Series	P.1500

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

## Recommendation ITU-T P.381

### Technical requirements and test methods for the universal wired headset or headphone interface of digital mobile terminals

#### Summary

Recommendation ITU-T P.381 specifies critical physical and electrical-acoustical characteristics for the universal headset interface and provides corresponding test methods. Both 3.5 mm and 2.5 mm diameter headset/headphone interfaces have been widely used in digital mobile terminals in recent years. Nowadays, the consumer is free to choose either the headset/headphone originally provided by the terminal manufacturer or others that are offered separately. However, the quality of service (QoS)/quality of experience (QoE) perceived by users is influenced by both the electrical performance of the interface and the compatibility between the terminal and the connected headset/headphone.

#### History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T P.381	2012-08-22	12	<a href="http://handle.itu.int/11.1002/1000/11685">11.1002/1000/11685</a>
2.0	ITU-T P.381	2014-02-13	12	<a href="http://handle.itu.int/11.1002/1000/12139">11.1002/1000/12139</a>
3.0	ITU-T P.381	2016-07-29	12	<a href="http://handle.itu.int/11.1002/1000/12970">11.1002/1000/12970</a>

---

\* 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, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at <http://www.itu.int/ITU-T/ipr/>.

© ITU 2016

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

## Table of Contents

	<b>Page</b>
1	Scope..... 1
2	References..... 1
3	Definitions ..... 2
3.1	Terms defined elsewhere ..... 2
3.2	Terms defined in this Recommendation..... 3
4	Abbreviations and acronyms ..... 3
5	General description ..... 5
6	Physical characteristics ..... 5
6.1	General rules..... 5
6.2	Pin assignments ..... 5
7	Electrical interface specification..... 7
7.1	Communication mode ..... 7
7.2	Multimedia playback mode ..... 26
8	Headset specification ..... 30
8.1	Communication mode ..... 30
8.2	Multimedia playback mode ..... 35
9	Function requirements for terminals with the universal headset interface ..... 38
Appendix I – Audio connectivity for sockets with four contact points ..... 41	
I.1	2.5 mm diameter plug connector with four poles ..... 41
I.2	2.5 mm diameter socket connector with four contact points ..... 42
I.3	3.5 mm diameter plug connector with four poles ..... 42
I.4	3.5 mm diameter socket connector with four contact points ..... 43
Appendix II – Audio connectivity for sockets with four contact points (optional dimensions to accommodate terminal designs with curved edges)..... 45	
Appendix III – Audio connectivity for sockets with three contact points ..... 46	
III.1	2.5 mm diameter plug connector with three poles ..... 46
III.2	2.5 mm diameter socket connector with three contact points ..... 46
III.3	3.5 mm diameter plug connector with three poles ..... 47
III.4	3.5 mm diameter socket connector with three contact points ..... 47
Appendix IV – Other considerations ..... 49	
IV.1	Filter recommendation..... 49
IV.2	Electrostatic discharge ..... 49
IV.3	Microphone basics – background ..... 49
IV.4	Vcc voltage for microphone bias..... 49
IV.5	DC resistance of microphone ..... 50
Bibliography..... 51	



## Recommendation ITU-T P.381

# Technical requirements and test methods for the universal wired headset or headphone interface of digital mobile terminals

## 1 Scope

This Recommendation specifies electrical requirements and test methods for the universal analogue headset/headphone interface used in digital mobile terminals.

The principle of this document is to ensure adequate compatibility between the digital mobile terminal and the wired analogue headset/headphone, and to have a better user experience. The universality of the headset/headphone interface will facilitate the separation of sales between digital mobile terminals and headsets/headphones. One of the benefits is that the user can be free to choose his or her favourite type of headset or headphone that is available on the market. In the long run, it will reduce e-waste. Furthermore, the universal interface can be used as the electric coupling design in hands-free systems and hearing aids for wider harmonization.

In order to provide instructions to manufacturers and encourage them to adopt the universal headset interface, the mechanical dimensions are shown in Appendices I, II and III.

This Recommendation is applicable to digital mobile terminals with a physical analogue audio output/input interface. Other similar ICT equipment may also refer to this Recommendation.

This Recommendation is not applicable for terminals designed solely for digital headset/headphone usage.

## 2 References

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

- [ITU-T G.122] Recommendation ITU-T G.122 (1993), *Influence of national systems on stability and talker echo in international connections.*
- [ITU-T P.10] Recommendation ITU-T P.10/G.100 (2006), *Vocabulary for performance and quality of service.*
- [ITU-T P.50] Recommendation ITU-T P.50 (1999), *Artificial voices.*
- [ITU-T P.56] Recommendation ITU-T P.56 (2011), *Objective measurement of active speech level.*
- [ITU-T P.57] Recommendation ITU-T P.57 (2011), *Artificial ears.*
- [ITU-T P.58] Recommendation ITU-T P.58 (2011), *Head and torso simulator for telephonometry.*
- [ITU-T P.79] Recommendation ITU-T P.79 (2007), *Calculation of loudness ratings for telephone sets.*
- [ITU-T P.340] Recommendation ITU-T P.340 (2000), *Transmission characteristics and speech quality parameters of hands-free terminals.*

- [ITU-T P.380] Recommendation ITU-T P.380 (2003), *Electro-acoustic measurements on headsets*.
- [ITU-T P.501] Recommendation ITU-T P.501 (2009), *Test signals for use in telephonometry*.
- [ITU-T P.502] Recommendation ITU-T P.502 (2000), *Objective test methods for speech communication systems using complex test signals*.
- [ITU-T P.581] Recommendation ITU-T P.581 (2009), *Use of head and torso simulator (HATS) for hands-free and handset terminal testing*.
- [ITU-T P.863] Recommendation ITU-T P.863 (2011), *Perceptual objective listening quality assessment*.
- [EN 50332-1] CENELEC EN 50332-1 (2013), *Sound system equipment: Headphones and earphones associated with personal music players – Maximum sound pressure level measurement methodology – Part 1: General method for "one package equipment"*.
- [EN 50332-2] CENELEC EN 50332-2 (2013), *Sound system equipment: Headphones and earphones associated with personal music players – Maximum sound pressure level measurement methodology – Part 2: Matching of sets with headphones if either or both are offered separately, or are offered as one package equipment but with standardised connectors between the two allowing to combine components of different manufacturers or different design*.
- [ETSI TS 103 106] ETSI TS 103 106 V1.2.1 (2013), *Speech and multimedia Transmission Quality (STQ); Speech quality performance in the presence of background noise: Background noise transmission for mobile terminals-objective test methods*.
- [IEC 60268-1] IEC 60268-1 (1985), *Sound system equipment – Part1: General*.
- [IEC 60268-7] IEC 60268-7 (2010), *Sound system equipment – Part 7: Headphones and earphones*.
- [IEC 61260-1] IEC 61260 -1(2014), *Electroacoustics – Octave-band and fractional-octave-band filters. – Part 1: Specifications*.
- [IEC 61672-1] IEC 61672-1 (2013), *Electroacoustics – Sound level meters – Part 1: Specifications*.

### 3 Definitions

#### 3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

**3.1.1 composite source signal (CSS)** [ITU-T P.10]: Signal composed in time by various signal elements.

**3.1.2 eardrum reference point (DRP)** [ITU-T P.10]: A point located at the end of the ear canal, corresponding to the eardrum position.

**3.1.3 earphone** [IEC 60268-7]: Electroacoustic transducer by which acoustic oscillations are obtained from electric signals and intended to be closely coupled acoustically to the ear.

**3.1.4 headset** [ITU-T P.10]: A device which includes a telephone receiver and transmitter which is typically secured to the head or the ear of the wearer.



**3.1.5 mouth reference point (MRP)** [ITU-T P.10]: Point 25 mm in front of and on the axis of the lip plane of the artificial mouth or a typical human mouth (see Figure 1 of ITU-T P.64).

## 3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

**3.2.1 artificial ear:** A device which incorporates an acoustic coupler and a calibrated microphone for measuring sound pressure, and which has an overall acoustic impedance similar to that of the average adult ear over a given frequency band (based on the definition in [ITU-T P.10]).

**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 head and torso simulator (HATS) for telephonometry:** A manikin that extends downwards from the top of the head to the waist. It is designed to simulate the sound pick-up characteristics and the acoustic diffraction produced by the average adult, and to reproduce the acoustic field generated by the human mouth (based on the definition in [ITU-T P.10]).

**3.2.4 headphone:** An object based on the assembly of one or two earphones on a headband or chinband, the use of which may be optional (e.g., with intra-concha earphones) (based on the definition in [IEC 60268-7]).

**3.2.5 mean opinion score – listening-only quality objective (MOS-LQO):** The score is calculated by means of an objective model which aims at predicting the quality for a listening-only test situation. Objective measurements made using the model given in [ITU-T P.863] give results in terms of MOS-LQO.

**3.2.6 receive:** The receiving direction of the signal transmission, usually from the measurement system to the device under test (DUT).

**3.2.7 send:** The sending direction of the signal transmission, usually from the device under test (DUT) to the measurement system.

## 4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

ABT	Audio Breakthrough Testing
AGC	Automatic Gain Control
A <sub>H,R,dt</sub>	Attenuation range in the receive direction during double talk
A <sub>H,S,dt</sub>	Attenuation range in the send direction during double talk
dBFS	decibels relative to Full Scale.
dBV	decibels relative to 1 Volt
dBVp	decibels relative to 1 Volt, psophometrically weighted
DRP	Drum Reference Position
DUT	Device Under Test
EC	Echo Cancellor
ECM	Electret condenser microphone
EL	Echo Loss
EMI	Electromagnetic Interference
ESD	Electrostatic Discharge

FFT	Fast Fourier Transform
GND	Ground
HATS	Head and Torso Simulator
HTCLw	Headset Terminal Coupling Loss weighted
JFET	Junction Field Effect Transistor
L	Left audio channel
$L_{\text{sendnom}}$	Level in send for nominal speech input level
$L_{\text{nominal}}$	nominal input Level
$L_{\text{S,min}}$	minimum activation Level in the send direction
$L_{\text{R,min}}$	minimum activation Level in the receive direction
LQ	Listening Quality
LQO	Listening Quality Objective
LQ <sub>R</sub>	Listening Quality in the receive direction
LQ <sub>S</sub>	Listening Quality in the send direction
LQO <sub>n</sub>	Listening Quality Objective, narrowband
LQO <sub>sw</sub>	Listening Quality Objective super-wideband
LSB	Lowest Significant Bit
MIC	Microphone
MONO	Mono Audio Channel
MOS	Mean Opinion Score
MRP	Mouth Reference Point
NLP	Non-linear Processing
POI	Point of Interconnection
R	Right audio channel
$R_{\text{bias}}$	bias Resistance
SNR	Signal to Noise Ratio
STMR	Sidetone Masking Rating
TCL	Terminal Coupling Loss
TCLw	weighted Terminal Coupling Loss
TELRdt	Talker Echo Loudness Rating under double talk conditions
$T_S$	delay in the sending direction
$T_{r,S,min}$	Minimum built-up time in Send
$T_R$	delay in the receiving direction
$T_{TES}$	system delay
UE	User Equipment
$V_m$	maximum output voltage
SPCV	Simulated programme signal characteristic voltage

## 5 General description

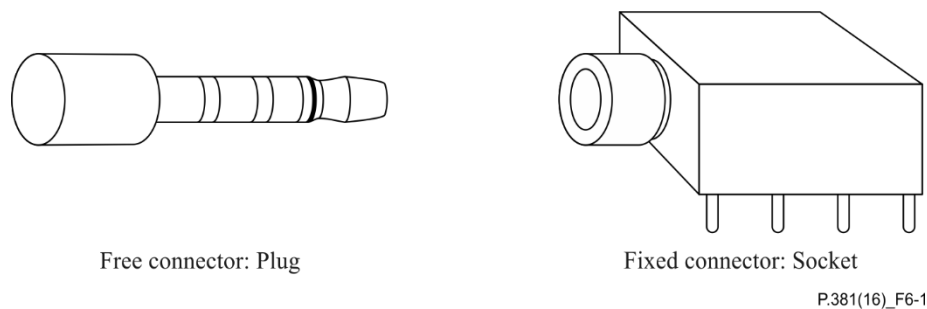
Generally, if a headset or a headphone is used, the overall user experience during a call highly depends on both the terminal and the connected headset/headphone. Although the acoustic quality of the headset/headphone is usually the weak link, more consideration with regard to the physical and electrical performance of the universal interface is needed.

This Recommendation specifies the universal concentric connector interface for successful interconnection between the digital mobile terminal and the headset/headphone, including the plug connector and the socket connector. Normally, the socket connector is fixed inside the terminal, with the outside rim level with the surrounding shell of the terminal. Particularly, if the outside rim is lower than the surrounding shell, the dimension of the plug hand grip shall not influence the precise mating.

## 6 Physical characteristics

### 6.1 General rules

Two types of concentric socket connectors are recommended for use, the 2.5 mm and the 3.5 mm diameter socket connectors. Isometric view of the plug and the socket connector are shown in Figure 6-1:



**Figure 6-1 – Isometric view of the plug and socket connector**

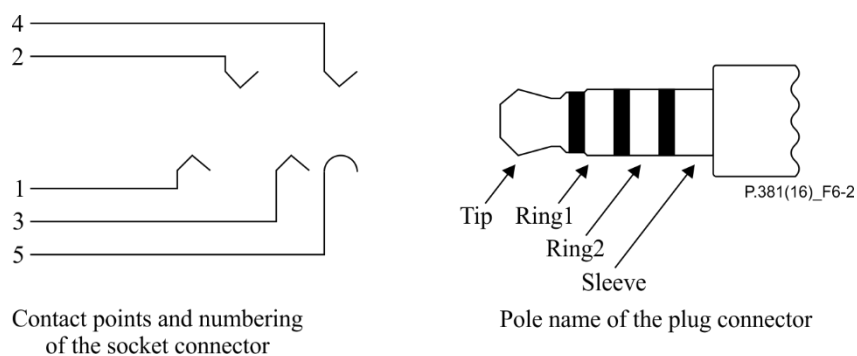
If the terminal is equipped with a headset interface and designed for both communication and audio-playing, the fixed connector shall be a 3.5 mm diameter or 2.5 mm diameter concentric socket connector with four contact points. Detailed dimension information about the socket connector with four contact points is given in Appendix I. Some terminals with curved edges may not work well with the dimensions given in Appendix I. For these cases, optional dimensions are given in Appendix II, which is fully compatible with connectors designed to the dimensions of Appendix I.

Although socket connectors with three contact points are no longer recommended for digital mobile terminals, dimension information about these connectors is given in Appendix III to help achieve complete compatibility.

NOTE – The contact points here do not include special points reserved for other functions.

### 6.2 Pin assignments

This clause gives an illustration of pin assignments of the socket connector with four contact points and those of the mated plug, as shown in Figure 6-2.



**Figure 6-2 – Pin assignments of a socket connector with four contact points**

The physical pinout order of the universal interface is important and should coordinate with the connected headset/headphone.

A socket connector with four contact points shall be compatible with both the plug connector with three poles and the plug connector with four poles specified in this Recommendation.

There are two types of headsets with different pin assignments commonly used today across geographical regions. Depending on the status, it was recommended that terminals should be able to identify both plugs intelligently and automatically.

### 6.2.1 Recommended pin assignment

Point 1 of the socket is to be connected to the tip of the plug, linking it to the left-hand channel of the receiver (L audio). Point 2 is to be connected to Ring 1, linking it to the right-hand channel of the receiver (R audio). Point 3 is to be connected to Ring 2, linking it to the transducer (MIC+). Points 4 and 5 are to be connected to the sleeve, linking it to ground (GND).

Referring to Figure 6-2:

- 1 is the contact point of the tip, linking it to the left-hand channel of the receiver (L audio).
- 2 is the contact point of Ring 1, linking it to the right-hand channel of the receiver (R audio).
- 3 is the contact point of Ring 2, linking it to the transducer (MIC+).
- 4 is the contact point of the sleeve, linking it to the GND.
- 5 is the bushing of the socket, linking it to the GND when it is made of conductive material.

It is recommended that the headset pole order from the tip to the sleeve is: L/R/MIC/GND.

It is agreed that the pinout order of L/R/MIC/GND has an advantage in electrostatic discharge (ESD) protection and allows for both plastic and metallic converters.

### 6.2.2 Alternate pin assignment

Point 1 of the socket is to be connected to the tip of the plug, linking it to the left-hand channel of the receiver (L audio). Point 2 is to be connected to Ring 1, linking it to the right-hand channel of the receiver (R audio). Points 3 and 5 are to be connected to Ring 2, linking it to GND. Point 4 is to be connected to the sleeve, linking it the transducer (MIC+).

Referring to Figure 6-2:

- 1 is the contact point of the tip, linking it to the left-hand channel of the receiver (L audio).
- 2 is the contact point of Ring 1, linking it to the right-hand channel of the receiver (R audio).
- 3 is the contact point of Ring 2, linking it to the GND.
- 4 is the contact point of the sleeve, linking it to the transducer (MIC+).
- 5 is the bushing of the socket, linking it to the GND when it is made of conductive material.

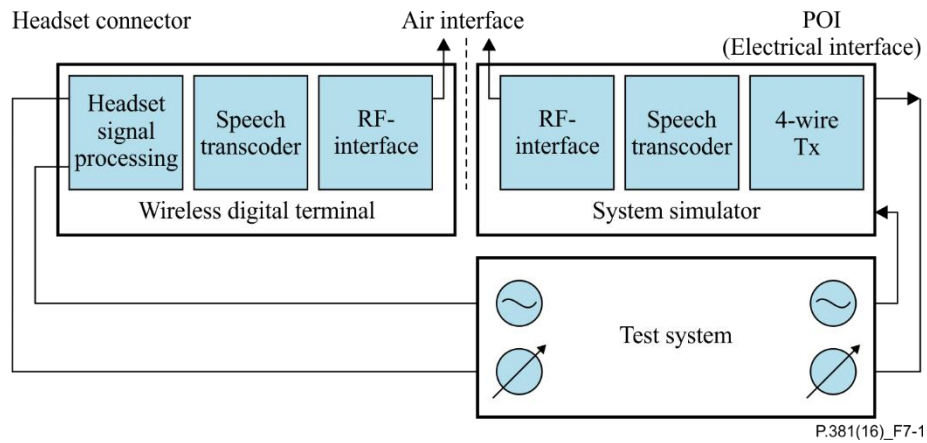
The pinout order from the tip to the sleeve for the alternate headset plug is: L/R/GND/MIC.

## 7 Electrical interface specification

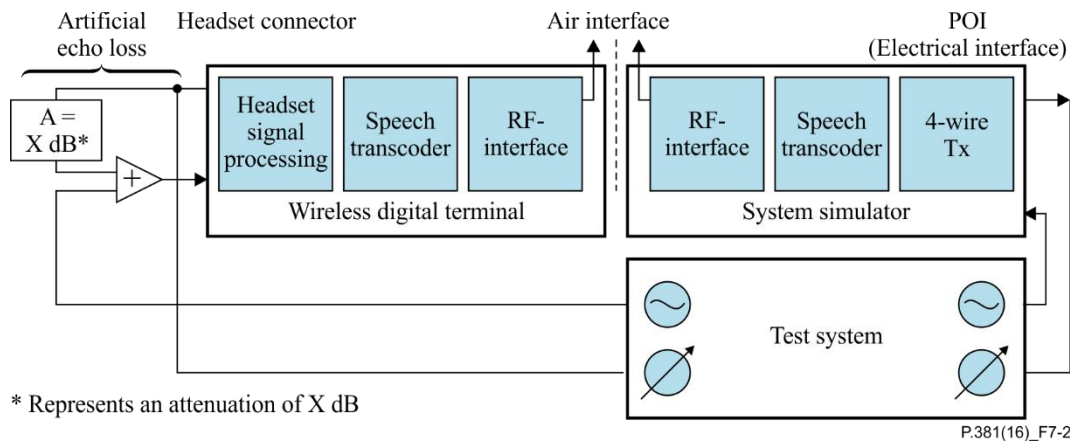
### 7.1 Communication mode

#### 7.1.1 Test set-up

Test set-ups are shown in Figures 7-1 and 7-2.



**Figure 7-1 – Test set-up for testing the electrical headset interface**



**Figure 7-2 – Test set-up with artificial echo loss for echo and double talk testing**

#### 7.1.1.1 Input and output characteristics of the test system for connecting to the headset connector

The output of the test system connected to the interface in Send of the headset connector must be DC resistant. The output impedance shall be between 1  $\Omega$  and 10 k $\Omega$ . The dynamic range of the test system shall be consistent with (or exceed) the level range provided by headset microphones.

The input of the test system connected to the receiving interfaces of the headset connectors shall have an input impedance of 32  $\Omega$ . The dynamic range shall be consistent with (or exceed) the output level range provided by the electrical output of digital mobile terminals' headset outputs.

The common ground impedance (between sending and receiving sides) for the test system shall be  $\leq 0.05 \Omega$ .

### 7.1.1.2 Test signals and test signal levels

Unless otherwise specified, full-band real speech signals, which can be found in [ITU-T P.501], are used for measurements. Detailed information about the test signal used can be found in the corresponding clause of [ITU-T P.501]. For test cases where composite source signals are specified, the speech-spectrum-shaped CSS signals specified in [ITU-T P.501] shall be used.

All test signals – which are used in Receive – have to be band-limited. The band limitation is achieved by bandpass filtering in the frequency range between 100 Hz and 4 kHz using a bandpass filter providing  $\geq 24$  dB/octave for narrowband mode. In wideband mode, the band limitation is achieved by bandpass filtering in the frequency range between 100 Hz and 8 kHz, using a bandpass filter providing  $\geq 24$  dB/octave. In Send, the test signals are used without band limitation.

For real speech, the test signal levels are referred to the [ITU-T P.56] active speech level of the (band limited in receiving direction) test signal, calculated over the complete test sequence, unless described otherwise. For other test signals, the test signal levels are referred to the average level of the (band limited in receiving direction) test signals, averaged over the complete test sequence length.

Unless stated otherwise, the nominal average signal levels for the measurements are as follows:

- $-16$  dBm0 in Receive.
- $-60$  decibels relative to 1 Volt (dBV) in Send (typical equivalent microphone signal level corresponding to  $-4.7$  dBPa at the mouth reference point (MRP)).
- The receive volume control is adjusted to the setting that produces the level closest to  $-39$  dBV considering binaural headsets.

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

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

### 7.1.2 Delay

#### 7.1.2.1 Requirements

In view of the following considerations:

- that delay has impact on echo performance and the dynamics of voice conversation;
- the amount of delay introduced by wireless systems depends on specific technology and may be inherent to the adopted coding technique.

The following is recommended:

- delay added by the terminal equipment should be minimized in accordance with the guidelines provided in [ITU-T G.114] even with the use of echo control;
- the terminal specific implementation dependent delay, including both the delay in sending direction and the delay in receiving direction, should be less than 70 ms. That means the sum of overall terminal delay in Send  $T_s$  and overall terminal delay in Receive  $T_r$ , should be less than  $70 \text{ ms} + T_{as} + T_{ar}$ . ( $T_{as}$  is the implementation independent system delay in Send and  $T_{ar}$  is the implementation independent system delay in Receive)

NOTE 1 – The overall terminal delay consists of the implementation independent system delay and the implementation dependent delay. The implementation independent system delay is introduced by the specific accessing technology in air interface and the coding technique in both sending and receiving directions. The implementation dependent delay is introduced by the speech processing, data transport/handling, speech enhancement, audio filtering, etc. in both sending and receiving direction, it is obtained by excluding implementation independent system delay from the measured overall terminal delay.

NOTE 2 – For 3GPP UMTS circuit-switched speech and 3GPP LTE MTSI-based speech, definitions, performance objectives and requirements are found in [b-3GPP TS 26.131].

### **7.1.2.2 Test**

#### **7.1.2.2.1 Test of overall terminal delay in Send**

- 1) The test signal to be used for the measurements shall be a composite source signal (CSS), as described in [ITU-T P.501]. The test signal consists of the voiced part followed by a pseudo random noise sequence with a minimum periodicity of 500 ms.
- 2) The test signal level shall be  $-60$  dBV, measured at the sending input of the headset interface.
- 3) The delay is calculated using the cross-correlation function between the signal at the terminal input and the signal at the system simulator output.
- 4) The measurement is corrected by the delay introduced by the test equipment ( $T_s$ , sys). The sending delay ( $T_s$ , wdt) is expressed in milliseconds, determined from the maximum of the cross-correlation function.

#### **7.1.2.2.2 Test of overall terminal delay in Receive**

- 1) The test signal to be used for the measurements shall be a composite source signal (CSS), as described in [ITU-T P.501]. The test signal consists of the voiced part followed by a pseudo random noise sequence with a minimum periodicity of 500 ms.
- 2) The test signal level shall be  $-16$  dBm0, measured at the digital reference point.
- 3) The delay is calculated using the cross-correlation function between the signal at the terminal input and the signal at the system simulator output. The measurement is corrected by the delay introduced by the test equipment ( $T_r$ , sys).
- 4) The receiving delay ( $T_r$ , wdt) is expressed in milliseconds, determined from the maximum of the cross-correlation function.

### **7.1.3 Level in Send for nominal speech input level**

#### **7.1.3.1 Requirements**

The sending level is measured at the point of interconnection (POI) (output of the reference speech decoder of the system simulator).

The sending level shall be  $-16$  dBm0  $\pm 3$  dB when inserting the sending signal at the nominal level, as described in clause 7.1.1.2.

#### **7.1.3.2 Test**

- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.
- 2) 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 the active speech level according to [ITU-T P.56].
- 3) The active speech level at the electrical reference point (POI) is measured.
- 4) The sending level is expressed in dBm0.

### **7.1.4 Level in Receive for nominal speech input level**

#### **7.1.4.1 Requirements**

The receiving level is measured at the receiving output of the headset interface.

The receiving level shall be  $-30$  dBV  $\pm 6$  dB at the maximum volume setting when inserting the receiving signal at nominal level, as described in clause 7.1.1.2.

The receiving level shall be  $-39 \text{ dBV} \pm 3 \text{ dB}$  at the nominal volume setting when inserting the receiving signal at nominal level, as described in clause 7.1.1.2.

The receiving level shall be  $-55 \text{ dBV} \pm 6 \text{ dB}$  at the minimum volume setting when inserting the receiving signal at nominal level, as described in clause 7.1.1.2.

#### 7.1.4.2 Test

- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.
- 2) 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 the active speech level according to [ITU-T P.56].
- 3) For the calculation, the active level at the sending output of the headset interface is used.
- 4) The receiving level is expressed in dBV.

The measurement is repeated for the second channel.

### 7.1.5 Level in Send for low and high speech input levels

#### 7.1.5.1 Requirements

The sending level is measured at the POI (output of the reference speech decoder of the system simulator). The test result is compared to the test result level in send for nominal speech input level ( $L_{\text{sendnom}}$ ) obtained for the nominal input level ( $L_{\text{nominal}}$ ) described in clause 7.1.3. The results should be according to Table 7-1.

**Table 7-1 – Limits for Level in Send for low and high speech input levels**

Input level [dBV]	Upper limit [dBm0]	Lower limit [dBm0]
$L_{\text{nominal}} - 10$	$L_{\text{sendnom}} - 5$	$L_{\text{sendnom}} - 12$
$L_{\text{nominal}} + 5$	$L_{\text{sendnom}} + 7$	$L_{\text{sendnom}} + 0$

#### 7.1.5.2 Test

- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.
- 2) 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 the active speech level according to [ITU-T P.56], adjusted as described in clause 7.1.5.1.
- 3) The active speech level at the electrical reference point (POI) is measured.
- 4) The sending level is expressed in dBm0 and compared to the limits described in clause 7.1.5.1.

#### 7.1.6 Linearity in Receive

(provisional, for further study)

### 7.1.7 Sending frequency response

#### 7.1.7.1 Requirements

The sending frequency response is measured from the sending input of the headset interface to the POI (input of the reference speech coder of the system simulators). Considering the special location of the headset microphone, the sending frequency response should provide an allowance for a high-frequency boost in order to comply with a large variety of headsets, which in combination with the digital mobile terminal should comply with the relevant standards in Send.



The measured frequency response shall be within the limits as defined in Table 7-2 for narrowband and Table 7-3 for wideband.

**Table 7-2 – Tolerance mask for the narrowband sending frequency response**

Frequency (Hz)	Upper limit	Lower limit	Target
100	-9	$-\infty$	-20
200	3	$-\infty$	-5
300	3	-3	0
1 000	5	-3	1
3 100	11	3	7
3 400	12	-5	4
4 000	12	$-\infty$	

NOTE 1 – All sensitivity values are expressed in dB on an arbitrary scale.  
 NOTE 2 – The limits for intermediate frequencies lie on a straight line drawn between the given values on a logarithmic (frequency) – linear (dB) scale.  
 NOTE 3 – The sending frequency response should have allowance for a high-frequency boost considering typical variations in mouth-to-microphone transfer characteristics.

**Table 7-3 – Tolerance mask for the wideband sending frequency response**

Frequency (Hz)	Upper limit	Lower limit	Target
100	-3	$-\infty$	-12
200	3	-3	0
1 000	3	-3	0
3 000	10	2	6
5 000	12	2	7
6 300	12	0	5
8 000	12	$-\infty$	

NOTE 1 – All sensitivity values are expressed in dB on an arbitrary scale.  
 NOTE 2 – The limits for intermediate frequencies lie on a straight line drawn between the given values on a logarithmic (frequency) – linear (dB) scale.  
 NOTE 3 – The sending frequency response should have allowance for a high-frequency boost considering typical variations in mouth-to-microphone transfer characteristics.

#### 7.1.7.2 Test

- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.
- 2) 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 the nominal signal level, applied to the sending input of the headset interface.  
 The power density spectrum at the sending input of the headset interface is used as the reference power density spectrum for determining the sending sensitivity.
- 3) In wideband, the sending sensitivity is determined in third octave intervals, as given by [IEC 61260-1] for frequencies from 100 Hz to 8 kHz inclusive, measured at the POI. In narrowband, it is determined for frequencies from 200 Hz to 4 kHz. In each third octave band, the level of the measured signal is referred to the level of the reference signal.

4) The sensitivity is determined in dBV/V.

## 7.1.8 Receiving frequency response

### 7.1.8.1 Requirements

The receiving frequency response is measured from the POI (output of the reference speech coder of the system simulators) to the receiving output of the headset interface. The receiving sensitivity response should be mostly flat in the entire frequency range in order to comply with a large variety of headsets, which in combination with the digital mobile terminal should comply with the relevant standards in Receive.

The measured frequency response shall be within the limits as defined in Table 7-4 for narrowband and Table 7-5 for wideband.

**Table 7-4 – Tolerance mask for the narrowband receiving frequency response**

Frequency (Hz)	Upper limit	Lower limit	Target
100	2	$-\infty$	–
200	2	$-\infty$	0
300	2	–6	0
1 000	2	–2	0
2 000	6	–2	–1
3 400	6	–5	–1.5
4 000	6	$-\infty$	–

NOTE 1 – All sensitivity values are expressed in dB on an arbitrary scale.  
 NOTE 2 – The limits for intermediate frequencies lie on a straight line drawn between the given values on a logarithmic (frequency) – linear (dB) scale.  
 NOTE 3 – The target response assumes the system under test uses the AMR-NB [b-3GPP TS 26.071] codec operating at 12.2 kbps.

**Table 7-5 – Tolerance mask for the wideband receiving frequency response**

Frequency (Hz)	Upper limit	Lower limit	Target
100	2	$-\infty$	0
200	2	–7	0
300	2	–5.5	0
1 000	2	–2	0
2 000	2	–2	–0.5
5 000	2	–6	–4
6 300	2	–12	–6
8 000	2	$-\infty$	–

NOTE 1 – All sensitivity values are expressed in dB on an arbitrary scale. The limits for intermediate frequencies lie on straight lines drawn between the given values on a linear (dB) – logarithmic (Hz) scale.  
 NOTE 2 – The target response assumes the system under test uses the AMR-WB [b-ITU-T G.722.2] codec operating at 12.65 kbps.

### 7.1.8.2 Test

- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.
- 2) 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 the nominal signal level, applied to the POI. The level is averaged over the complete test signal.
- 3) For wideband, the receiving sensitivity is determined in third octave intervals as given by [IEC 61260-1] for frequencies from 100 Hz to 8 kHz inclusive, measured at the headset interface. In narrowband, it is determined for frequencies from 200 Hz to 4 kHz. In each third octave band, the level of the measured signal is referred to the level of the reference signal, averaged over the complete test sequence length.
- 4) The sensitivity is determined in dBV/V.

The measurement is repeated for the second channel.

### 7.1.9 Sidetone loss

#### 7.1.9.1 Requirements

The talker sidetone masking rating (STMR) (electrical sidetone) is measured from the sending input of the headset interface to the receiving output of the headset interface.

The STMR shall be  $\geq 20$  dB and should be  $\leq 35$  dB for the nominal setting of the volume control. For all other positions of the volume control, the STMR shall be  $\geq 10$  dB.

NOTE 1 – Where a user-controlled receiving volume control is provided, it is recommended that the sidetone loss is independent of the volume control setting.

NOTE 2 – For connections with headsets where the human air-conducted sidetone paths are obstructed (one example being some binaural insert type headsets), it is important to provide a terminal electrical sidetone path.

#### 7.1.9.2 Test

- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.
- 2) 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 the nominal signal level, applied to the sending input of the headset interface. The level is averaged over the complete test signal.  
  
The measured power density spectrum at the sending input of the headset interface is used as the reference power density spectrum for determining the sidetone sensitivity.
- 3) Measurements shall be made at one twelfth-octave intervals as given by [IEC 61260-1] for frequencies from 100 Hz to 8 kHz inclusive. For the calculation the averaged measured level at each frequency band ([ITU-T P.79] Table 3, bands 1 to 20) is referred to the averaged test signal level in each frequency band.
- 4) The measured sensitivity is corrected by adding the nominal sensitivities ( $-55$  dBV/Pa +  $19$  dBPa/V =  $-36$  dB) of the headset thus transferring the measured electrical signal levels to their equivalent acoustical signal levels when assuming a headset with flat characteristics in sending and receiving (flat at ERP) directions.
- 5) The sidetone path loss (LmeST), as expressed in dB, and the STMR (in dB) shall be calculated from formula 5-1 of [ITU-T P.79], using  $m = 0.225$  and the weighting factors in Table 3 of [ITU-T P.79]. Leakage correction shall not be applied.

## 7.1.10 Sidetone delay

### 7.1.10.1 Requirements

The sidetone delay is measured from the sending input of the headset interface to the receiving output of the headset interface.

The maximum sidetone round-trip delay shall not exceed 5 ms in wideband, and should not exceed 5 ms in narrowband. The measured result is only applicable where the level of the electrical sidetone is sufficiently high to be measured.

### 7.1.10.2 Test

- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.
- 2) The test signal is a CS-signal complying with [ITU-T P.501] using a pn sequence with a length of 4 096 points (for the 48 kHz sampling rate) which equals to the period  $T$ . The duration of the complete test signal is as specified in [ITU-T P.501]. The test signal level is the nominal signal level, applied to the sending input of the headset interface.
- 3) The cross-correlation function  $\Phi_{xy}(\tau)$  between the input signal  $S_x(t)$  generated by the test system in Send and the output signal  $S_y(t)$  measured at the receiving output of the headset interface is calculated in the time domain:

$$\Phi_{xy}(\tau) = \frac{1}{T} \int_{t=-\frac{T}{2}}^{\frac{T}{2}} S_x(t) \cdot S_y(t + \tau) dt \quad (7-1)$$

The measurement window  $T$  shall be exactly identical to the time period  $T$  of the test signal, the measurement window is positioned to the pn-sequence of the test signal.

The sidetone delay is calculated from the envelope  $E(\tau)$  of the cross-correlation function  $\Phi_{xy}(\tau)$ . The first maximum of the envelope function occurs in correspondence with the test signal at the sending input of the headset interface, the second one occurs with a possible delayed sidetone signal. The difference between the two maxima corresponds to the sidetone delay. The envelope  $E(\tau)$  is calculated by the Hilbert transformation  $H\{xy(\tau)\}$  of the cross-correlation:

$$H\{xy(\tau)\} = \int_{u=-\infty}^{u=+\infty} \frac{\Phi_{xy}(u)}{\pi(\tau - u)} du \quad (7-2)$$

$$E(\tau) = \sqrt{[\Phi_{xy}(\tau)]^2 + [H\{xy(\tau)\}]^2} \quad (7-3)$$

It is assumed that the measured sidetone delay is  $< T/2$ .

## 7.1.11 Noise in Send

### 7.1.11.1 Requirements

The noise in Send is measured from the sending input of the headset interface to the POI (input of the reference speech coder of the system simulators).

The signal to noise ratio (SNR) shall be  $\geq 30$  dB.

Spectral peaks in the frequency domain shall not exceed the averaged spectrum by more than 10 dB.

### 7.1.11.2 Test

- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.

- 2) The British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501] at nominal signal level, as described in clause 7.1.1.2, is applied at the sending input of the headset interface. The output level is measured using a speech level voltmeter according to [ITU-T P.56]. This level is the reference speech signal level.
- 3) For the noise measurement, no test signal is used. However, all sources which potentially contribute to noise should be considered. Interference from radio frequencies is not accurately covered by an interface specification as the complete terminal/headset system needs to be assessed. Moreover, the necessary test system cabling is likely to introduce further deviations from real-life conditions. Therefore, radio induced noise is not expected to be accurately covered by the test cases in this Recommendation.
- 4) The noise is measured at the output in the frequency range between 100 Hz and 4 kHz for narrowband and between 100 Hz and 8 kHz for wideband. The length of the time window is 1 s which is the averaging time for the idle channel noise. The power density spectrum of the noise signal is determined using the fast Fourier transform (FFT) (8 k samples/48 kHz sampling rate or equivalent). A Hann window is used.
- 5) The noise is determined by A-weighting [IEC 61672-1] and referring to the reference speech signal level as determined with the speech sequence.
- 6) Spectral peaks are measured in the frequency domain. The frequency spectrum of the 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). 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 from 100 Hz to 3.4 kHz in narrowband and from 100 Hz to 6.3 kHz in wideband.

## 7.1.12 Noise in Receive

### 7.1.12.1 Requirements

The noise in Receive is measured from the receiving output of the headset interface.

The SNR shall be higher than 30 dB at volume setting according to 7.1.1.2,

The SNR shall be higher than [39] dB at maximum volume setting,

Spectral peaks in the frequency domain shall not exceed the averaged spectrum by more than 10 dB.

### 7.1.12.2 Test

- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.
- 2) The British-English single talk sequence described in clause 7.3.2 [ITU-T P.501] at the nominal signal level, as described in clause 7.1.1.2, is applied at the POI. The output level is measured using a speech level voltmeter according to [ITU-T P.56]. This level is the reference speech signal level.
- 3) For the noise measurement, no test signal is used. However, all sources which potentially contribute to noise should be considered. Interference from radio frequencies is not accurately covered by an interface specification as the complete terminal/headset system needs to be assessed. Moreover, the necessary test system cabling is likely to introduce further deviations from real-life conditions. Therefore, radio induced noise is not expected to be accurately covered by the test cases in this Recommendation.
- 4) The noise is measured at the output in the frequency range between 100 Hz and 8 kHz. The length of the time window is 1 s 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 Hann window is used.

- 5) The noise is determined by A-weighting [IEC 61672-1] and referring to the reference speech signal level as determined with the speech sequence.
- 6) Spectral peaks are measured in the frequency domain. The frequency spectrum of the 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). 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 from 100 Hz to 3.4 kHz in narrowband and from 100 Hz to 6.3 kHz in wideband.

The measurement is repeated for the second channel.

### 7.1.13 Sending distortion

#### 7.1.13.1 Requirements

The distortion in Send is measured from the sending input of the headset interface to the POI (output of the reference speech decoder of the system simulator).

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

**Table 7-6 – Limits for the signal to harmonic distortion**

Frequency (Hz)	Signal to harmonic distortion ratio limit, Send (dB)
315	30
400	40
1 000	40

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

#### 7.1.13.2 Test

- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.
- 2) For the test, a sinusoidal signal at frequencies of 315 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz and 1 000 Hz is used. The duration of the sine wave shall be < 1 s. The sinusoidal signal level shall be the nominal signal level. In order to ensure a reliable activation, a conditioning sequence is inserted before the actual measurement. The conditioning sequence is according to clause 7.3.7 of [ITU-T P.501]. The short conditioning sequences (either male or female) should be used. The level of the activation signal is the nominal signal level.
- 3) The signal to harmonic distortion ratio is measured selectively up to 3.5 kHz for narrowband and 7 kHz for wideband.
- 4) The test is repeated using a signal level 10 dB higher than the nominal signal level. The level of the activation signal is kept at the nominal signal level.

### 7.1.14 Receive distortion

#### 7.1.14.1 Requirements

The distortion in Receive is measured from the receiving output of the headset interface.

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

**Table 7-7 – Limits for the signal to harmonic distortion**

Frequency (Hz)	Signal to harmonic distortion ratio limit, receive (dB)
315	30
400	40
1 000	40

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

#### **7.1.14.2 Test**

- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.
- 2) For the test, a sinusoidal signal at frequencies of 315 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz and 1 000 Hz is used. The duration of the sine wave shall be < 1 s. The sinusoidal signal level shall be the nominal signal level. In order to ensure a reliable activation, a conditioning sequence is inserted before the actual measurement. The conditioning sequence is according to clause 7.3.7 of [ITU-T P.501]. The short conditioning sequences (either male or female) should be used. The level of the activation signal is the nominal signal level.
- 3) The signal to harmonic distortion ratio is measured selectively up to 7 kHz.
- 4) The test is repeated using a signal level 10 dB higher than the nominal signal level. The level of the activation signal is kept at the nominal signal level.

#### **7.1.15 Noise cancellation test in Send**

##### **7.1.15.1 Requirements**

(provisional, for further study)

The noise cancellation in Send is measured from the sending input of the headset interface to the POI (input of the reference speech coder of the system simulators).

The objective of this test is to check the performance of the noise cancellation in Send.

When testing through the objective methodology, the terminal shall comply with the following requirements:

For narrowband terminals:

N-MOS-LQOn      Average N-MOS-LQOn  $\geq$  tbd

S-MOS-LQOn      Average S-MOS-LQOn  $\geq$  tbd

G-MOS-LQOn      Average G-MOS-LQOn  $\geq$  tbd

For wideband terminals:

N-MOS-LQOw      Average N-MOS-LQOw  $\geq$  tbd

S-MOS-LQOw      Average S-MOS-LQOw  $\geq$  tbd

G-MOS-LQOw      Average G-MOS-LQOw  $\geq$  tbd

##### **7.1.15.2 Test**

(provisional, for further study)

In order to create a representative electrical test signal for the headset interface containing speech at a nominal level mixed with the amount of background noise picked up by a representative headset, the set-up in [b-ETSI EG 202 396-1] is used. The representative headset is connected to a reference interface providing nominal properties for the electrical interface, as described in clause 7.1.1.1. The

signal (speech plus noise) is recorded at this interface and inserted through the appropriate reference interfaces, as described in clause 7.1.1.1, in such a way that the signal level and spectral content delivered to the terminal under test is equivalent to the one it would have seen if the headset was connected directly. Either headsets considered to be representative for the type of headset attached to the terminal are used or individual headsets are used. In addition, the unprocessed speech plus noise signal is recorded at the headset microphone position using a reference microphone positioned close to the headset microphone.

- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.
- 2) S-MOS, N-MOS and G-MOS are calculated as described in [ETSI TS 103 106]. The speech level is  $-1.7$  dBPa at the MRP.
- 3) The test signal is applied to the headset interface.
- 4) A proper conditioning sequence should be used in advance of the measurement, as described in [ETSI TS 103 106].
- 5) The noise types as described in Table 7-8 shall be used. See [b-ETSI EG 202 396-1].
- 6) The measurement over the eight noise types shall be made in the same unique and dedicated call and not in the same call as, for example, the one established for acoustic measurement. The noise types shall be presented according to the order specified in Table 7-8.

**Table 7-8 – Noises used for background noise simulation**

Description	File name	Duration	Level	Type
Recording in pub	Pub_Noise_binaural_V2	30 s	L: 75,0 dB(A) R: 73,0 dB(A)	Binaural
Recording at pavement	Outside_Traffic_Road_binaural	30 s	L: 74,9 dB(A) R: 73,9 dB(A)	Binaural
Recording at pavement	Outside_Traffic_Crossroads_binaural	20 s	L: 69,1 dB(A) R: 69,6 dB(A)	Binaural
Recording at departure platform	Train_Station_binaural	30 s	L: 68,2 dB(A) R: 69,8 dB(A)	Binaural
Recording at the drivers position	Fullsize_Car1_130Kmh_binaural	30 s	L: 69,1 dB(A) R: 68,1 dB(A)	Binaural
Recording at sales counter	Cafeteria_Noise_binaural	30 s	L: 68,4 dB(A) R: 67,3 dB(A)	Binaural
Recording in a cafeteria	Mensa_binaural	22 s	L: 63,4 dB(A) R: 61,9 dB(A)	Binaural
Recording in business office	Work_Noise_Office_Callcenter_binaural	30 s	L: 56,6 dB(A) R: 57,8 dB(A)	Binaural

## 7.1.16 One-way speech quality in Send

### 7.1.16.1 Requirements

The listening speech quality in Send LQs is measured from the sending input of the headset interface to the POI (input of the reference speech coder of the system simulators). The speech processing prior to the encoder in the terminal shall be disabled for this test.

The listening speech quality in Send shall be

MOS-LQOs > [TBD] in narrowband

MOS-LQOs > [TBD] in wideband



NOTE 1 –The purpose of this test is limited to verification of a proper implementation of the speech encoding operation. Speech processing in the terminal is necessary to compensate a number of acoustic aspects. The processing, while used to provide a suitable user experience, may not result in the highest MOS-LQOs score for a given speech codec operating point. Examples of such processing include, but are not limited to: (1) filtering to compensate for acoustic path loss from the MRP to the microphone, (2) Use of automatic gain control to compensate for soft talkers, loud talkers, variations of positioning, (3) Presence of noise suppression.

NOTE 2 – The MOS-LQOs values depend on speech codec, its operating point and the particular speech signal used

### **7.1.16.2 Test**

The test methods to be used are described in [ITU-T P.863].

NOTE –As recommended in [ITU-T P.863], both narrowband and wideband systems are evaluated on a super-wideband scale. Therefore the mean opinion score (MOS) requirements are given in MOS-LQOs. See [ITU-T P.863] for more information.

- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.
- 2) The test signals used are the British-English test sequences as specified in [ITU-T P.501] Annex C (two male speakers, two female speakers, two sentences each). The test signal level is the nominal signal level. The test signal level is measured as "active speech level" according to [ITU-T P.56]. The speech activity should be between 30% and 70%. Level pre-alignment to -26 dBoV of recordings shall be used – see [ITU-T P.863.1] clause 10.2.

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

- 3) The test arrangement is according to clause 7.1.1.

The calculation is made using the signal recorded at the POI. The MOS-LQOs calculation is performed on a sentence-pair basis and the average score is calculated.

## **7.1.17 One-way speech quality in Receive**

### **7.1.17.1 Requirements**

The listening quality in the Receive direction ( $LQ_R$ ) is measured from the POI (output of the reference speech coder of the system simulators) to the receiving output of the headset interface.

The listening speech quality in Receive shall be:

MOS-LQOs > [TBD] in narrowband

MOS-LQOs > [TBD] in wideband

NOTE 1 –The purpose of this test is limited to verification of a proper implementation of the speech encoding operation. Speech processing in the terminal is necessary to compensate a number of acoustic aspects. The processing, while used to provide a suitable user experience, may not result in the highest MOS-LQOs score for a given speech codec operating point. Examples of such processing include, but are not limited to: (1) filtering to compensate for acoustic path loss from the MRP to the microphone, (2) Use of automatic gain control to compensate for soft talkers, loud talkers, variations of positioning, (3) Presence of noise suppression,

NOTE 2 – The MOS-LQOs values depend on speech codec, its operating point and the particular speech signal used

### **7.1.17.2 Test**

The test methods to be used are described in [ITU-T P.863].

NOTE –As recommended in [ITU-T P.863], both narrowband and wideband systems are evaluated on a super-wideband scale. Therefore the MOS requirements are given in MOS-LQOs. See [ITU-T P.863] for more information.

- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.

- 2) The test signals used are the British-English test sequences as specified in [ITU-T P.501] Annex C (two male speakers, two female speakers, two sentences each). The test signal level is the nominal signal level. The test signal level is measured as "active speech level" according to [ITU-T P.56]. The speech activity should be between 30% and 70%. Level pre-alignment to  $-26$  dBov of recordings shall be used – see [ITU-T P.863.1] clause 10.2.

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

- 3) The test arrangement is according to clause 7.1.1. The signal measured at the headset interface is used for the calculation. The MOS-LQOs calculation is performed on a sentence-pair basis and the average score is calculated.

The measurement is repeated for the second channel.

### **7.1.18 Terminal coupling loss**

#### **7.1.18.1 Requirements**

The weighted terminal coupling loss (TCL<sub>w</sub>) is measured from the POI (input of the reference speech coder of the system simulator) to the POI (output of the reference speech coder of the system simulator).

The TCL<sub>w</sub> provided by the headset signal processing in conjunction with typical echo paths, as described in Figure 7-2, shall be  $\geq 46$  dB at the volume control setting according to clause 7.1.1.2. The TCL<sub>w</sub> shall be also  $\geq 46$  dB at maximum setting of volume control.

NOTE – A TCL<sub>w</sub>  $\geq 50$  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  $\geq 50$  dB.

#### **7.1.18.2 Test**

- 1) The test arrangement is according to clause 7.1.1. For the test, an artificial echo path is inserted as shown in Figure 7-2. The test is performed using an artificial echo loss of 40dB.
- 2) The attenuation between the input of the test point (POI) and the output of the test point (POI) is determined.
- 3) The test signal is the compressed real speech signal described in clause 7.3.3 of [ITU-T P.501]. The signal level shall be  $-10$  dBm0.
- 4) 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).
- 5) TCL<sub>w</sub> is calculated according to clause B.4 of [ITU-T G.122] (trapezoidal pseudo rule). For the calculation, the average measured echo level at each frequency band is referred to the average level of the test signal measured in each frequency band. For the measurement, a time window has to be applied which is adapted to the duration of the actual test signal (250 ms).

### **7.1.19 Temporal echo effects**

(provisional, for further study)

#### **7.1.19.1 Requirements**

Temporal echo effects are measured from the POI (input of the reference speech coder of the system simulator) to the POI (output of the reference speech coder of the system simulator).

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

### 7.1.19.2 Test

- 1) The test arrangement is according to clause 7.1.1, Figure 7-2. The test is performed using an artificial echo loss of 40dB.
- 2) The test signal consists of the 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 test signal shall be applied for at least 5.0s before starting the analysis in order to allow full convergence. The echo signal is analysed during a period of at least 2.8 s, which represents eight periods of the CS signal. The integration time for the level analysis shall be 35 ms, the analysis is referred to the level analysis of the reference signal. The TCL variation is compared to the maximum TCL achieved in the test.
- 3) The measurement result is displayed as attenuation vs. time. The exact synchronization between the input and output signal has to be guaranteed.

NOTE 1 – In addition, tests with speech signals, as described in [ITU-T P.501], should be carried out to see the time variant behaviour of the echo canceller (EC). However, for such tests the simple broadband attenuation based test principle, as described above, cannot be applied due to the time varying spectral content of the speech-like signals.

NOTE 2 – The analysis is conducted only during the active signal part, the pauses between the composite source signals are not analysed. The analysis time is reduced by the integration time of the level analysis (hang over time due to 35 ms integration time).

### 7.1.20 Double talk performance

(provisional, for further study)

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" (TEL<sub>Rdt</sub>) should be high and the attenuation inserted should be as low as possible. Terminals which do not allow double talk in any case should provide a good echo attenuation which is realized by a high attenuation range in this case.

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

- attenuation range in 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.

The double talk performance may be highly influenced by the performance of the echo canceller (EC), especially by the non-linear processing (NLP) implementation.

#### 7.1.20.1 Attenuation range in the Send direction during double talk

(provisional, for further study)

##### 7.1.20.1.1 Requirements

Based on the level variation in  $A_{H,S,dt}$  the behaviour of the headset signal processing can be classified according to Table 7-9.

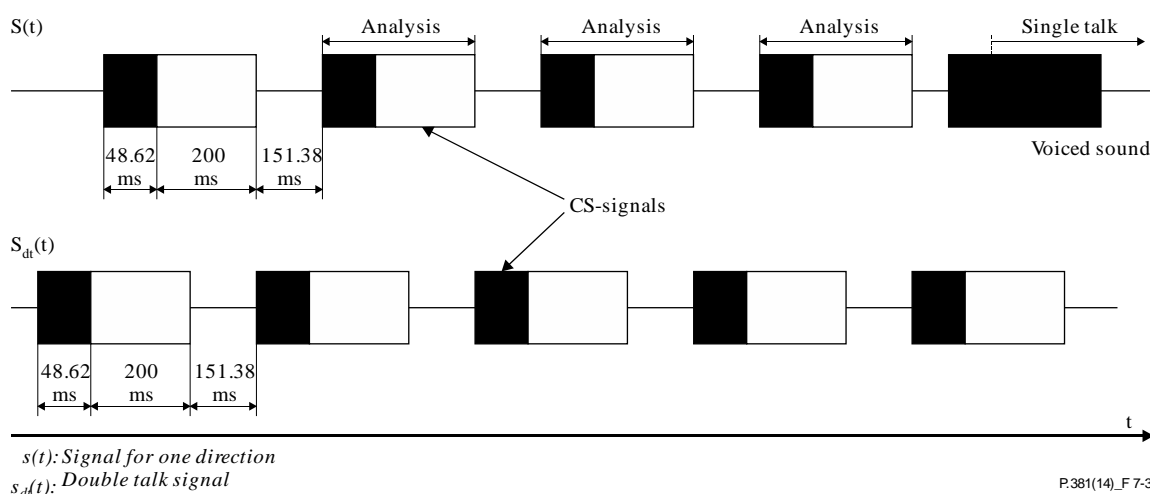
**Table 7-9 – Categorization of double talk capability according to [ITU-T P.340]**

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

In general, Table 7-9 provides a quality classification of the headset signal processing 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.

### 7.1.20.1.2 Test

The test signal to determine the attenuation range during double talk is shown in Figure 7-3. A sequence of uncorrelated CS signals is used which is inserted in parallel in both Send and Receive.



**Figure 7-3 – Double talk test sequence with overlapping CS signals in Send and Receive**

Figure 7-3 indicates that the sequences overlap partially. The beginning of the CS sequence (voiced sound, black) is overlapped by the end of the pn-sequence (white) of the opposite direction. During the active signal parts of one signal the analysis can be conducted in Send and Receive. The analysis times are shown in Figure 7-3 as well. The test signals are synchronized in time at the headset interface. The delay of the test arrangement should be constant during the measurement.

NOTE – The length of voiced sound of the double talk signal is achieved by repeating one period of the voiced sound for double talk according to [ITU-T P.501] ten times and cutting off the initial 3.3 ms of the period of the first voiced sound.

The settings for the test signals are as follows.

**Table 7-10 – Signal levels for double talk tests in Send and Receive**

	<b>Receive (sdt(t))</b>	<b>Send (s(t))</b>
Pause length between two signal bursts	151.38 ms	151.38 ms
Average signal level (Assuming an original pause length of 101.38 ms)	-16 dBm0	-60 dBV
Active signal parts	-14.7 dBm0	-58.7 dBV

- 1) The test arrangement is according to clause 7.1.1, Figure 7-2; the test signal is shown in Figure 7-3. The test is performed using an artificial echo loss of 40 dB.
- 2) When determining the attenuation range in Send, the signal measured at the POI is referred to the test signal inserted.
- 3) The level is determined as level vs time from the time domain. The integration time of the level analysis is 5 ms. The attenuation is determined from the level difference measured at the beginning of the double talk always with the beginning of the CS-signal in Send until its complete activation (during the pause in the receiving channel). The analysis is performed over the complete signal starting with the second CS-signal. The first CS-signal is not used for the analysis.
- 4) The categorization is made according to Appendix III of [ITU-T P.502].

#### **7.1.20.2 Attenuation range in the Receive direction during double talk**

(provisional, for further study)

##### **7.1.20.2.1 Requirements**

Based on the level variation in  $A_{H,R,dt}$  the behaviour of the headset signal processing can be classified according to Table 7-11.

**Table 7-11 – Categorization of double talk capability according to [ITU-T P.340]**

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

In general, Table 7-11 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.

##### **7.1.20.2.2 Test**

The test signal to determine the attenuation range during double talk is shown in Figure 7-2. A sequence of uncorrelated CS signals is used which is inserted in parallel in Send and Receive. The test signals are synchronized in time at the POI. The delay of the test arrangement should be constant during the measurement.

The settings for the test signals are as follows:

**Table 7-12 – Signal levels for double talk tests in Send and Receive**

	<b>Receive (s(t))</b>	<b>Send (s(t))</b>
Pause length between two signal bursts	151.38 ms	151.38 ms
Average signal level (Assuming an original pause length of 101.38 ms)	-16 dBm0	-60 dBV
Active signal parts	-14.7 dBm0	-58.7 dBV

- 1) The test arrangement is according to clause 7.1.1, Figure 7-2; the test signal is shown in Figure 7-5.
- 2) When determining the attenuation range in Receive the signal measured at the sending interface referred to the test signal inserted.
- 3) The level is determined as level vs time from the time domain. The integration time of the level analysis is 5 ms. The attenuation is determined from the level difference measured at the beginning of the double talk always with the beginning of the CS-signal in Receive until its complete activation (during the pause in the sending channel). The analysis is performed over the complete signal starting with the second CS-signal. The first CS-signal is not used for the analysis.
- 4) The categorization is made according to [ITU-T P.502], Appendix III.

### 7.1.20.3 Detection of echo components during double talk

(provisional, for further study)

#### 7.1.20.3.1 Requirements

Echo loss (EL) during double talk is the echo suppression provided by the headset signal processing during double talk measured at the receiving interface.

NOTE – The echo attenuation during double talk is based on the parameter TELRdt. It is assumed that the terminal at the opposite end of the connection provides the nominal loudness rating (SLR + RLR = 10 dB).

Under these conditions the requirements given in Table 7-13 are applicable (more information can be found in Annex A of [ITU-T P.340]).

**Table 7-13 – Categorization of double talk capability according to [ITU-T P.340]**

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

#### 7.1.20.3.2 Test

- 1) The test arrangement is according to clause 7.1.1, Figure 7-2.
- 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. A detailed description can be found in [ITU-T P.501]. For narrowband, the narrowband test signals are used; for wideband, the wideband test signals as described in [ITU-T P.501] are used.

- 3) The signals are fed simultaneously in Send and Receive. The level in Send at the headset interface is  $-60$  dBV (nominal level), the level in Receive at the POI is  $-16$  dBm0 (nominal level).
- 4) The test signal is measured at the receiving interface. The measured signal consists of the double talk signal which was fed in at the sending interface and the echo signal. The echo signal is filtered by comb filter using mid-frequencies and a bandwidth according to the signal components of the signal in Receive (see [ITU-T P.501]). The filter will suppress frequency components of the double talk signal.
- 5) For each frequency band which is used in Receive the echo attenuation can be measured separately. The requirement for category 1 is fulfilled if in any frequency band the echo signal is either below the signal noise or below the required limit. If echo components are detectable, the classification is based on Table 7-12. The echo attenuation is to be achieved for each individual frequency band according to the different categories.

### 7.1.21 Activation in Send

The activation in Send is mainly determined by the minimum built-up time in send ( $T_{r,S,min}$ ) and the minimum activation level in the Send direction ( $L_{S,min}$ ). The minimum activation level is the level required to remove the inserted attenuation in Send 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 headset interface.

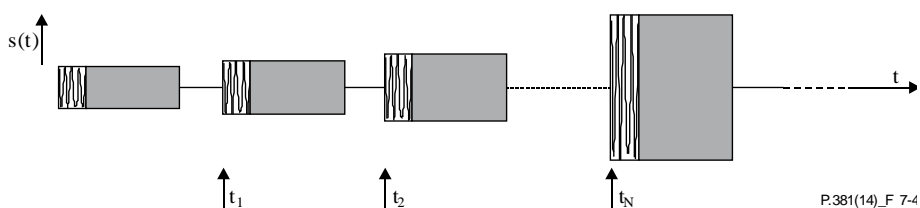
#### 7.1.21.1 Requirements

$L_{S,min}$  should be  $\leq -75$  dBV when analyzing a complete CSS burst

$L_{S,minonset}$  should be  $\leq$  [TBD] dBV when analyzing the onset of a CSS burst (first 50 ms)

#### 7.1.21.2 Test

The structure of the test signal is shown in Figure 7-4. The test signal consists of composite source signal (CSS) components according to [ITU-T P.501], clause 5.2.1.2 with increasing levels for each CSS burst.



**Figure 7-4 – Test signal to determine the minimum activation level and the build-up time**

The settings of the test signal are as shown in Table 7-14:

**Table 7-14 – Settings of the CSS in Send**

	<b>CSS duration/ minimum pause duration</b>	<b>Level of the first CS signal (active signal part at the sending input of the headset interface)</b>	<b>Level difference between two periods of the test signal</b>
CSS to determine switching characteristic in Send	248.62 ms/451.38 ms	-78.3 dBV (Note)	1 dB
NOTE – The level of the active signal part corresponds typically to an average level of -38.2 dBV at the headset interface 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.

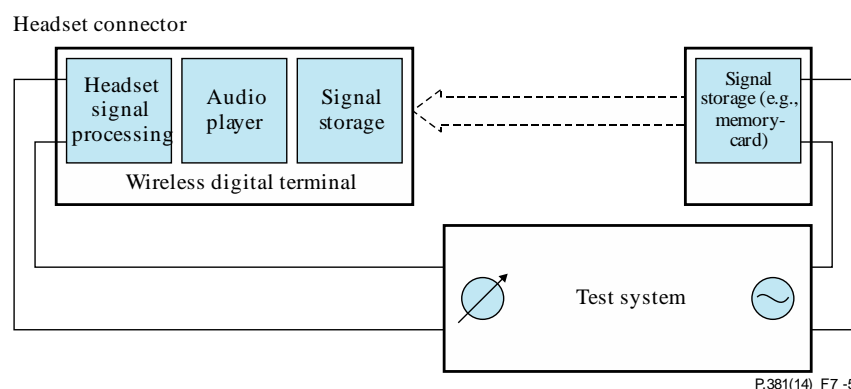
- 1) The test arrangement is according to clause 7.1.1, Figure 7-1.
- 2) The level of the transmitted single CSS burst is measured at the POI over a 400 ms time window. The level at POI is first measured with the nominal input level. The level at POI is then measured again for each input level. The inserted attenuation is calculated by referring the measured signal level to the firstly measured output level (nominal case).
- 3) The minimum activation level is determined from the CSS burst which has less than 6 dB attenuation compared to the nominal case.
- 4) The activation time is measured indirectly by repeating the complete test described above (from point 1) but with a time window covering only the first 50 ms of the CSS bursts.

NOTE – If the measurement using the CS signal does not allow the minimum activation level to be clearly identified, the measurement may be repeated by using the one syllable word, as described in [ITU-T P.501], 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.

## 7.2 Multimedia playback mode

### 7.2.1 Test set-up

The test set-up is shown in Figure 7-5.



**Figure 7-5 – Test arrangement for testing the electrical headset interface**



### **7.2.1.1 Input and output characteristics of the test system for connecting to the headset connector**

The input of the test system connected to the Receive interfaces of the headset connectors shall have an input impedance of 32 Ohm. The dynamic range shall be consistent with (or exceed) the output level range provided by the electrical output of digital mobile terminals' headset outputs.

In case the microphone stays connected during the tests the output of the test system connected to the sending interface of the headset connector must be DC resistant. The output impedance shall be between 1  $\Omega$  and 10 k $\Omega$ . The dynamic range shall be consistent with (or exceed) the level range provided by headset microphones.

### **7.2.1.2 Test signals and test signal levels**

For multimedia playback, the test signals have to be downloaded in the appropriate format (e.g., \*.wav, \*.mp3, \*.aac) for the phone under test. All test signals used are in the 16 bit \*.pcm format and then coded into the appropriate format. All signal levels stated in this section are relative to decibels relative to full scale (dBFS), where 0 dBFS represents the root mean square (RMS) level of a full-scale sinusoidal.

Programme simulation noise as defined in [IEC 60268-1] is used for the measurements. Detailed information about the test signal used can be found in the corresponding clause of this Recommendation.

Artificial test signals – which are used in receive – have to be band-limited. The band limitation is achieved by a low-pass filter up to 22 kHz providing  $\geq 24$  dB/octave filter roll-off. The programme simulation noise according to [EN 50332] is band limited by design and requires no filtering.

All test signal levels are referred to the average level of the test signals, averaged over the complete test sequence length, unless described otherwise.

The nominal average signal level for the measurements is  $-23$  dBFS.

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

## **7.2.2 Output level in multimedia playback mode**

### **7.2.2.1 Requirements**

The level is measured as the output level generated by the output of the mobile terminal player when playing back the pre-recorded signal at the output of the headset interface.

The output level shall be  $-22.5$  dBV  $\pm 6$  dB at a maximum volume setting when playing programme simulation noise at  $-10$  dBFS.

NOTE – Considering exposure times associated with music listening, acoustic safety standards/regulations may require deployment of further measures to inform about and/or reduce the risk of hearing damage.

### **7.2.2.2 Test**

- 1) The test arrangement is according to clause 7.2.1, Figure 7-5.
- 2) The test signal used for the measurements shall be programme simulation noise providing sufficient signal energy to 22 kHz. The test signal level is  $-10$  dBFS. Volume control as well as tone controls and other sound effects are set to produce maximum electrical output level.
- 3) For the calculation, the averaged level at the output of the headset interface is used. The output level is determined up to 22 kHz.

For the calculation, the average signal level measured at the output of the headset interface is used.

- 4) The output level is expressed in dBV.

The measurement is repeated for the second channel.

### 7.2.3 Frequency response in multimedia playback mode

#### 7.2.3.1 Requirements

The frequency response is measured as the output level generated by the output of the mobile terminal player when playing back the pre-recorded music signal. The frequency response should be mostly flat in the entire frequency range in order to comply with a large variety of headsets which in combination with the digital mobile terminal should comply with the relevant standards in Receive.

The measured frequency response shall be within the limits as defined in Table 7-15.

**Table 7-15 – Tolerance mask for frequency response in multimedia playback mode**

Frequency (Hz)	Upper limit	Lower limit
50	2	-2
12 000	2	-2
16 000	2	-5

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

#### 7.2.3.2 Test

- 1) The test arrangement is according to clause 7.2.1, Figure 7-5.
- 2) The test signal used for the measurements shall be a full-band music signal providing sufficient signal energy up to 22 kHz. The test signal is the nominal signal level. The level is averaged over the complete test signal.
- 3) The frequency response is determined in third octave intervals as given by [IEC 61260-1] for frequencies between 50 Hz and 16 kHz, inclusive. In each third octave band, the level of the measured signal is referred to the level of the reference signal (downloaded to the signal storage of the mobile terminal), averaged over the complete test sequence length.
- 4) The sensitivity is determined in dBV/V.

The measurement is repeated for the second channel.

### 7.2.4 Noise in multimedia playback mode

#### 7.2.4.1 Requirements

The noise is measured as the output level generated by the output of the mobile terminal player when playing back the programme simulation noise at the output of the headset interface and referring this to the idle channel noise produced when playing back a dithering noise signal.

The SNR shall be  $\geq 40$  dB.

Noise spectral peaks in the frequency domain shall not exceed the averaged spectrum by more than 10 dB.

#### 7.2.4.2 Test

- 1) The test arrangement is according to clause 7.2.1, Figure 7-5.
- 2) Programme simulation noise providing sufficient signal energy up to 22 kHz at the nominal signal level is used for playback. The test signal level is the A-weighted average level of the

complete test signal. The output level is measured as an unweighted broadband signal level up to 22 kHz. This level is the reference signal level.

- 3) For the noise measurement, a dithering noise signal with a stochastically varying lowest significant bit (LSB) is used for playback.
- 4) The idle channel noise is measured at the output in the frequency range up to 22 kHz. The length of the time window is 1 s which is the averaging time for the idle channel noise. The test lab has to ensure the correct activation of the device under test (DUT) during the measurement.

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

- 5) The noise is determined by A-weighting [IEC 61672-1] and referring to the reference signal level as determined with the music signal. Spectral peaks are measured in the frequency domain. The average noise spectrum used for determining the spectral peak should be calculated as the arithmetic mean of the noise spectrum values when stated in dBV.

The measurement is repeated for the second channel.

## 7.2.5 Distortion in multimedia playback mode

### 7.2.5.1 Requirements

The distortion is measured as harmonic distortion generated by the output of the mobile terminal player when playing back the pre-recorded sinusoidal signal at the output of the headset interface.

The ratio of signal to harmonic distortion shall be above the following mask as shown in Table 7-16.

**Table 7-16 – Limits for the signal to harmonic distortion**

Frequency (Hz)	Signal to harmonic distortion ratio limit, Send (dB)
100	40
315	50
5 000	50

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

### 7.2.5.2 Test

- 1) The test arrangement is according to clause 7.2.1, Figure 7-5.
- 2) For the test, a sinusoidal signal at frequencies of 100 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz, 1 000 Hz, 2 000 Hz and 5 000 Hz is used. The duration of the sine wave shall be < 1 s. The sinusoidal signal level shall be the nominal signal level.
- 3) The signal to harmonic distortion ratio is measured selectively up to 20 kHz.
- 4) The test is repeated using a signal level 20 dB higher than the nominal signal level.

The measurement is repeated for the second channel.

## 7.2.6 Receiving crosstalk

### 7.2.6.1 Requirements

The receiving crosstalk is measured as the L-R crosstalk and the R-L crosstalk generated by the output of the mobile terminal player when playing back programme simulation noise and measuring the resulting level at the two output channels of the headset interface.

In duration of 0-5 s, the attenuation measured at the right output, referenced to the spectrum generated at the left output shall be above 30 dB. In duration of 5-10 s, the attenuation measured at the left output, referenced to the level at the right output shall be also above 30 dB.

NOTE – the crosstalk attenuation of  $\geq 30$  dB should be kept over the frequency range from 50 Hz to 16 000 Hz.

**Table 7-17 – Signal sequence for the L-R and R-L crosstalk**

level period	Left channel (dBFS)	Right channel (dBFS)
0-5 s	-10	0
5-10 s	0	-10

### 7.2.6.2 Test

- 1) The test arrangement is according to clause 7.2.1, Figure7-5.
- 2) The test signal used for the measurements shall be programme simulation noise. The test signal is the nominal signal level. The level is averaged over the complete test signal. The signal sequence is shown in Table 7-16.
- 3) Playback the test signal, the crosstalk is determined by analysing the measured signal at the output of the headset interface. In duration of 0-5 s, the measured level at the right channel is referred to the level at the left channel, and the attenuation is L-R crosstalk. In duration of 5-10 s, the measured level at the left channel is referred to the level at the right channel, and the attenuation is R-L crosstalk.
- 4) The crosstalk is determined in dBV/V.

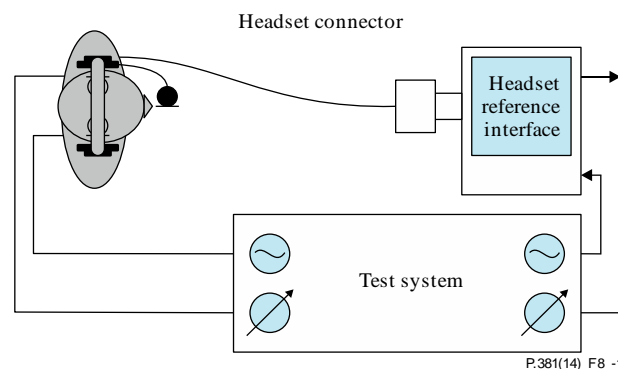
## 8 Headset specification

### 8.1 Communication mode

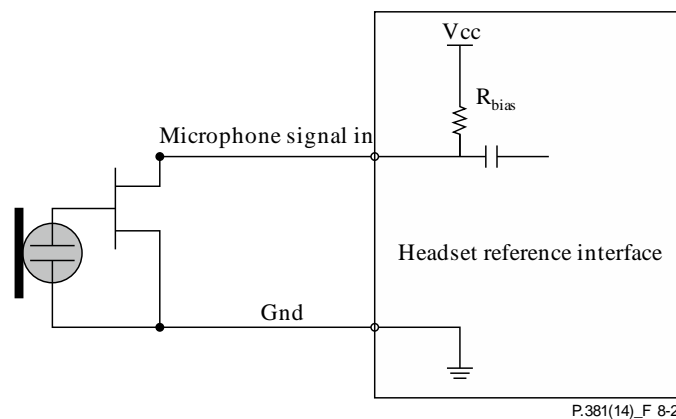
The receiving requirements of multimedia playback mode of headset are applied in communication mode here.

#### 8.1.1 Test set-up

The test set-up is shown in Figure 8-1.



**Figure 8-1 – Test arrangement for testing the headset**



**Figure 8-2 – Input connection of a headset reference interface**

### 8.1.1.1 Input and output characteristics of the test system for connecting the headset

The output of the test system must fulfil the requirements. The output impedance shall be  $< 2 \Omega$ . The maximum output voltage shall be  $150 \text{ mV} \pm 1 \text{ mV}$  when loaded with a  $32 \Omega$  resistor.

The bias voltage provided by the test system shall be  $2.6 \text{ V} \pm 1\%$ . bias resistance ( $R_{bias}$ ) is the bias resistance inside the input of the test system. The bias resistance shall be  $2.2 \text{ k}\Omega \pm 2\%$ . The nominal level shall be  $-60 \text{ dBV}$  (expected from a headset with a nominal sensitivity of  $-55 \text{ dBV/Pa}$ ).

### 8.1.1.2 Test signals and test signal levels

Unless otherwise specified, full-band real speech signals are used for the measurements which can be found in [ITU-T P.501]. Detailed information about the test signal used can be found in the corresponding clause of [ITU-T P.501].

All test signals – which are used in Receive – have to be band-limited. The band limitation is achieved by bandpass filtering in the frequency range between 50 Hz and 8 kHz using a bandpass filter providing at least 24 dB/octave. In Send, the test signals are used without band limitation.

All test signal levels are referred to the average level of the test signals, averaged over the complete test sequence length, unless described otherwise.

The nominal average signal levels for the measurements are as follows:

- $-37 \text{ dBV}$  in receive for binaural headsets
- $-31 \text{ dBV}$  in receive for monaural headsets
- $-4.7 \text{ dBPa}$  at the MRP

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

### 8.1.1.3 Positioning of the headsets

Recommendations for the set-up and positioning of headsets are given in [ITU-T P.380]. Unless stated otherwise, headsets shall be placed in their recommended wearing position.

Some insert earphones may not fit properly in Type 3.3 ear simulators. For such insert type headsets, an ITU-T P.57 Type 2 ear simulator may be used in conjunction with the head and torso simulator (HATS) mouth simulator. In this case, the Type 2 ear simulators can be mounted on HATS. Separate Type 2 simulators may also be used; in this case they should be placed outside HATS close to the HATS pinnae.

#### **8.1.1.4 Position and calibration of HATS**

The HATS shall be equipped with a Type 3.3 artificial ear. For the measurement of binaural headsets the HATS shall be equipped with two artificial ears. The pinnae are specified in [ITU-T P.57] for Type 3.3 artificial ears. The pinnae shall be positioned on HATS according to [ITU-T P.58].

The exact calibration and equalization procedures as well as how to combine the two ear signals for the purpose of measurements, can be found in [ITU-T P.581]. Unless stated otherwise, the HATS shall be diffuse-field equalized. The reverse nominal diffuse field curve as found in Table 3 of [ITU-T P.58] shall be used. For measurements requiring diffuse-field correction values for closer frequency spacing than that which is specified in [ITU-T P.58], the interpolation method found in Annex A shall be used.

#### **8.1.2 Sensitivity in Send**

##### **8.1.2.1 Requirements**

The sending sensitivity is measured from the MRP to the sending input of the headset reference interface input.

The sending sensitivity shall be  $-55 \text{ dBV/Pa} \pm 6 \text{ dB}$  when inserting the sending signal at the nominal level, as described in clause 8.1.1.2.

##### **8.1.2.2 Test**

- 1) The test arrangement is according to clause 8.1.1, Figure 8-1.
- 2) 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 the nominal signal level, the level is averaged over the complete test signal.

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

- 3) For the calculation, the average measured level at the headset reference interface is used.
- 4) The sensitivity is expressed in dBV/Pa.

#### **8.1.3 Sensitivity in Receive**

##### **8.1.3.1 Requirements**

The requirements are as defined in clause 8.2.2.

##### **8.1.3.2 Test**

The measurement method is defined in clause 8.2.2.

#### **8.1.4 Sending frequency response**

##### **8.1.4.1 Requirements**

The sending frequency response is measured from the MRP to the sending input of the headset reference interface input.

The measured frequency response shall be within the limits defined in Table 8-1.

**Table 8-1 – Tolerance mask for the wideband sending frequency response**

Frequency (Hz)	Upper limit	Lower limit
100	4	$-\infty$
200	4	-4
1 000	4	-4
3 000	4	-
8 000	-1	-15

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

**8.1.4.2 Test**

- 1) The test arrangement is according to clause 8.1.1, Figure 8-1.
- 2) 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 the nominal signal level, applied at the MRP.  
The measured power density spectrum at the MRP is used as the reference power density spectrum for determining the sending sensitivity.
- 3) The sending sensitivity is determined in third octave intervals, as given by [IEC 61260-1] for frequencies between 100 Hz and 8 kHz inclusive, measured at the POI. In each third octave band, the level of the measured signal is referred to the level of the reference signal averaged over the complete test sequence length.
- 4) The sensitivity is determined in dBV/Pa.

**8.1.5 Receiving frequency response****8.1.5.1 Requirements**

The receiving frequency response is measured from the receiving output of the headset reference interface to the drum reference position (DRP) with diffuse-field correction.

The measured frequency response shall be within the limits as defined in Table 8-2.

**Table 8-2 – Tolerance mask for the wideband receiving frequency response**

Frequency (Hz)	Upper limit	Lower limit
100	12	$-\infty$
200	10	-10
300	9	-6
1 000	6	-6
2 000	8	-6
5 000	8	-6
8 000	8	-12
10 000	8	$-\infty$

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

### 8.1.5.2 Test

- 1) The test arrangement is according to clause 8.1.1, Figure 8-1.
- 2) 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 the nominal signal level, applied to the headset reference interface. The level is averaged over the complete test signal.
- 3) For wideband, the receiving sensitivity is determined in third octave intervals as given by [IEC 61260-1] for frequencies between 100 Hz and 10 kHz inclusive, measured at the headset interface. In each third octave band, the level of the measured signal is referred to the level of the reference signal, averaged over the complete test sequence length.
- 4) The sensitivity is determined in dBPa/V.

The measurement is repeated for the second channel.

### 8.1.6 Idle channel noise in Send

#### 8.1.6.1 Requirements

The idle channel noise in Send is measured from the MRP to the sending input of the headset reference interface.

The idle noise in sending direction shall be  $< -90$  dBV(A).

#### 8.1.6.2 Test

- 1) The test arrangement is according to clause 8.1.1, Figure 8-1.
- 2) The British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501] at the nominal signal level, as described in clause 8.1.1.2, is applied at the MRP. The test signal level is the average level of the complete test signal. The output level is measured using a speech level voltmeter according to [ITU-T P.56]. This level is the reference speech signal level.
- 3) For the noise measurement, no test signal is used. However, all sources which potentially contribute to noise should be considered. Interference from radio frequencies is not accurately covered by an interface specification as the complete terminal/headset system needs to be assessed. Moreover the necessary test system cabling is likely to introduce further deviations from real-life conditions. Therefore, radio induced noise is not expected to be accurately covered by the test cases in the present Recommendation.
- 4) The idle channel noise is measured at the output in the frequency range between 100 Hz and 8 kHz. The length of the time window is 1 s which is the averaging time for the idle channel noise. The test lab has to ensure the correct activation of the headset during the measurement. If the headset is deactivated during measurement, the measurement window has to be cut to the duration when the headset remains activated.

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

- 5) The idle channel noise is determined by A-weighting [IEC 61672-1] and referring to the reference speech signal level as determined with the speech sequence. Spectral peaks are measured in the frequency domain. The average noise spectrum used for determining the spectral peak should be calculated as the arithmetic mean of the noise spectrum values when stated in dBV.

### 8.1.7 Distortion in Send

#### 8.1.7.1 Requirements

The distortion in Send is measured from the MRP to the sending input of the headset reference interface.



The ratio of signal to harmonic distortion shall be above the following mask as shown in Table 8-3.

**Table 8-3 – Limits for the signal to harmonic distortion**

Frequency (Hz)	Signal to harmonic distortion ratio limit, Send (dB)
315	40
400	50
1 000	50

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

### 8.1.7.2 Test

- 1) The test arrangement is according to clause 8.1.1, Figure 8-1.
- 2) For the test, a sinusoidal signal at frequencies of 315 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz and 1 000 Hz is used. The duration of the sine wave shall be < 1 s. The sinusoidal signal level shall be the nominal signal level. In order to ensure a reliable activation, a conditioning sequence is inserted before the actual measurement, in case of active headsets. The conditioning sequence is according to clause 7.3.7 of [ITU-T P.501]. The short conditioning sequences (either male or female) should be used. The level of the activation signal is the nominal signal level.
- 3) The signal to harmonic distortion ratio is measured selectively up to 7 kHz.
- 4) The test is repeated using a signal level 10 dB higher than the nominal signal level. The level of the activation signal is kept at the nominal signal level.

### 8.1.8 Coupling loss

#### 8.1.8.1 Requirements

The headset terminal coupling loss weighted (HTCL<sub>w</sub>) is measured at the headset reference interface, from receiving to sending.

The HTCL<sub>w</sub> provided by the headset shall be  $\geq 40$  dB.

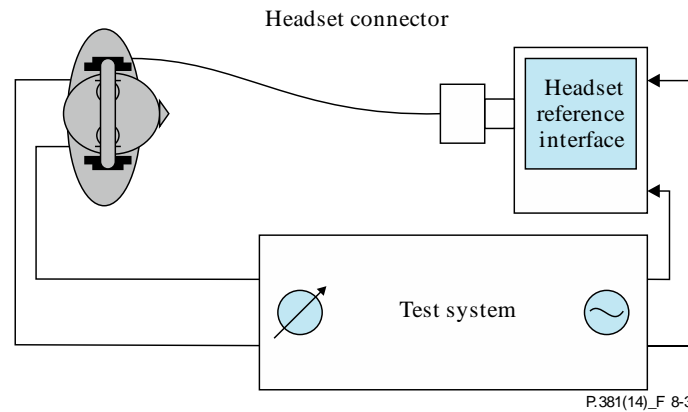
#### 8.1.8.2 Test

- 1) The test arrangement is according to clause 8.1.1. Identical signals are applied to both left and right receiving direction headset terminals.
- 2) The attenuation between the receiving direction and sending direction is determined.
- 4) The test signal is a pn sequence, according to [ITU-T P.501], with a length of 4 096 points (48 kHz sampling rate) and a crest factor of 6 dB. The duration of the test signal is 250 ms, the test signal level is -24 dBV. The low crest factor is achieved by random alternation of the phase between -180° and +180°.
- 5) HTCL<sub>w</sub> is calculated according to TCL<sub>w</sub> in clause B.4 of [ITU-T G.122] (trapezoidal rule). For the calculation, the average measured echo level at each frequency band is referred to the average level of the test signal measured in each frequency band. For the measurement, a time window has to be applied which is adapted to the duration of the actual test signal (250 ms).

## 8.2 Multimedia playback mode

### 8.2.1 Test set-up

The test set-up is shown in Figure 8-3.



**Figure 8-3 – Test arrangement for testing the headset**

### 8.2.1.1 Input and output characteristics of the test system for connecting the headset

The output impedance shall be  $< 2 \Omega$ . The maximum RMS output voltage shall be  $150 \text{ mV} \pm 1 \text{ mV}$  when loaded with a  $32 \Omega$  resistor. The common ground impedance (between sending and receiving sides) for the test system shall be  $\leq 0.05 \text{ Ohm}$ .

### 8.2.1.2 Test signals and test signal levels

Programme simulation noise is used for the measurements. Detailed information about the test signal used can be found in the corresponding clause of this Recommendation.

Artificial test signals – which are used in Receive – have to be band-limited. The band limitation is achieved by bandpass filtering in the frequency up to 22 kHz using low-pass filter providing higher than 24 dB/octave filter roll-off. The programme simulation noise according to [EN 50332] is band limited by design and requires no filtering.

All test signal levels are referred to the average level of the test signals, averaged over the complete test sequence length, unless described otherwise.

The nominal average signal level for the measurements is  $-32 \text{ dBV}$ .

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

### 8.2.1.3 Positioning of the headsets

Recommendations for the set up and positioning of headsets are given in [ITU-T P.380]. Unless stated otherwise, headsets shall be placed in their recommended wearing position. Some insert earphones may not fit properly in Type 3.3 ear simulators. For such insert type headsets, an ITU-T P.57 Type 2 ear simulator may be used in conjunction with the HATS mouth simulator. The HATS should be equipped with two artificial ears as specified in [ITU-T P.57]. For binaural headsets two artificial ears are required.

Unless stated otherwise the measurements in Receive are repeated five times and averaged. The averaged result is used.

### 8.2.1.4 Position and calibration of HATS

The HATS shall be equipped with a Type 3.3 Artificial Ear. For the measurement of binaural headsets the HATS shall be equipped with two artificial ears. The pinnae are specified in [ITU-T P.57] for Type 3.3 artificial ears. The pinnae shall be positioned on HATS according to [ITU-T P.58].

The exact calibration and equalization procedures as well as how to combine the two ear signals for the purpose of measurements can be found in [ITU-T P.581]. Unless stated otherwise, the HATS shall be diffuse-field equalized. The DRP to diffuse field correction curve as found in Table 14A and

Table 14B of [ITU-T P.58] shall be used for 1/3rd octaves and 1/12th octaves respectively. For measurements requiring diffuse-field correction values for closer frequency spacing than that which is specified in [ITU-T P.58], the interpolation method found in Annex A shall be used.

## 8.2.2 Sensitivity in multimedia playback mode

### 8.2.2.1 Requirements

The receiving sensitivity is measured from the receiving output of the headset reference interface to the diffuse field equalized sound pressure.

The maximum headset receiving sensitivity is governed from an acoustic safety point of view considering expected exposure times and levels using portable music players. It is specified in terms of simulated programme signal characteristic voltage (SPCV). This is the voltage, of a specific programme simulation noise signal, required for producing a sound pressure of 1 Pa, after applying diffuse-field correction and A-weighting [IEC 61672-1].

The receiving sensitivity is measured from the receiving output of the headset reference interface to the DRP with diffuse-field correction and A-weighting.

The sensitivity shall be  $75 \text{ mV} \leq \text{SPCV} \leq 300 \text{ mV}$  when measured according to [EN 50332].

### 8.2.2.2 Test

- 1) The test arrangement is according to clause 8.2.1, Figure 8-3.
- 2) The test signal used for the measurements shall be programme simulation noise providing sufficient signal energy up to 22 kHz. The test signal level is the nominal signal level.
- 3) For the calculation, the averaged level at the output of the headset interface is used. The sensitivity is determined from 20 Hz to 20 kHz.
- 4) The sensitivity is expressed in dBPa/V.

The measurement is repeated for the second channel.

## 8.2.3 Distortion in multimedia playback mode

### 8.2.3.1 Requirements

The distortion is measured from the receiving output of the headset reference interface to the DRP with diffuse-field correction.

The ratio of signal to harmonic distortion shall be above the following mask as shown in Table 8-4.

**Table 8-4 – Limits for the signal to harmonic distortion**

Frequency (Hz)	Signal to harmonic distortion ratio limit (dB)
100	40
315	50
5 000	50

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

### 8.2.3.2 Test

- 1) The test arrangement is according to clause 8.2.1, Figure 8-3.
- 2) For the test, a sinusoidal signal at frequencies of 100 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz, 1 000 Hz, 2 000 Hz and 5 000 Hz is used. The duration of the sine wave shall be < 1 s. The sinusoidal signal level shall be the nominal signal level of –32 dBV.

3) The signal to harmonic distortion ratio is measured selectively up to 10 kHz.

The measurement is repeated for the second channel.

The measurement is **not** repeated five times.

## 8.2.4 Receiving crosstalk

### 8.2.4.1 Requirements

The receiving crosstalk is measured as the L-R crosstalk and the R-L crosstalk generated by headset by playing programme simulation noise at the output of the headset interface and measuring the resulting level at the two output channels of the headset interface.

In duration of 0-5 s, the attenuation, measured at the right ear and referenced to the level at the left ear, shall be above 20 dB. In duration of 5-10 s, the attenuation, measured at the left ear and referenced to the level at the right ear, shall be also above 20 dB.

NOTE – The crosstalk attenuation of  $\geq 20$  dB should be kept over the frequency range from 50 Hz to 16 000 Hz.

**Table 8-5 – Signal sequence for the L-R and R-L crosstalk**

<b>level period</b>	<b>Left channel (dBV)</b>	<b>Right channel (dBV)</b>
0-5 s	-22	0
5-10 s	0	-22

### 8.2.4.2 Test

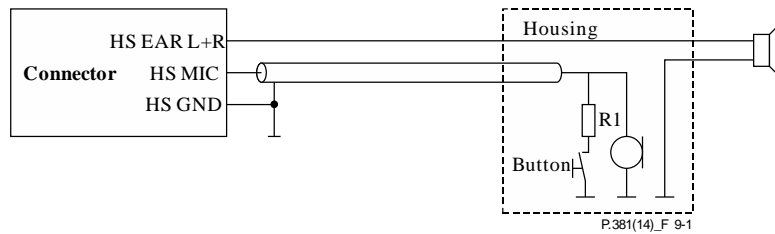
- 1) The test arrangement is according to clause 8.2.1, Figure 8-3.
- 2) The test signal used for the measurements shall be programme simulation up to 22 kHz. The test signal is the nominal signal level. The level is averaged over the complete test signal. The signal sequence is shown in Table 8-5.
- 3) Output the test signal to the headset, the crosstalk is determined by analysing the measured signal at the output of artificial ear. In duration of 0-5 s, the measured level at the right ear is referenced to the level at the left ear, and the attenuation is L-R crosstalk. In duration of 5-10 s, the measured level at the left ear is referenced to the level at the right ear, and the attenuation is R-L crosstalk.
- 4) The crosstalk is determined in dBPa/Pa.

The measurement is **not** repeated five times.

## 9 Function requirements for terminals with the universal headset interface

The terminal shall provide the intelligent detection mechanism as follows:

- 1) The terminal should be capable of detecting any plug-in action automatically and then activate the corresponding function according to the current state.
- 2) The terminal should be capable of detecting any plug-out action automatically and then activate the corresponding function according to the current state.
- 3) The terminal should be capable of determining whether the inserted plug has three poles or four poles.
- 4) Some headsets have Function/End buttons for submitting the signals to the terminal. Figure 9-1 is an example illustrating how the signal is submitted through the MIC input pin. The resistance to GND should be according to Table 9-1.



**Figure 9-1 – Button-control headset function**

NOTE – For harmonization with some already-in-market headsets and the better user experience, it has been suggested that the terminal with a universal interface be able to detect headsets of different pole orders automatically and adjust the speech or audio stream accordingly.

**Table 9-1 – Recommended resistance ranges for Function/End button detection by the terminal and recommended values for the headset. All values refer to the resistance observed at the audio jack (including also the microphone impedance for a case with 2.2 V and 2.2 k $\Omega$  bias network)**

Resistance range	Min threshold for detection by the terminal [ $\Omega$ ]	Headset typical [ $\Omega$ ]	Max threshold for detection by the terminal [ $\Omega$ ]
End/Play/Pause	0	0	70
Voice control	110	135	180
Vol+	210	240	290
Vol-	360	470	680

## **Annex A**

### **Interpolation method for diffuse-field correction**

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

For measurements requiring diffuse-field correction values for closer frequency spacing than 1/12-octave bands, linear interpolation on a log scale from the 1/12-octave band interpolated values in Table 14B of [ITU-T P.58] shall be used.

## Appendix I

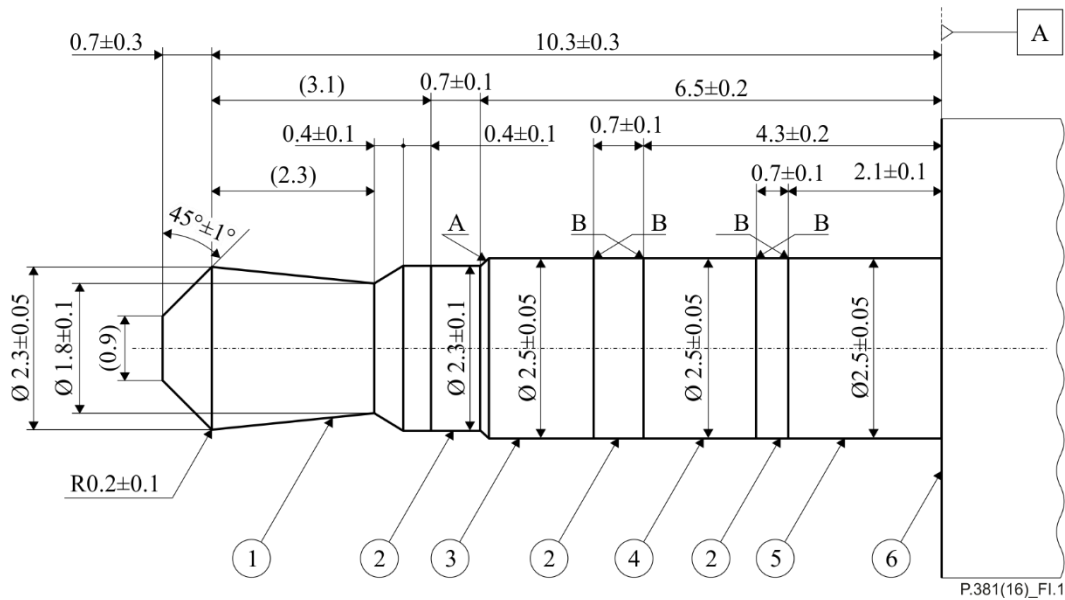
### Audio connectivity for sockets with four contact points

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

This appendix illustrates the dimensions of the concentric plug and socket connector with four contact points.

#### I.1 2.5 mm diameter plug connector with four poles

Figure I.1 shows the shape and dimensions of the 2.5 mm diameter plug connector with four poles. The width of strip A along the axial direction is 0.1 mm. Junction B should be free of burr or fash.

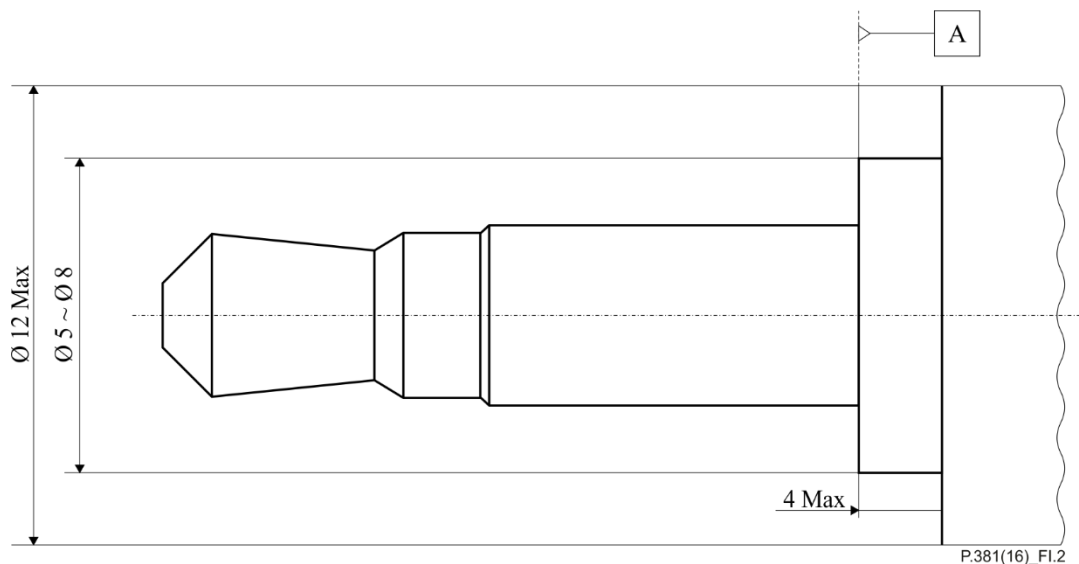


NOTE 1 – ① is the Tip and made of conductive material; ② is the insulating ring; ③ is Ring 1 and made of conductive material; ④ is Ring 2 and made of conductive material; ⑤ is the sleeve and made of conductive material; ⑥ is an illustration of the hand grip at the end of a plug.

NOTE 2 – "Fash" here refers to a rough edge or ridge on the surface.

**Figure I.1 – Shape and dimensions of the 2.5 mm diameter plug connector with four poles**

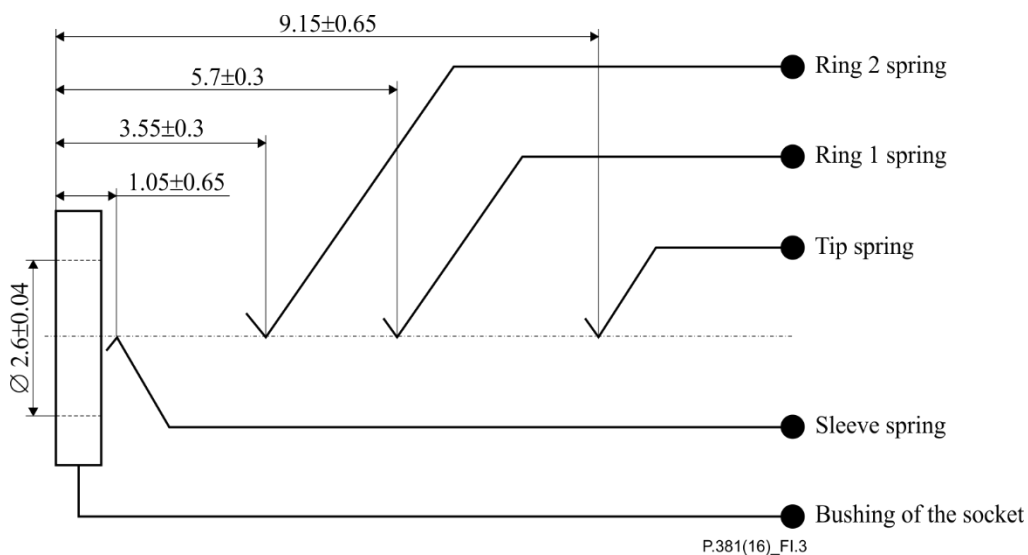
Figure I.2 shows the dimensions of the 2.5 mm diameter plug grip, specified to ensure the plug can be properly inserted.



**Figure I.2 – Dimensions of the 2.5 mm diameter plug connector grip**

### I.2 2.5 mm diameter socket connector with four contact points

The socket should be able to mate and cooperate with the plug reliably. The dimensions and positioning for each contact spring are illustrated in Figure I.3. Considering the tolerance of the plug dimension and positioning of the socket contact spring, in addition to the shift of the practical contact point location caused by the width of the spring, the minimum distance between the contact point of the Ring 2 spring and that of the sleeve spring is recommended to be longer than 1.6 mm. If bushing of the socket is made of conductive material, the contact area of the sleeve spring may exceed the given range indicated in Figure I.3, so bushing of the socket should not be longer than 2.0 mm.

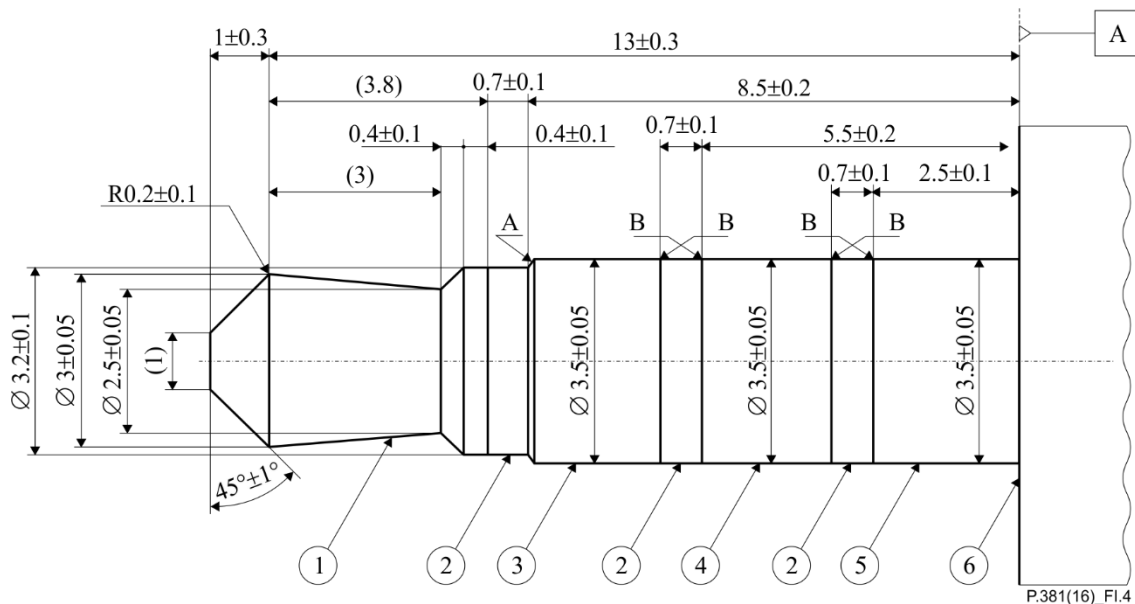


**Figure I.3 – Dimensions of the 2.5 mm diameter socket with four contact points and positioning of each contact spring**

### I.3 3.5 mm diameter plug connector with four poles

Figure I.4 shows the shape and dimensions of the 3.5 mm diameter plug connector with four poles. The width of strip A along the axial direction is 0.15 mm. Junction B should be free of burr or fash.



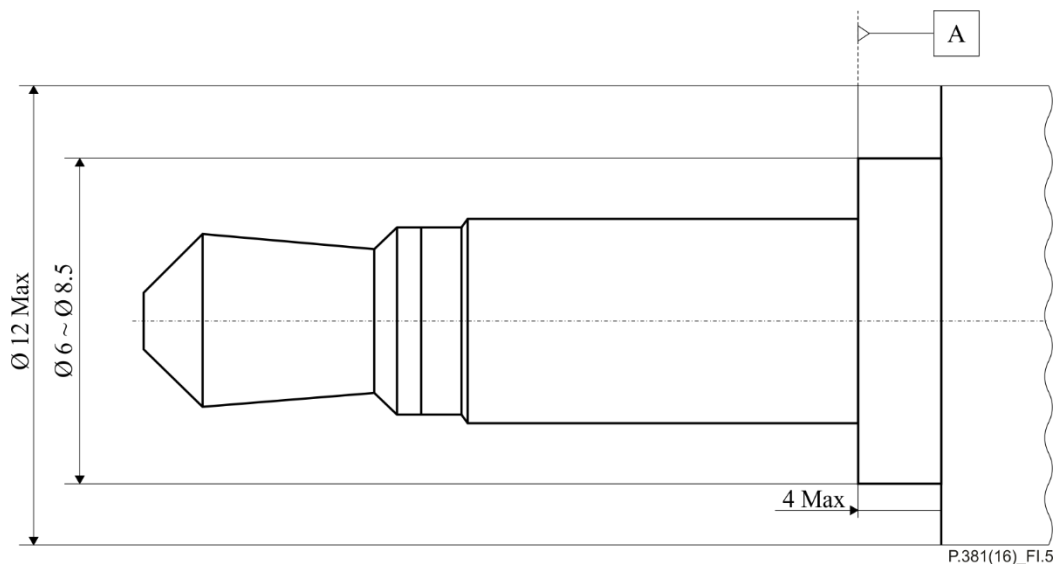


NOTE 1 – ① is the Tip and made of conductive material; ② is the insulating ring; ③ is Ring 1 and made of conductive material; ④ is Ring 2 and made of conductive material; ⑤ is the sleeve and made of conductive material; ⑥ is an illustration of the hand grip at the end of a plug.

NOTE 2 – "Fash" here refers to a rough edge or ridge on the surface.

**Figure I.4 – Shape and dimensions of the 3.5 mm diameter plug connector with four poles**

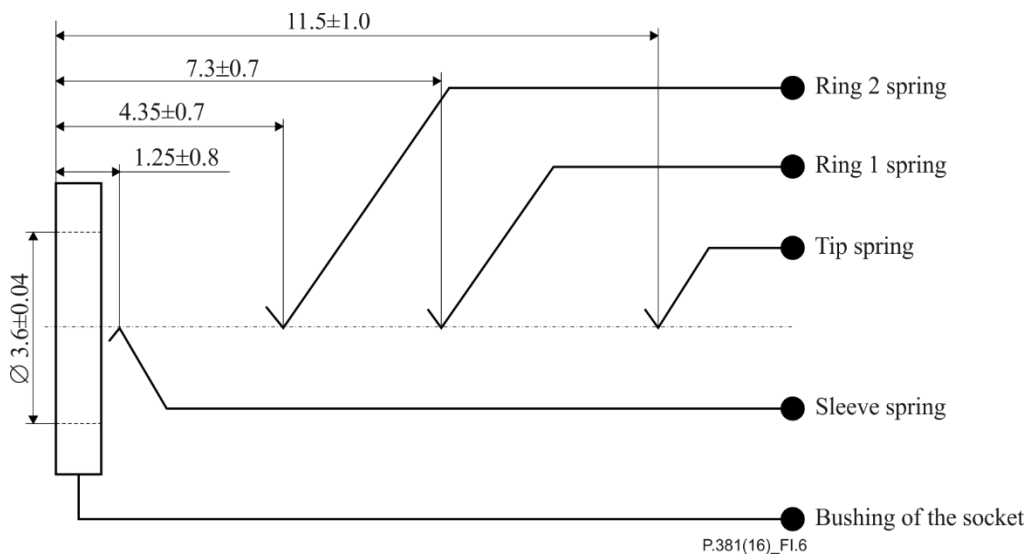
Figure I.5 shows the dimensions of the 3.5 mm diameter plug grip, specified to ensure the plug can be properly inserted.



**Figure I.5 – Dimensions of the 2.5 mm diameter plug connector grip**

#### I.4 3.5 mm diameter socket connector with four contact points

The socket should be able to mate and cooperate with the plug reliably. The dimensions and the positioning for each contact spring are illustrated in Figure I.6. Considering the tolerance of the plug dimension and positioning of the socket contact spring, in addition to the shift of the practical contact point location caused by the width of the spring, the minimum distance between the contact point of the Ring 2 spring and that of the sleeve spring is recommended to be more than 2.4 mm. If bushing of the socket is made of conductive material, the contact area of the sleeve spring may exceed the given range indicated in Figure I.6 so bushing of the socket should not be longer than 2.4 mm.



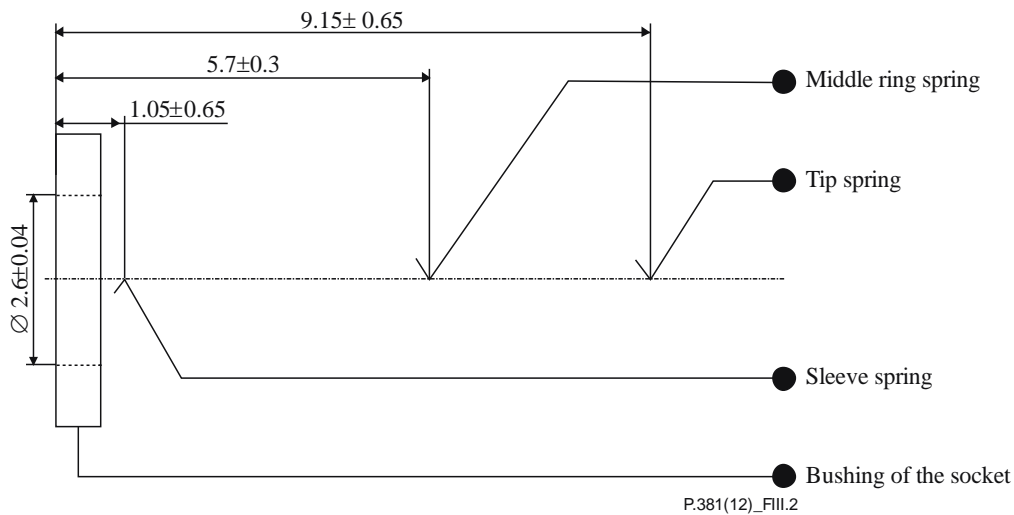
**Figure I.6 – Dimensions of the 3.5 mm diameter socket with four contact points and positioning of each contact spring**

## **Appendix II**

### **Audio connectivity for sockets with four contact points (optional dimensions to accommodate terminal designs with curved edges)**

Void

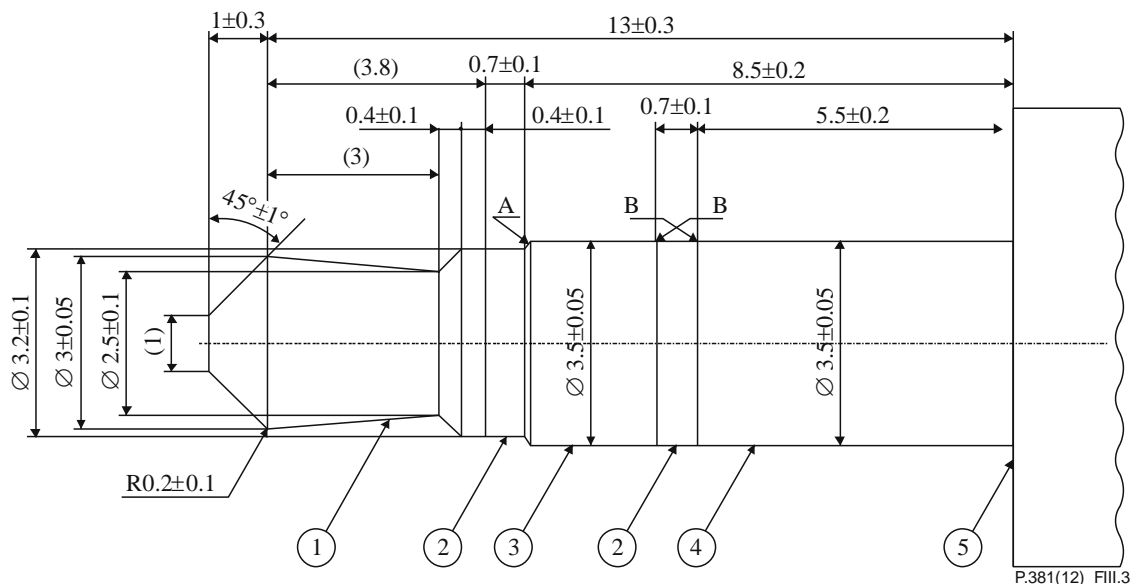




**Figure III.2 – Dimensions of the 2.5 mm diameter socket with three contact points and positioning of each contact spring**

### III.3 3.5 mm diameter plug connector with three poles

Figure III.3 shows the shape and dimensions of the 3.5 mm diameter plug connector with three poles. The width of strip A along the axial direction is 0.15 mm. Junction B should be free of burr or fash.

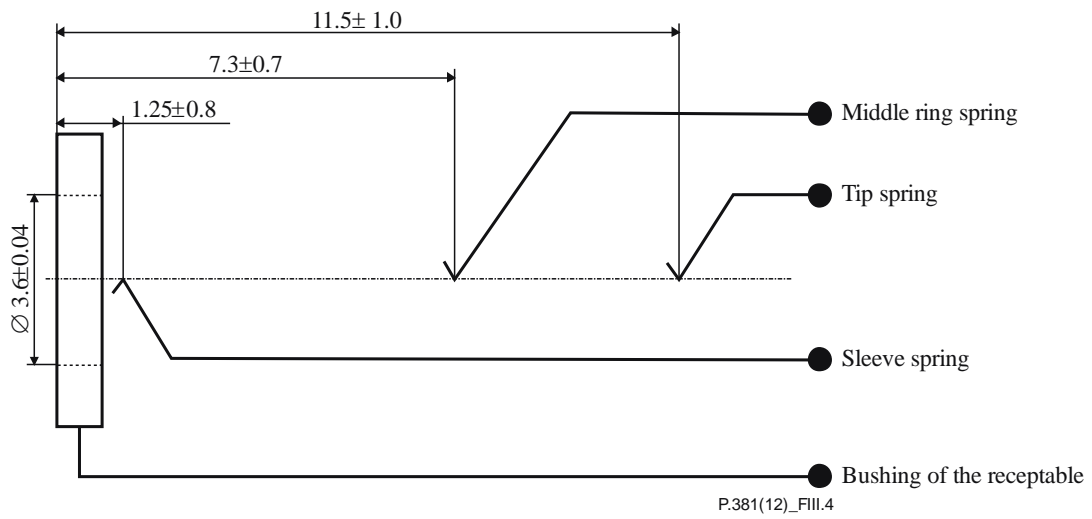


NOTE – ① is the Tip and made of conductive material; ② is the insulating ring; ③ is the middle ring and made of conductive material; ④ is the sleeve and made of conductive material; ⑤ is an illustration of the hand grip at the end of a plug.

**Figure III.3 – Shape and dimensions of the 3.5 mm diameter plug connector with three poles**

### III.4 3.5 mm diameter socket connector with three contact points

The socket should be able to mate and cooperate with the plug reliably. The dimensions and positioning for each contact spring are illustrated in Figure III.4. To achieve better compatibility with various headphones or headsets, the 3.5 mm diameter sockets with three contact points should be in line with this requirement. If bushing of the socket is made of conductive material, the contact area of the sleeve spring may exceed the given range indicated in Figure III.4, so bushing of the socket should not be longer than 5.3 mm.



**Figure III.4 – Dimensions of the 3.5 mm diameter socket with three contact points and positioning of each contact spring**

## Appendix IV

### Other considerations

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

#### IV.1 Filter recommendation

It is recommended to filter the connection to the microphone in order that the headset shall not generate any appreciable non-linear distortion as a result of interfering RF signals. This is covered by audio breakthrough testing (ABT) specifications.

#### IV.2 Electrostatic discharge

It is recommended to protect the microphone using an electrostatic discharge (ESD) diode. A low value can result in poor electromagnetic interference (EMI) performance (this can result in ABT failures) and a high value will result in voltage spikes as seen by the microphone itself.

#### IV.3 Microphone basics – background

The standard electret microphones are not voltage sources, but current sources when biased correctly. They contain an internal junction field effect transistor (JFET) that does not amplify the signal much, but changes impedance levels. When the bias voltage is low, below 0.8 V for example, the JFET will instead work as a resistor. When they operate as a current source, the output impedance can, for example, be around 30 k $\Omega$  and the current can, for example, be 220-250  $\mu$ A.

Since they operate as a current source, a larger bias resistor will result in a bigger signal swing which is the desired behaviour. However, at a certain point, the voltage drop across the resistor means that the JFET inside the microphone goes out of saturation and the benefits of a large resistor are lost. The optimal resistor value depends on the bias point. Typically, it is in the range 3-6 k $\Omega$ . However, in some cases a lower value is desired, for example, if you want to detect multiple passive button values.

Traditionally, 2.2 k $\Omega$  has been used in the past 30 years by electret manufacturers. The value comes from a max current of 500  $\mu$ A from the old JFET design. As JFET technology has improved, manufacturers have increased the DC resistance value. Although conservative, some typical values seen have been 2.6 k $\Omega$ .

#### IV.4 Vcc voltage for microphone bias

In principle, the same goes for the biasing voltage, but due to other constraints, there may be a reason to use a range (e.g., EMC/ESD reasons). A range of 1.5-3.6 V can be used, but the range 2.7-3.0 V is preferred, for compatibility considerations with all headsets (see note below).

A high bias voltage allows a larger bias resistor and pushes JFET further into saturation and thereby gives better S/N.

NOTE – The preferred range of 2.7-3.0 V for microphone bias is not in conflict with the 2.6 V specified in clause 8.1.1.1 for the headset testing system. The headset testing system voltage of 2.6 V was selected for backward compatibility with existing mobile phone sockets. The performance of headsets measured using a 2.6 V microphone bias will also be representative of performance when using a slightly higher microphone bias in the range of 2.7-3.0 V. However, the reverse may not be true, because 2.6 V is near the edge of what is needed for some microphone headsets to operate properly. Therefore, testing with a slightly lower microphone bias voltage may catch interoperability issues with existing sockets not designed to the preferred 2.7-3.0 V range (while still being representative of performance for new sockets that are designed as per the preferred voltage range). Mobile phone sockets using this preferred operating range will be backwards compatible with a wide range of existing headsets.

#### **IV.5 DC resistance of microphone**

The microphone current should be in the range of 120-300  $\mu\text{A}$  (i.e., the DC resistance should be such that the resulting microphone current is in this range).



## Bibliography

- [b-ITU-T G.722.2] Recommendation ITU-T G.722.2 (2003) *Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)*.
- [b-ITU-T P.310] Recommendation ITU-T P.310 (2009), *Transmission characteristics for narrow-band digital handset and headset telephones*.
- [b-ITU-T P.350] Recommendation ITU-T P.350 (2001), *Handset dimensions*.
- [b-ITU-T P.370] Recommendation ITU-T P.370 (1996), *Coupling hearing aids to telephone sets*.
- [b-ITU-T P.1100] Recommendation ITU-T P.1100 (2008), *Narrow-band hands-free communication in motor vehicles*.
- [b-3GPP TS 26.131] 3GPP TS 26.131 V12.3.0 (2014), *Terminal acoustic characteristics for telephony; Requirements*.  
<http://www.3gpp.org/DynaReport/26131.htm>>
- [b-3GPP TS 26.132] 3GPP TS 26.132 V10.1.0 (2011), *Speech and video telephony terminal acoustic test specification (Release 10)*.
- [b-EIAJ RC-5325A] EIAJ RC-5325A, (2004), *4-Pole miniature concentric plugs and jacks*.  
<http://www.jeita.or.jp/japanese/standard/book/RC-5325A/#page=1>
- [b-ETSI EG 202 396-1] ETSI EG 202 396-1 (2011), *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; Version 1.2.4*.
- [b-ETSI EN 301 489-7] ETSI EN 301489-7 (2005), *Electromagnetic compatibility and Radio spectrum Matters (ERM); ElectroMagnetic Compatibility (EMC) standard for radio equipment and services; Part 7: Specific conditions for mobile and portable radio and ancillary equipment of digital cellular radio telecommunications systems (GSM and DCS)*.
- [b-IEC 60130-8] IEC 60130-8 (1976), *Connectors for frequencies below 3 MHz. Part 8: Concentric connectors for audio circuits in radio receivers*.
- [b-IEC 60268-11] IEC 60268-11 (1987), *Sound system equipment. Part 11: Application of connectors for the interconnection of sound system components*.
- [b-IEC 60603-11] IEC 60603-11 (1992), *Connectors for frequencies below 3 MHz for use with printed boards – Part 11: Detail specification for concentric connectors (dimensions for free connectors and fixed connectors)*.
- [b-IEC 61842] IEC 61842 (2002), *Microphones and earphones for speech communications*.
- [b-IEC 61938] IEC 61938 (1996), *Audio, video and audiovisual systems – Interconnections and matching values – Preferred matching values of analogue signals*.
- [b-IEC 62368-1] IEC 62368-1 (2010), *Audio/video, information and communication technology equipment – Part 1: Safety requirements*.
- [b-JIS C 6560] JIS C 6560 (1994), *Concentric plugs and jacks*.

[b-OMTP]

OMTP (2007), *Local connectivity: Wired analogue audio*.  
<http://www.gsma.com/newsroom/wp-content/uploads/2012/03/omtplocalconnectivitywiredanalogueaudiov10.pdf>

[b-3GPP TS 26.071]

3GPP TS 26.071, *Mandatory speech CODEC speech processing functions; AMR speech Codec; General description*.



## SERIES OF ITU-T RECOMMENDATIONS

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