ITUPublications
Recommendations

International Telecommunication Union
Standardization Sector

Recommendation

ITU-T P.381 (03/2023)

SERIES P: Telephone transmission quality, telephone installations, local line networks

Voice terminal characteristics

Technical requirements and test methods for analogue wired headsets or headphones and corresponding universal interface of terminals

ITU-T P-SERIES RECOMMENDATIONS

TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS

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

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

Recommendation ITU-T P.381

Technical requirements and test methods for analogue wired headsets or headphones and corresponding universal interface of 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 or headphone (HP) interfaces have been widely used in digital mobile terminals in recent years. Nowadays, the consumer is free to choose either the headset or HP originally provided by the terminal manufacturer or others that are offered separately. However, the quality of service or experience perceived by users is influenced by both the electrical performance of the interface and the compatibility between the terminal and the connected headset or HP.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T P.381	2012-08-22	12	11.1002/1000/11685
2.0	ITU-T P.381	2014-02-13	12	11.1002/1000/12139
3.0	ITU-T P.381	2016-07-29	12	11.1002/1000/12970
4.0	ITU-T P.381	2017-03-01	12	11.1002/1000/13172
5.0	ITU-T P.381	2020-10-14	12	11.1002/1000/14465
6.0	ITU-T P.381	2023-03-01	12	11.1002/1000/15464

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

FOREWORD

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

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

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

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

NOTE

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

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

INTELLECTUAL PROPERTY RIGHTS

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

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

© ITU 2023

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

Electronic attachment: Set of signals referenced in Annex B.

Recommendation ITU-T P.381

Technical requirements and test methods for analogue wired headsets or headphones and corresponding universal interface of terminals

1 Scope

This Recommendation[1] specifies electrical requirements and test methods for the universal analogue headset or headphone (HP) interface used in digital mobile terminals. It is also applicable to terminals with a digital headset interface combined with an external adapter for an analogue (universal) headset interface. The combination is measured as one device.

The principle of this Recommendation is to ensure adequate compatibility between the digital mobile terminal and the wired analogue headset or HP, and to provide better user experience. The universal headset or HP interface enables the separation of sales between digital mobile terminals and headsets or HPs. One of the benefits is that the user can be free to choose his or her favourite type of headset or HP 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 to III.

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

This Recommendation is not applicable to terminals designed solely for digital headset or HP 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.114] Recommendation ITU-T G.114 (2003), *One-way transmission time*.

[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.56] Recommendation ITU-T P.56 (2011), *Objective measurement of active speech level*.

[ITU-T P.57] Recommendation ITU-T P.57 (2021), *Artificial ears*.

[ITU-T P.58] Recommendation ITU-T P.58 (2023), *Head and torso simulator for telephonometry*.

[ITU-T P.64] Recommendation ITU-T P.64 (2022), *Determination of sensitivity/frequency characteristics of local telephone systems*.

[1] This Recommendation includes an electronic attachment available from https://www.itu.int/net/itu-t/sigdb/genaudio/AudioForm-g.aspx?val=10000381

[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 (2022), *Electro-acoustic measurements on headsets*.

[ITU-T P.501] Recommendation ITU-T P.501 (2020), *Test signals for use in telephony and other speech-based applications.*

[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 (2022), *Use of head and torso simulator for hands-free and handset terminal testing.*

[ITU-T P.863] Recommendation ITU-T P.863 (2018), *Perceptual objective listening quality prediction.*

[IEC 60268-1] International Standard IEC 60268-1:1985, *Sound system equipment – Part 1: General*.

[IEC 61260-1] International Standard IEC 61260-1:2014, *Electroacoustics – Octave-band and fractional-octave-band filters – Part 1: Specifications.*

[IEC 61672-1] International Standard IEC 61672-1:2013, *Electroacoustics – Sound level meters – Part 1: Specifications.*

[EN 50332-1] European Standard 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] European Standard 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.*

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

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

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

3.1.3 earphone [b-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 [b-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) [b-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 A.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 that 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.

NOTE – Definition based on that in [b-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.

NOTE – Definition based on that in [b-ITU-T P.10].

3.2.4 headphone (HP): 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).

NOTE – Definition based on that in [b-IEC 60268-7].

3.2.5 mean opinion score-listening-only quality objective (MOS-LQO): A score calculated by means of an objective model that aims to predict the quality for a listening-only test situation.

NOTE – 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 a signal transmission, usually from the measurement system to the device under test.

3.2.7 Send: The sending direction of a signal transmission, usually from the device under test to the measurement system.

4 Abbreviations, acronyms and symbols

4.1 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

A	Attenuation
ABT	Audio Breakthrough Testing
AGC	Automatic Gain Control
AMR	Adaptive Multi-Rate
CSS	Composite Source Signal
DC	Direct Current
DRP	eardrum Reference Point
DUT	Device Under Test
EC	Echo Canceller
EL	Echo Loss

ERP	Ear Reference Point
ESD	Electrostatic Discharge
EVS	Enhanced Voice Service
FB	Fullband
FFT	Fast Fourier Transform
G	Global
GND	Ground
HATS	Head and Torso Simulator
HTCL	Headset Terminal Coupling Loss
HTCLw	weighted Headset Terminal Coupling Loss
HP	Headphone
ID	Identifier
JFET	Junction Field Effect Transistor
L	Left
LQO	Listening-only Quality Objective
LQOf	Listening-only Quality Objective fullband
LQOn	Listening-only Quality Objective narrowband
LQOs	Listening-only Quality Objective super-wideband
LQOw	Listening-only Quality Objective wideband
LQr	Listening Quality in the Receive direction
LQs	Listening Quality in the Send direction
MIC	Microphone
MOS	Mean Opinion Score
MRP	Mouth Reference Point
N	Noise
NB	Narrowband
PN	Pseudo-Noise
POI	Point of Interconnection
R	Right
RF	Radio Frequency
RLR	Receive Loudness Rating
RMS	Root Mean Square
S	Speech
SNR	Signal to Noise Ratio
SPCV	Simulated Programme signal Characteristic Voltage
STMR	Sidetone Masking Rating
SWB	Super-Wideband

TCL	Terminal Coupling Loss
TCLw	weighted Terminal Coupling Loss
TELRdt	Talker Echo Loudness Rating under double talk conditions
Tx	Transmitter
WB	Wideband

4.2 Symbols

$A_{\mathrm{H,R,dt}}$	attenuation range in the Receive direction during double talk
$A_{\mathrm{H,S,dt}}$	attenuation range in the Send direction during double talk
dBFS	decibels relative to full scale
dBV	decibels relative to 1 V
L_{meST}	sidetone path loss
L_{nominal}	nominal input level
$L_{\mathrm{s,min}}$	minimum activation level in the Send direction
L_{sendnom}	level in Send for nominal speech input level
R_{bias}	bias resistance
T_{ar}	implementation independent system delay in Receive
T_{as}	implementation independent system delay in Send
T_{r}	overall terminal delay in Receive
$T_{\mathrm{r,sys}}$	test system delay in the Receive direction
$T_{\mathrm{r,s,min}}$	minimum built-up time in Send
T_{s}	terminal delay in Send
$T_{\mathrm{s,sys}}$	test system delay in the Send direction
V_{cc}	voltage at the common collector

5 General description

Generally, if a headset or a HP is used, the overall user experience during a call depends greatly on both the terminal and the connected headset or HP. Although the acoustic quality of the headset or HP 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 between digital mobile terminals and headsets or HPs, including connectors to the plug and socket. Normally, the socket connector is fixed inside the terminal, with the outside rim of the socket level with the surrounding shell of the terminal.

6 Physical characteristics

6.1 General rules

Two types of concentric socket connector are recommended for use, of diameter 2.5 mm and 3.5 mm. Figure 6-1 shows an isometric view of plug and socket connectors:

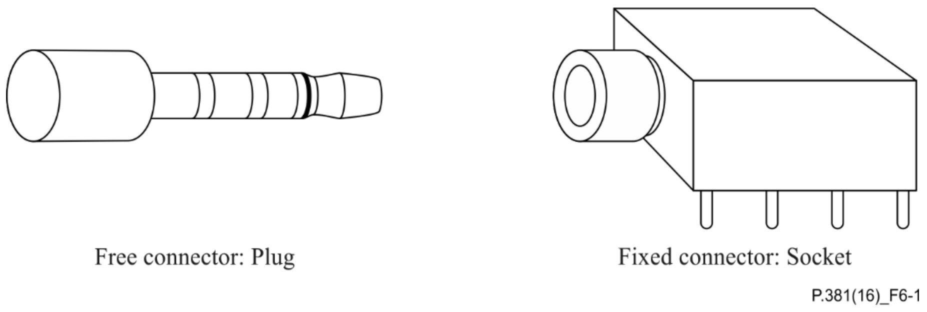


Figure 6-1 – Isometric view of the plug and socket connectors

If a terminal is equipped with a headset interface and designed for both communication and audio, and a universal interface is utilized, the fixed connector shall be a concentric socket connector with four contact points of diameter either 3.5 mm or 2.5 mm. Detailed dimensions of sockets and plug connectors with four contact points are given in Appendix I. Some terminals with curved edges may not work well with the dimensions given in Appendix I.

Although socket connectors with three contact points are no longer recommended for digital mobile terminals, their dimensional information is given in Appendix III.

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

6.2 Pin assignments

Figure 6-2 illustrates pin assignments of the socket connector with four contact points and those of the mated plug.

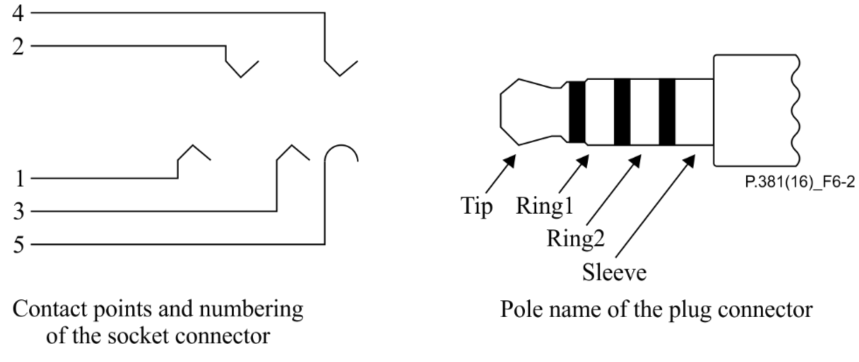


Figure 6-2 – Pin assignments of a socket connector with four contact points

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

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

There are two different pin assignments commonly used today across geographical regions. It is recommended that terminals be able to identify both pin assignments intelligently and automatically.

6.2.1 Recommended pin assignment

On the left-hand side of Figure 6-2:

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

In summary, the pole order from the tip to the sleeve of the headset plug is: L/R/MIC/GND.

The pinout order of L/R/MIC/GND has an advantage in electrostatic discharge (ESD) protection and allows for both plastic and metallic convertors.

6.2.2 Alternative pin assignment

On the left-hand side of Figure 6-2:

- 1 is the contact point of the tip, linking it to the L audio channel of the receiver;
- 2 is the contact point of ring 1, linking it to the R audio channel of the receiver;
- 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.

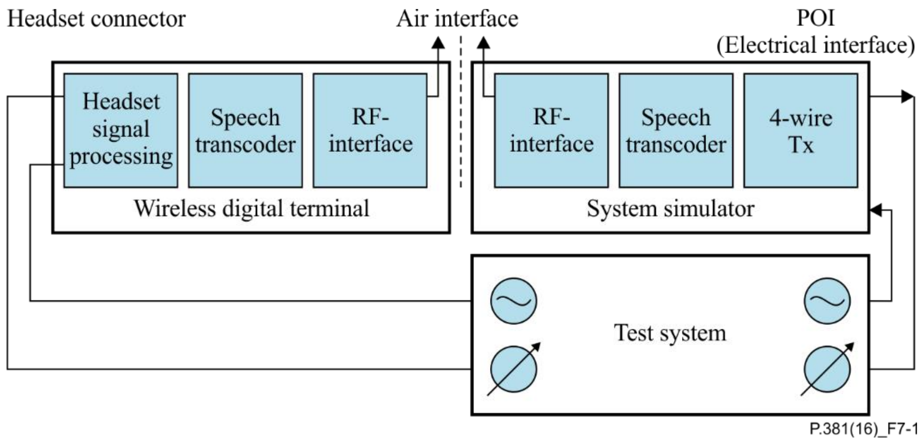
In summary, the alternative pole order from the tip to the sleeve of the headset plug is: L/R/GND/MIC.

7 Terminal electrical interface specification

7.1 Communication mode (terminal)

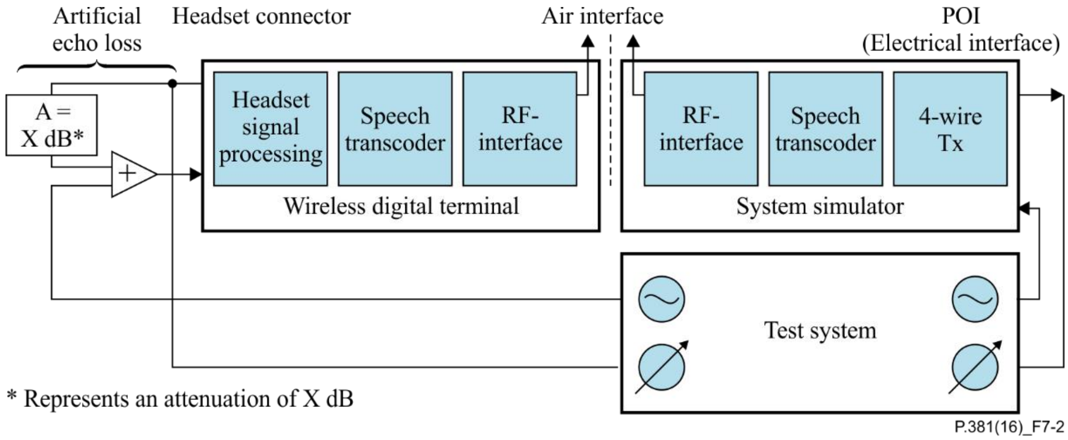
7.1.1 Test set-up for the communication mode (terminal)

Figures 7-1 and 7-2 show test set-ups.



Tx: transmitter

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



A: attenuation

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 direct current (DC) resistant. The output impedance shall be between 1 Ω and 10 kΩ. The dynamic range of the test system shall be consistent with (or exceed) the level range provided by headset MICs.

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

The common ground impedance (between sending and receiving sides) for the test system shall be ≤0.05 Ω.

7.1.1.2 Test signals and test signal levels

Unless otherwise specified, fullband (FB) 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 (CSSs) are specified, those that are speech-spectrum shaped specified in [ITU-T P.501] shall be used.

All test signals used in Receive tests have to be band limited with ≥24 dB/octave roll off, achieved by bandpass filter. In narrowband (NB) mode, signal should be band limited between 100 Hz and 4 kHz. In wideband (WB) mode, the signal should be band limited between 100 Hz and 8 kHz. In super-wideband (SWB) mode, the signal should be band limited between 50 Hz and 16 kHz. In FB mode, the signal should be band limited between 20 Hz and 20 kHz.

In Send, the test signals are used without band limitation.

For real speech, the test signal levels take as reference 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 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 dBV (decibels relative to 1 Volt (dBV) in Send (typical equivalent MIC 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 network signal levels are used, it should be noted and stated in the 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 account of terminal delays. When analysing signals, any delays introduced by the test system, codecs and terminals have to be taken into account accordingly.

7.1.2 Delay for the communication mode (terminal)

7.1.2.1 Requirements

In view of the following considerations:

– that delay has an impact on echo performance and the dynamics of voice conversation;

– that 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 the overall terminal delay in Send T_s and overall terminal delay in Receive T_r, should be less than 70 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 speech processing, data transport or handling, speech enhancement, audio filtering, etc., in both sending and receiving directions, it is obtained by excluding the implementation independent system delay from the measured overall terminal delay.

NOTE 2 – For 3GPP universal mobile telecommunications system circuit-switched speech and 3GPP long-term evolution multimedia telephony service for IMSOLR-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 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 test system delay in the Send direction ($T_{s,sys}$). The sending delay (T_s), expressed in milliseconds, is 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 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 test system delay in the Receive direction ($T_{r,sys}$).

4) The receiving delay (T_r), expressed in milliseconds, is determined from the maximum of the cross-correlation function.

7.1.3 Level in Send for nominal speech input level for the communication mode (terminal)

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 ± 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 set-up 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 decibel-milliwatts measured at a zero transmission level point.

7.1.4 Level in Receive for nominal speech input level for the communication mode (terminal)

7.1.4.1 Requirements

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

The receiving level shall be −30dBV ± 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 dBV ± 3 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 dBV ± 6 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 set-up 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 for the communication mode (terminal)

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 ($L_{sendnom}$) obtained for the nominal input level ($L_{nominal}$) 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]
$L_{nominal} - 10$	$L_{sendnom} - 5$	$L_{sendnom} - 12$
$L_{nominal} + 5$	$L_{sendnom} + 7$	$L_{sendnom} + 0$

7.1.5.2 Test

1) The test set-up 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 for the communication mode (terminal)

(Provisional, for further study.)

7.1.7 Sending frequency response for the communication mode (terminal)

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 MIC, 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 specified in Table 7-2 for NB, Table 7-3a for WB, Table 7-3b for SWB and Table 7-3c for FB.

Figure 7-3-a shows the tolerance mask for the NB sending frequency response and Figure 7-3-b shows the tolerance mask for the WB sending frequency response.

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

Frequency (Hz)	Upper limit	Lower limit	Target
100	–9	–∞	–20
200	3	–∞	–5
300	3	–3	0
1 000	5	–3	1
3 100		3	7
3 400	12	–5	4
4 000	12	–∞	

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 make allowance for a high-frequency boost considering typical variations in mouth-to-MIC transfer characteristics.

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

Frequency (Hz)	Upper limit	Lower limit	Target
NOTE 4 – The target response assumes the system under test uses the adaptive multi-rate narrowband (AMR-NB) speech transmission [b-3GPP TS 26.071] codec operating at 12.2 kbit/s.			

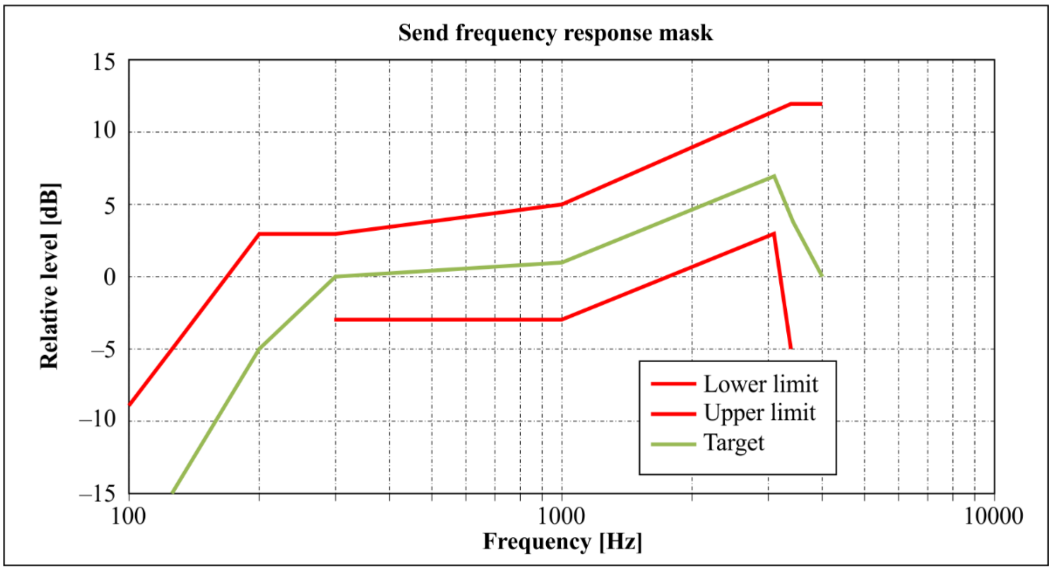


Figure 7-3a – Tolerance mask for the narrowband sending frequency response (informative)

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

Frequency (Hz)	Upper limit	Lower limit	Target
100	–3	–∞	–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	–∞	
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 make allowance for a high-frequency boost considering typical variations in mouth-to-MIC transfer characteristics. NOTE 4 – The target response assumes the system under test uses the adaptive multi-rate wideband (AMR-WB) speech transmission [b-ITU-T G.722.2] codec operating at 12.65 kbit/s.			

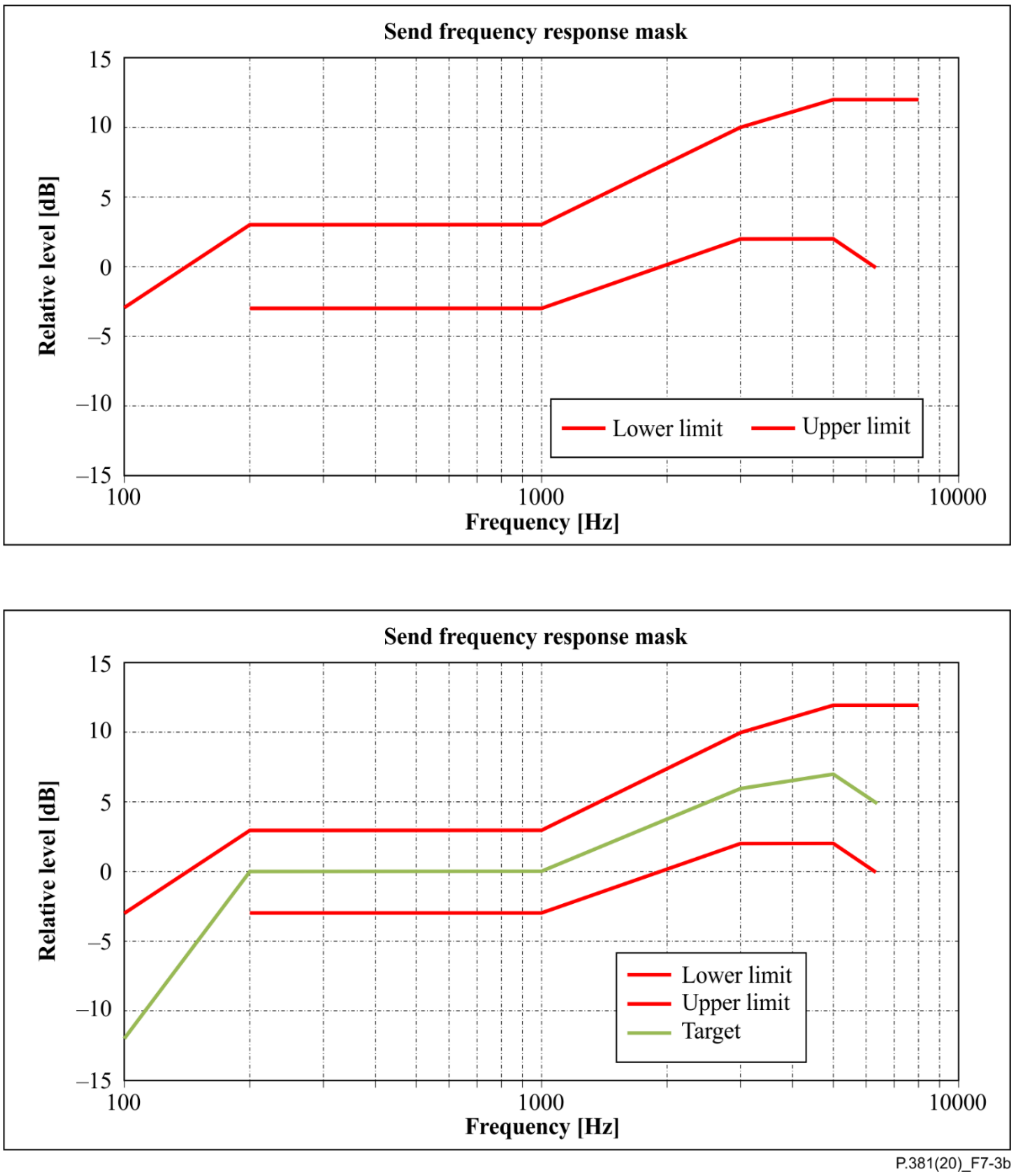


Figure 7-3b – Tolerance mask for the WB sending frequency response (informative)

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

Frequency (Hz)	Upper limit	Lower limit	Target
100	[3]		[0]
200	[3]	[–3]	[0]
1 000	[3]	[–3]	[0]
3 000	[8]	[0]	[4]
5 000	[10]	[0]	[5]
8 000	[10]	[0]	[6]
12 500	[10]	[–1]	[6]
16 000	[10]		

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

Frequency (Hz)	Upper limit	Lower limit	Target
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 make allowance for a high-frequency boost considering typical variations in mouth–to–MIC transfer characteristics. NOTE 4 – The target response assumes the system under test uses the enhanced voice service (EVS) [b–3GPP TS 26.441] codec operating in SWB mode at 24.4 kbit/s. NOTE 5 – Values within square brackets [] are provisional and require further validation.			

Table 7–3c – Tolerance mask for the fullband sending frequency response

Frequency (Hz)	Upper limit	Lower limit	Target
100	[3]		[0]
200	[3]	[–3]	[0]
1 000	[3]	[–3]	[0]
3 000	[8]	[0]	[4]
5 000	[10]	[0]	[5]
8 000	[10]	[0]	[6]
12 500	[10]	[–1]	[6]
20 000	[10]	[–3]	[4]
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 make allowance for a high-frequency boost considering typical variations in mouth-to-MIC transfer characteristics. NOTE 4 – The target response assumes the system under test uses the EVS [b-3GPP TS 26.441] codec operating in FB mode at 24.4 kbit/s. NOTE 5 – Values within square brackets [] are provisional and require further validation.			

7.1.7.2 Test

1) The test set-up 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 to determine the sending sensitivity.

3) In FB, the sending sensitivity is determined in one-third octave bands, as given by [IEC 61260-1] for frequencies from 100 Hz to 20 kHz inclusive, measured at the POI. In SWB, it is determined for frequencies from 100 Hz to 16 kHz. In WB, it is determined for frequencies from 100 Hz to 8 kHz. In NB, it is determined for frequencies from 100 Hz to 4 kHz. In each one-third octave band, the level of the measured signal takes as reference the level of the reference signal.

4) The sensitivity is determined in dBV/v.

7.1.8 Receiving frequency response for the communication mode (terminal)

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 NB, Table 7-5a for WB, Table 7-5b for SWB and Table 7-5c for FB.

Figure 7-4a shows the tolerance mask for the NB receiving frequency response and Figure 7-4b shows the tolerance mask for the WB receiving frequency response.

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

Frequency (Hz)	Upper limit	Lower limit	Target
100	2	–∞	–
200	2	–∞	0
300	2	–6	0
1 000	2	–2	0
2 000	6	–2	–1
3 400	6	–5	–1.5
4 000	6	–∞	–
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 speech transmission [b-3GPP TS 26.071] codec operating at 12.2 kbit/s.			

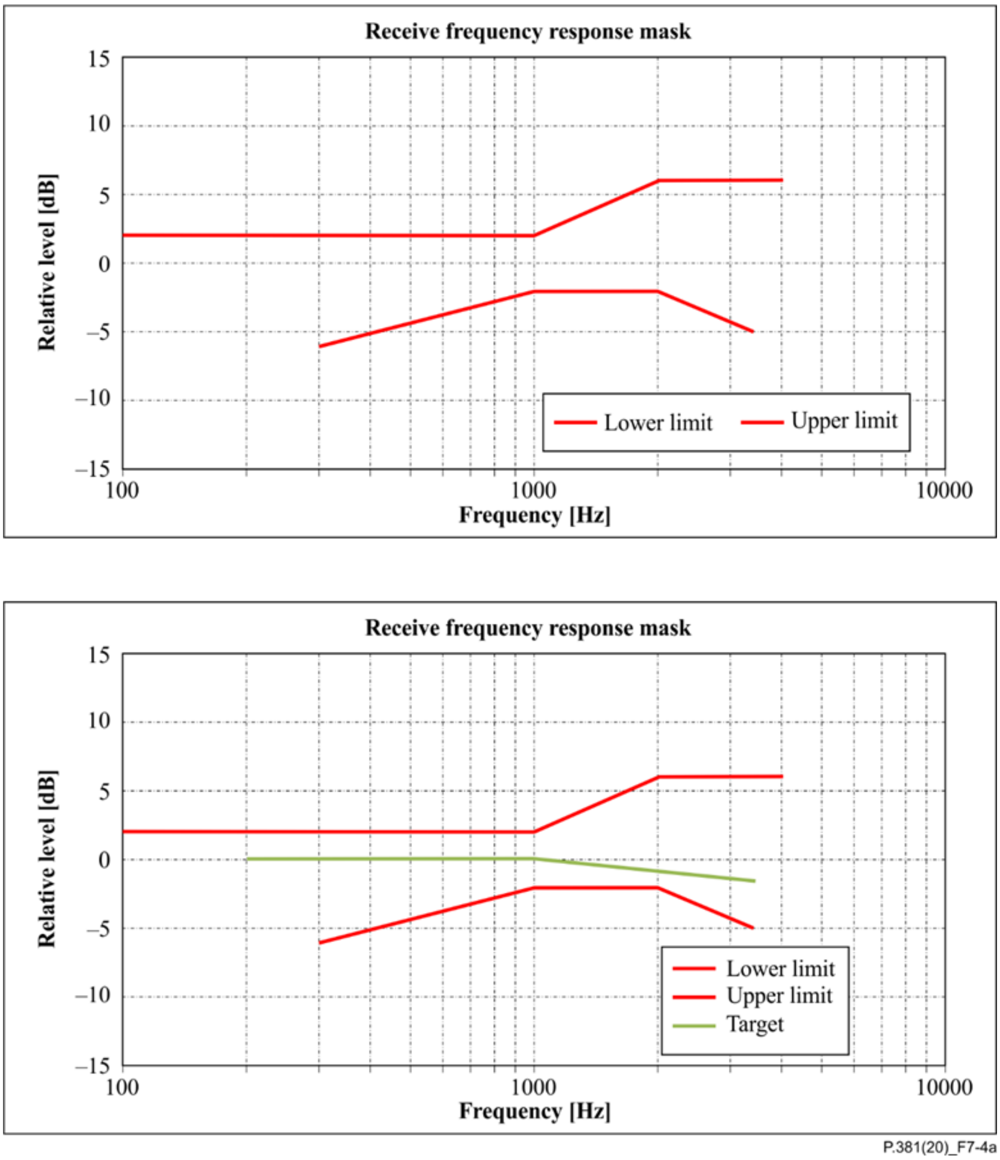


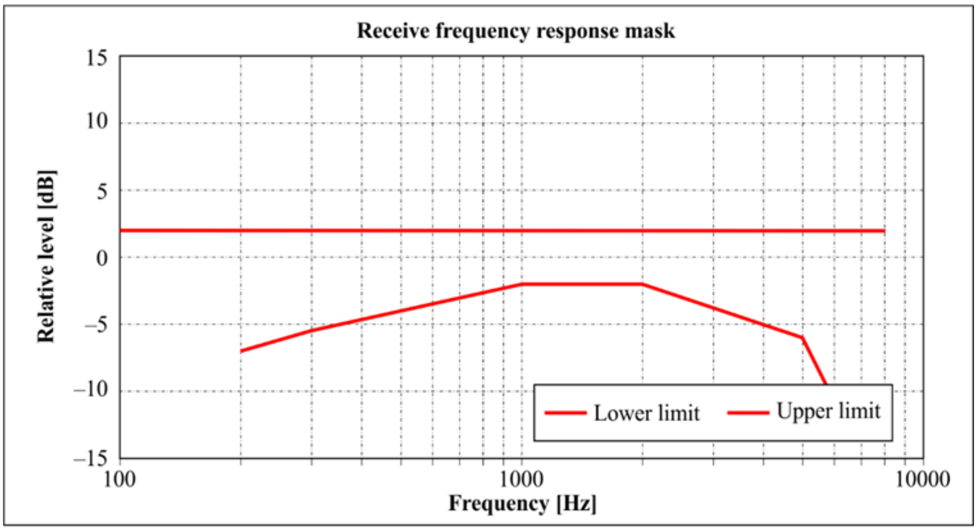
Figure 7-4a – Tolerance mask for the narrowband receiving frequency response (informative)

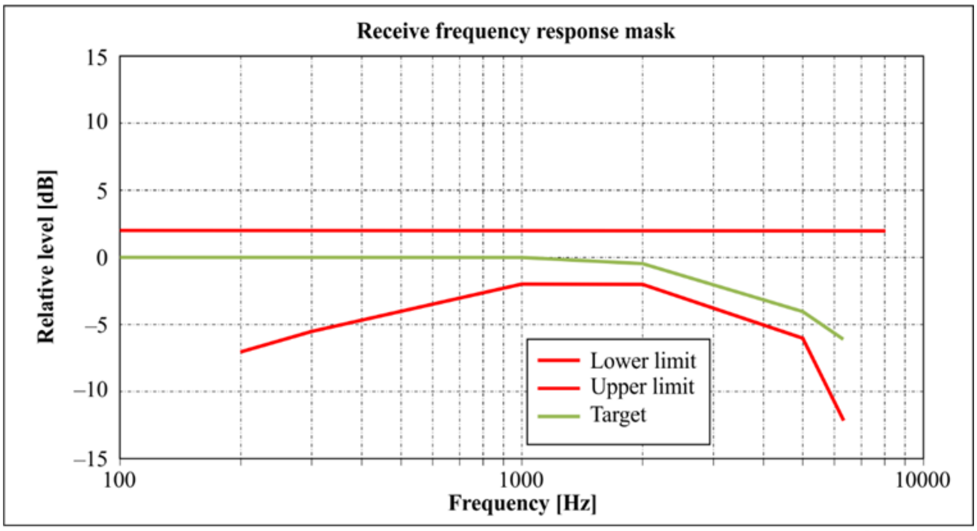
Table 7-5a – Tolerance mask for the wideband receiving frequency response

Frequency (Hz)	Upper limit	Lower limit	Target
100	2	–∞	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	–∞	–

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 speech transmission [b-ITU-T G.722.2] codec operating at 12.65 kbit/s.





P.381(20)_F7-4b

Figure 7-4b – Tolerance mask for the wideband receiving frequency response (informative)

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

Frequency (Hz)	Upper limit	Lower limit	Target
100	[2]		[0]
200	[2]	[–7]	[0]
1 000	[2]	[–5.5]	[0]
3 000	[2]	[–2]	[0]
5 000	[2]	[–2]	[0]
8 000	[2]	[–3]	[–1]
12 500	[2]	[–6]	[–2]
16 000	[2]		

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 EVS [b–3GPP TS 26.441] codec operating in SWB mode at 24.4 kbit/s.
NOTE 5 – Values within square brackets [] are provisional and require further validation.

Table 7–5c – Tolerance mask for the fullband receiving frequency response

Frequency (Hz)	Upper limit	Lower limit	Target
100	[2]		[0]
200	[2]	[–7]	[0]
1 000	[2]	[–5.5]	[0]
3 000	[2]	[–2]	[0]
5 000	[2]	[–2]	[0]
8 000	[2]	[–3]	[–1]
12 500	[2]	[–4]	[0]
20 000	[2]	[–7]	[–2]
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 EVS [b-3GPP TS 26.441] codec operating in FB mode at 24.4 kbit/s. NOTE 5 – Values within square brackets [] are provisional and require further validation.			

7.1.8.2 Test

1) The test set-up 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) In FB, the receiving sensitivity is determined in one-third octave bands as given by [IEC 61260-1] for frequencies from 100 Hz to 20 kHz inclusive, measured at the headset interface. In SWB, it is determined for frequencies from 100 Hz to 16 kHz. In WB, it is determined for frequencies from 100 Hz to 8 kHz. In NB, it is determined for frequencies from 100 Hz to 4 kHz. In each one-third octave band, the level of the measured signal takes as reference 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 for the communication mode (terminal)

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 ≥20 dB and should be ≤35 dB for the nominal setting of the volume control. For all other positions of the volume control, the STMR shall be ≥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 set-up 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 to determine the sidetone sensitivity.

3) Measurements shall be made at one one-12th octave bands 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 (bands 1 to 20 in Table 3 of [ITU-T P.79]) takes as reference 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 the ear reference point (ERP)) directions.

5) The sidetone path loss (L_{meST}), expressed in decibels, and the STMR, expressed 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 for the communication mode (terminal)

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 FB, SWB and WB, and should not exceed 5 ms in NB. 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 set-up is according to clause 7.1.1, Figure 7-1.

2) The test signal is a CSS complying with [ITU-T P.501] using a pseudo-noise (PN) sequence with a length of 85.33 ms (4 096 samples in 48 kHz sampling rate, or equivalent for other sampling rates) which equals 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:

$$\Phi_{xy}(\tau) = \frac{1}{T}\int_{-T/2}^{T/2} S_x(t) \cdot S_y(t+\tau)\mathrm{d}t \tag{7-1}$$

The measurement window T shall be identical to the time period T of the test signal, the measurement window is positioned at the start of 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 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:

$$H[xy(\tau)] = \int_{-\infty}^{+\infty} \frac{\Phi_{xy}(u)}{\pi(\tau-u)}\mathrm{d}u \tag{7-2}$$

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

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

7.1.11 Noise in Send for the communication mode (terminal)

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 ≥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 set-up 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 that potentially contribute to noise should be considered. Interference from radio frequencies is not accurately covered by an interface specification, as the complete terminal or 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 NB, between 100 Hz and 8 kHz for WB, between 100 Hz and 16 kHz for SWB and between 100 Hz and 20 kHz for FB. 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) of size 8 192 samples for a 48 kHz sampling rate or equivalent for other sampling rates. A Hann window with 75% overlap is used.
5) The noise is determined by A-weighting [IEC 61672-1] and takes as reference the reference speech signal level as determined in steps 1) and 2), resulting in a value for SNR.
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 8 192 samples for a 48 kHz sampling rate, or equivalent for other sampling rates). A Hann window with 75% overlap is used. A smoothed average idle channel noise spectrum is calculated by a moving average (arithmetic mean) one-third 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 NB, from 100 Hz to 6.3 kHz in WB, from 100 Hz to 13 kHz in SWB and from 100 Hz to 18 kHz in FB.

7.1.12 Noise in Receive for the communication mode (terminal)

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 a volume setting according to clause 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 set-up 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 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 that potentially contribute to noise should be considered. Interference from radio frequencies is not accurately covered by an interface specification, as the complete terminal or 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 for NB and WB, between 100 Hz and 16 kHz for SWB and between 100 Hz and 20 kHz for FB. 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 an FFT of size 8 192 samples for a 48 kHz sampling rate or equivalent for other sampling rates. A Hann window with 75% overlap is used.

5) The noise is determined by A-weighting [IEC 61672-1] and referring to the reference speech signal level as determined in steps 1) and 2), resulting in a value for SNR.

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 using an FFT of size 8 192 samples for a 48 kHz sampling rate or equivalent for other sampling rates. A Hann window with 75% overlap is used. A smoothed average idle channel noise spectrum is calculated by a moving average (arithmetic mean) one-third octave wide across the idle noise channel spectrum stated in dB (linear average in decibels 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 NB, from 100 Hz to 6.3 kHz in WB, from 100 Hz to 13 kHz in SWB and from 100 Hz to 18 kHz in FB.

The measurement is repeated for the second channel.

7.1.13 Sending distortion for the communication mode (terminal)

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 as described in Table 7-6.

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 (decibel) – logarithmic (hertz) scale.	

7.1.13.2 Test

1) The test set-up 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 are 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 NB, up to 7 kHz for WB, up to 14 kHz for SWB and up to 20 kHz for FB.

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 for the communication mode (terminal)

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 as described in Table 7-7.

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 set-up 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 are 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, for NB and WB, up to 20 kHz for SWB and FB.

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 for the communication mode (terminal)

7.1.15.1 Requirements

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 NB terminals, where LQOn is listening quality objective narrowband; N is noise; S is speech; and G is global:

N-MOS-LQOn Average N-MOS-LQOn ≥ 3.0

S-MOS-LQOn Average S-MOS-LQOn ≥ 3.5

G-MOS-LQOn For further study

For WB terminals, where LQOw is listening quality objective wideband:

N-MOS-LQOw Average N-MOS-LQOw ≥ 3.0

S-MOS-LQOw Average S-MOS-LQOw ≥ 3.5

G-MOS-LQOw For further study

For SWB terminals, where LQOs is listening quality objective super-wideband:

N-MOS-LQOs Average N-MOS-LQOs ≥ 3.3

S-MOS-LQOs Average S-MOS-LQOs ≥ 3.9

G-MOS-LQOs For further study

For FB terminals, where LQOf is listening quality objective fullband:

N-MOS-LQOf Average N-MOS-LQOf ≥ 3.3

S-MOS-LQOf Average S-MOS-LQOf ≥ 3.9

G-MOS-LQOf For further study

7.1.15.2 Test

The pre-recorded noisy speech signals according to Annex B shall be used for the electrical insertion. A proper conditioning sequence for convergence of noise cancellation as described in [b-ETSI TS 103 106] and [b-ETSI TS 103 281] is included.

In order to analyse S-MOS, N-MOS and G-MOS in NB and WB mode according to [b-ETSI TS 103 106], the unprocessed references (speech plus noise recorded close at the headset MIC position using a measurement MIC) are also provided in Annex B for each noise type.

1) The test set-up is according to clause 7.1.1, Figure 7-1.
2) For NB and WB speech transmission, S-MOS, N-MOS and G-MOS are calculated as described in [b-ETSI TS 103 106]. For SWB and FB speech transmission, S-MOS, N-MOS and G-MOS are calculated in accordance with model A as described in [b-ETSI TS 103 281]. In both cases, the speech level is −1.7 dBPa at the MRP.
3) In SWB and FB mode, the level calibration procedure according to clause 9.5 of [b-ETSI TS 103 281] shall be used in advance of noisy measurements (usage of silence measurement).
4) The test signal is applied to the headset interface.
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.

Table 7-8 – Noises used for background noise simulation

Description	File name	Duration (s)	Level [dB SPL(A)]	Type
Recording in public house	Pub_Noise_binaural_V2	30	L: 75,0 R: 73,0	Binaural

Table 7-8 – Noises used for background noise simulation

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

7.1.16 One-way speech quality in Send for the communication mode (terminal)

7.1.16.1 Requirements

The listening quality in the Send direction (LQs) of speech 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:

For NB the requirement for MOS-LQOf is further study.

For WB the requirement for MOS-LQOf is for further study.

For SWB the requirement for MOS-LQOf is for further study.

For FB the requirement for MOS-LQOf requirement for further study.

NOTE 1 – The purpose of this test is limited to verification of an appropriate implementation of the speech encoding operation. Speech processing in the terminal is necessary to compensate for a number of acoustic aspects. The processing, while used to provide a suitable user experience, may not result in the highest MOS-LQOf 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 MIC; 2) use of automatic gain control (AGC) to compensate for soft talkers, loud talkers, variations of positioning; 3) presence of noise suppression.

NOTE 2 – MOS-LQOf 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 NB and WB systems are evaluated on an SWB scale. Therefore the MOS requirements are given in MOS-LQOf. See [ITU-T P.863] for more information.

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

2) The test signals used are the British-English test sequences as specified in Annex C of [ITU-T P.501] (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 the 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 Appendix III of [ITU-T P.863].

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

3) The test set-up is according to clause 7.1.1.

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

7.1.17 One-way speech quality in Receive for the communication mode (terminal)

7.1.17.1 Requirements

The listening quality in the Receive direction (LQr) is measured from the POI (output of the reference speech coder of the system simulators) to the receiving output of the headset interface. The speech processing after to the decoder in the terminal shall be disabled for this test.

The listening speech quality in Receive shall be:

For NB the requirement for MOS-LQOf is further study.

For WB the requirement for MOS-LQOf is further study.

For SWB the requirement for MOS-LQOf is further study.

For FB the requirement for MOS-LQOf > is further study.

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 for a number of acoustic aspects. The processing, while used to provide a suitable user experience, may not result in the highest MOS-LQOf 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 MIC; 2) use of AGC to compensate for soft talkers, loud talkers, variations of positioning; 3) presence of noise suppression.

NOTE 2 – The MOS-LQOf 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 NB and WB systems are evaluated on a SWB scale. Therefore the MOS requirements are given in MOS-LQOf. See [ITU-T P.863] for more information.

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

2) The test signals used are the British-English test sequences as specified in Annex C of [ITU-T P.501] (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 the 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 Appendix III of [ITU-T P.863].

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

3) The test set-up is according to clause 7.1.1. The signal measured at the headset interface is used for the calculation. The MOS-LQOf 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 for the communication mode (terminal)

7.1.18.1 Requirements

The weighted terminal coupling loss (TCLw; for NB, the terminal coupling loss (TCL) for WB, SWB and FB) 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 provided by the headset signal processing in conjunction with typical echo paths, as described in Figure 7-2, shall be ≥46 dB at the volume control setting according to clause 7.1.1.2. The TCL shall also be ≥46 dB at the maximum setting of the volume control.

NOTE – A TCL ≥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 (EL) ≥50 dB.

7.1.18.2 Test

1) The test set-up 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 EL of 40 dB.

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 (six sentences) are discarded from the analysis to allow for convergence of the acoustic echo canceller (EC). The analysis is performed over the remaining length of the test sequence (last six sentences).

5) For NB mode, TCLw is calculated according to clause B.4 of [ITU-T G.122] (trapezoidal rule). In WB, SWB and FB mode, TCL is calculated as the EL from 100 Hz to 8 kHz. For the calculation, the averaged test signal level at each frequency band takes as reference the averaged measured echo signal level in each frequency band. For the measurement, a time window has to be applied that is adapted to the duration of the actual test signal. The EL is calculated by the following equations. The form of the first is generalized from that specified in clause B.4 of [ITU-T G.122] to calculate EL based on tabulated data, which allows the calculation of EL within any frequency range between f_0 and f_N.

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

$$\text{where} \quad C = 10\log_{10} 2\,(\log_{10} f_N - \log_{10} f_0),$$

where

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

A_1 is the output/input power ratio at frequency f_i;

A_N is the output/input power ratio at frequency $f_N = 8$ kHz

7.1.19 Temporal echo effects for the communication mode (terminal)

7.1.19.1 Requirements

This test aims to verify whether the system maintains sufficient echo attenuation during a time variant echo path when applying speech. The measured echo level measured with a time varying echo path shall not increase by more than 6 dB compared to the echo level observed under steady-state conditions.

7.1.19.2 Test

1) Before conducting the test, the EC shall be fully converged.

2) The test set-up is according to clause 7.1. The test signal used is the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501]. The first male sentence and the first female sentence are used. The average test signal level is −16 dBm0. The echo signal is analysed during the complete test signal. The integration time for the level analysis shall be 35 ms. The measurement result is calculated as level versus time. The echo level is determined under steady state conditions and stored as a reference.

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

4) The test is repeated simulating the time-variant echo path using the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501]. Again, the first male sentence and the first female sentence are used. The average test signal level is −16 dBm0. The echo signal is analysed during the complete test signal. The integration time for the level analysis shall be 35 ms. The measurement result is calculated as level versus time.

5) The difference of the echo level between the reference and the measured EL with the time-variant echo path is determined.

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

7.1.20 Double talk performance for the communication mode (terminal)

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, the talker echo loudness rating under double talk conditions (TELRdt) should be high and the attenuation inserted should be as low as possible. Terminals that 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.

Double talk performance may be highly influenced by the performance of the EC, especially in the non-linear processing implementation.

7.1.20.1 Attenuation range in the Send direction during double talk

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]	≤3	≤6	≤9	≤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 that is category 1 based on double talk performance is also of high quality overall.

7.1.20.1.2 Test

The test signal to determine the attenuation range during double talk is shown in Figure 7-5. The test signal to determine the attenuation range during double talk is the double talk speech sequence as specified in clause 7.3.5 of [ITU-T P.501]. The competing speaker is always inserted as the double talk sequence sdt(t) in Send and is used for analysis.

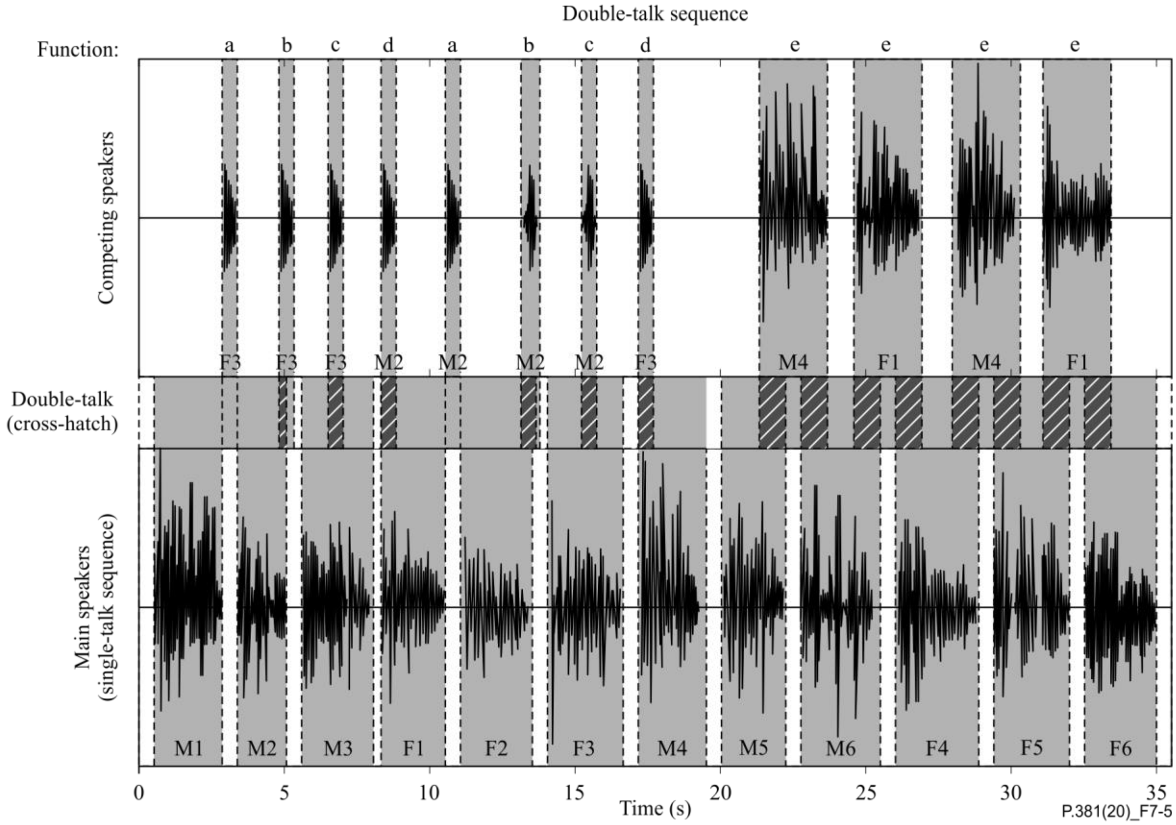


Figure 7-5 – Double talk test sequence with overlapping speech sequences in Send and Receive

Figure 7-5 indicates that the sequences overlap partially. The test signals are synchronized in time at the electrical interface. The delay of the test set-up shall be constant during the measurement. The settings for the test signals are as listed in Table 7-10.

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

	Receive (sdt(t))	**Send** (s(t))
Average signal level	−16 dBm0	−60 dBV

1) The test set-up is according to clause 7.1.1, Figure 7-2; the test signal is shown in Figure 7-5. The test is performed using an artificial EL of 40 dB. Before the actual test, a training sequence for the EC consisting of the British-English single talk sequence described in clause 7.3.2 of [ITU-T P.501] with a level of −16 dBm0 is applied to the POI.

2) When determining the attenuation range in Send, the signal measured at the POI takes as reference the test signal inserted.

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

4) The categorization is done according to Appendix III of [ITU-T P.502].

7.1.20.2 Attenuation range in the Receive direction during double talk

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]

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

In general, Table 7-11 provides a quality classification of terminals regarding double talk performance. However, this does not mean that a terminal that is category 1 based on double talk performance is also of high quality overall.

7.1.20.2.2 Test

The test signal to determine the attenuation range during double talk is shown in Figure 7-5. The test signals are synchronized in time at the electrical interface. The delay of the test arrangement shall be constant during the measurement.

The settings for the test signals are listed in Table 7-12.

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

	Receive (sdt(t))	**Send** (s(t))
Average signal level	−16 dBm0	−60 dBV

1) The test set-up 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 is the test signal inserted.

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

4) The categorization is done according to Appendix III of [ITU-T P.502].

7.1.20.3 Detection of echo components during double talk

7.1.20.3.1 Requirements

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

NOTE – Echo attenuation during double talk is based on the value of the TELRdt. It is assumed that the terminal at the opposite end of the connection provides the nominal loudness rating (SLR + Receive loudness rating (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]

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

7.1.20.3.2 Test

1) The test set-up is according to clause 7.1.1, Figure 7-2.
2) The double talk signal consists of a sequence of orthogonal signals that 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 NB, the NB test signals are used; for WB, the WB test signals as described in [ITU-T P.501] are used. For SWB and FB, the FB test signals as described in [ITU-T P.501] are used, observing the filtering described in clause 7.1.1.2.
3) The signals are fed simultaneously in Send and Receive. The level in Send at the headset interface is −60 dBV (nominal level); that in Receive at the POI is −16 dBm0 (nominal level).
4) The test signal is measured at the receiving interface and consists of the double talk signal that was fed in at the sending interface and the echo signal. The echo signal is filtered by a 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 suppresses frequency components of the double talk signal.
5) For each frequency band that 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 below either the signal noise or the required limit. If echo components are detectable, the classification is based on Table 7-12. Echo attenuation is to be achieved for each individual frequency band according to the different categories.

7.1.21 Activation in Send for the communication mode (terminal)

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 ≤ −75 dBV when analysing a complete CSS burst.

$L_{s,minonset}$ should be ≤ [TBD] dBV when analysing the onset of a CSS burst (first 50 ms).

7.1.21.2 Test

The structure of the test signal is shown in Figure 7-6. The test signal consists of CSS components according to clause 7.2.1.2 of [ITU-T P.501], with increasing levels for each CSS burst.

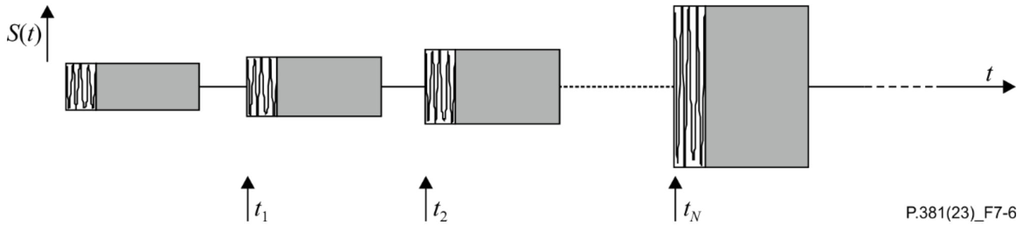


Figure 7-6 – Test signal to determine the minimum activation level and the build-up time

The settings of the test signal are listed in Table 7-14.

Table 7-14 – Settings of the CSS in Send

	CSS duration/ minimum pause duration (ms)	**Level of the first CS signal (active signal part at the sending input of the headset interface)** (dBV)	**Level difference between two periods of the test signal** (dB)
CSS to determine switching characteristic in Send	248.62/451.38	−78.3 (Note)	1
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 set-up 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 reference level at POI is the level 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 test signal level.
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 (terminal)

7.2.1 Test set-up for the multimedia playback mode (terminal)

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

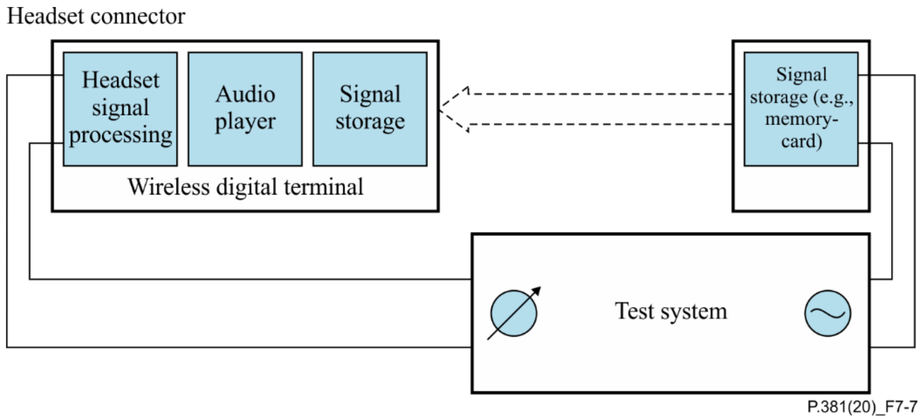


Figure 7-7 – Set-up for testing the electrical headset interface

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 Ω. The dynamic range shall be consistent with (or exceed) the output level range provided by the electrical output of the digital mobile terminal headset outputs.

If the MIC 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 Ω and 10 kΩ. The dynamic range shall be consistent with (or exceed) the level range provided by headset MICs.

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 quoted in this clause are relative to decibels relative to full scale (dBFS), where 0 dBFS represents the root mean square (RMS) level of a full-scale sinusoidal signal.

Programme simulation noise as specified 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 that are used in Receive have to be band limited. The band limitation is achieved by a low-pass filter up to 22 kHz providing ≥24 dB/octave filter roll-off. The programme simulation noise according to [EN 50332-1] is band limited by design and requires no filtering.

All test signal levels are 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 terminal delays into account. 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 for the multimedia playback mode (terminal)

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 ± 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 or regulations may require deployment of further measures to inform about or reduce the risk of hearing damage.

7.2.2.2 Test

1) The test set-up is according to clause 7.2.1, Figure 7-7.
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 the 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 for the multimedia playback mode (terminal)

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 that in combination with the digital mobile terminal should comply with the relevant standards in Receive.

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

Table 7-15 – Tolerance mask for frequency response for the 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 decibels on an arbitrary scale. The limits for intermediate frequencies lie on straight lines drawn between the given values on a linear (decibel) – logarithmic (hertz) scale.		

7.2.3.2 Test

1) The test set-up is according to clause 7.2.1, Figure 7-7.
2) The test signal used for the measurements shall be a FB 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 one-third octave bands as given by [IEC 61260-1] for frequencies between 50 Hz and 16 kHz, inclusive. In each one-third octave band, the level of the measured signal takes as reference 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 for the multimedia playback mode (terminal)

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 ≥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 set-up is according to clause 7.2.1, Figure 7-7.
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 least significant bit 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 laboratory 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 an FFT of size 8 192 samples for a 48 kHz sampling rate or equivalent for other sampling rates. A Hann window with 75% overlap is used.
5) The noise level is determined by applying A-weighting [IEC 61672-1] and then referred to the reference signal level as determined in step 2), resulting in a value for SNR. Spectral peaks are measured in the frequency domain. The average noise spectrum is used as a threshold to determine spectral peaks and should be calculated as the arithmetic mean of the noise spectrum values in dBV.

The measurement is repeated for the second channel.

7.2.5 Distortion for the multimedia playback mode (terminal)

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 masks listed 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 (decibel) – logarithmic (hertz) scale.	

7.2.5.2 Test

1) The test set-up is according to clause 7.2.1, Figure 7-7.
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 are 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 for the multimedia playback mode (terminal)

7.2.6.1 Requirements

The receiving crosstalk is measured as that L-R and R-L 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.

Over a period of 0 s to 5 s, the attenuation measured at the right output, referenced to the spectrum generated at the left output shall be above 30 dB. Over a period of 5 s to 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 ≥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

Period (s)	**Level**	
	Left channel (dBFS)	**Right channel (dBFS)**
0-5	–10	–∞
5-10	–∞	–10

7.2.6.2 Test

1) The test set-up is according to clause 7.2.1, Figure 7-7.
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 listed in Table 7-17.
3) Playback the test signal, the crosstalk is determined by analysing the measured signal at the output of the headset interface. Over a period of 0 s to 5 s, the measured level at the right channel takes as reference the level at the left channel, and the attenuation is L-R crosstalk. Over a period of 5 s to 10 s, the measured level at the left channel takes as reference 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 (headset) for the communication mode

8.1.1 Test set-up (headset)

The test set-up is shown in Figure 8-1 and the input connection of a headset reference interface is shown in Figure 8-2.

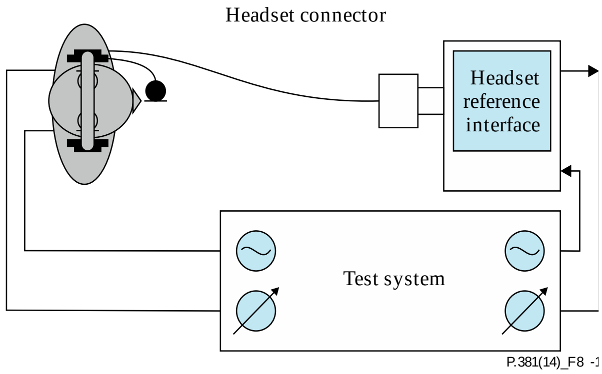
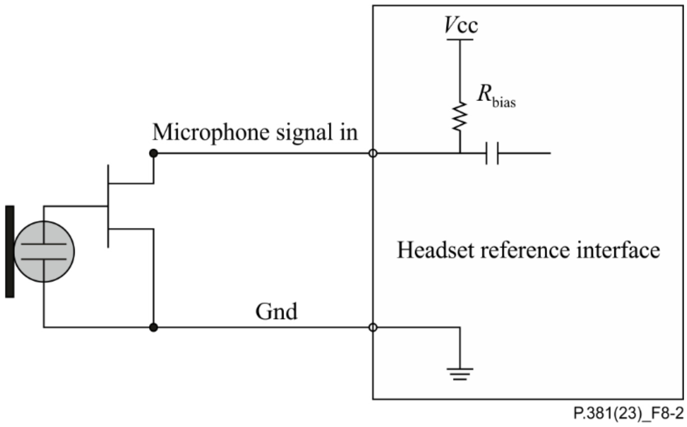


Figure 8-1 –Set-up for testing the headset



V_{cc}: voltage at the common collector

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

A head and torso simulator (HATS) according to [ITU-T P.58] shall be used as an acoustical interface for testing. As specified in clause 7 of [ITU-T P.380], it shall be equipped with one of the following artificial ear types according to [ITU-T P.57]:

- type 4.3;
- type 4.4;
- type 3.3;
- type 3.4.

For measurement of binaural headsets, the HATS shall be equipped with both left and right artificial ears.

If artificial ears of type 3 are used and produce measurement results that differ from those obtained with type 4, the type 4 results shall take precedence.

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

The output impedance shall be <2 Ω. To prevent damage to the headset during testing, the voltage applied to the Receive terminals shall not exceed 1.9 V peak to peak (derived from 150 mV for −10 dBFS, considering the maximum digital peak level and an additional 3 dB allowance).

The bias voltage provided by the test system shall be 2.6 V ± 1%, see clause IV.4 for more details. Bias resistance (R_{bias}) is the bias resistance inside the input of the test system. The bias resistance shall be 2.2 kΩ ± 2%. The nominal level shall be −60 dBV (expected from a headset with a nominal sensitivity of −55 dBV/Pa).

NOTE – The nominal Send sensitivity mentioned above includes the acoustic loss from MRP to the MIC of the headset.

8.1.1.2 Test signals and test signal levels

Unless otherwise specified, FB 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]. For test cases where CSSs are specified, those that are speech-spectrum shaped specified in [ITU-T P.501] shall be used.

All test signals used in Receive have to be band limited with ≥24 dB/octave roll off. The signal should be between 50 Hz and 8 kHz. In Send, the test signals are used without band limitation.

Unless described otherwise, all test signal levels are averaged over the complete test sequence.

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

- −37 dBV in Receive for binaural headsets;
- −31 dBV in Receive for monaural headsets;
- −4.7 dBPa at the MRP.

NOTE – The signal levels are derived from an assumption of a headset Receive sensitivity of simulated programme signal characteristic voltage (SPCV) = 150 mV, which is within the range specified in clause 8.2.2. A flat diffuse-field-corrected frequency response has been assumed. The overall nominal Receive path RLR for headsets of 2 dB (monaural) or 8 dB (binaural) has been considered, taking into account the non-flat response at the ERP.

Some tests require exact synchronization of test signals in the time domain. Therefore, it is required to take headset delays into account. 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 a type 3.3 or 3.4 artificial ear. For such insert type headsets, an ITU-T P.57 type 4 artificial ear is highly recommended (see also notes to clause 7 of [ITU-T P.380]).

When a headset is placed on a HATS, the results may vary from trial to trial due to slight variations in positioning. Relatively accurate and repeatable results can be obtained by making several measurements and averaging the results according to [ITU-T P.380]. Unless stated otherwise, the measurements in Receive should be repeated five times and averaged. The headset shall be completely removed from the artificial ear and re-mounted for each trial. The averaged result is used.

8.1.1.4 Position and calibration of HATS

The calibration and equalization procedures, as well as the combination method for signals from two ears can be found in [ITU-T P.581]. Unless stated otherwise, the measurements from HATS at the eardrum reference point (DRP) shall be corrected to diffuse field. The reverse of the diffuse field to DRP function found in Table 3 or Annex A of [ITU-T P.58] shall be used for bands of one-third octave and one-12th octave, respectively. For measurements requiring diffuse-field correction values in closer frequency spacing than that specified in [ITU-T P.58], the interpolation method found in Annex A of this Recommendation shall be used.

8.1.2 Sensitivity in Send (headset)

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 dBV/Pa ± 6 dB when inserting the sending signal at the nominal level, as described in clause 8.1.1.2.

NOTE – The Send sensitivity specified above includes the acoustic loss from MRP to the MIC of the headset.

8.1.2.2 Test

1) The test set-up 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 averaged over the complete test signal.

   The measured power density spectrum at the MRP is used as the reference power-density spectrum to determine 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 (headset)

8.1.3.1 Requirements

The requirements are specified in clause 8.2.2.

NOTE – The Receive part of a headset may be a completely passive device that does not change its characteristics depending on multimedia or communication modes. The Receive sensitivity is therefore only specified in the context of one of these modes, to avoid double and potentially conflicting requirements. The Receive sensitivity is specified for the multimedia case, considering music playback safety standards for headsets or HPs.

8.1.3.2 Test

The measurement method is specified in clause 8.2.2.

8.1.4 Sending frequency response (headset)

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 specified in Table 8-1a. If supporting SWB or FB speech, the measured frequency response shall also comply with the limits specified in Table 8-1b.

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

Frequency (Hz)	**Upper limit**	**Lower limit**
100	4	−∞
200	4	−4
1 000	4	−4
3 000	4	–
8 000	−1	−15
NOTE – All sensitivity values are expressed in decibels on an arbitrary scale. The limits for intermediate frequencies lie on straight lines drawn between the given values on a linear (decibel) – logarithmic (hertz) scale.		

Table 8-1b – Tolerance mask for the super-wideband and fullband sending frequency response

Frequency (Hz)	**Upper limit**	**Lower limit**
100	4	
200	4	–4
1 000	4	–4
3 000	4	–4
12 500	–4	–20
16 000	–5.5	
NOTE – All sensitivity values are expressed in decibels on an arbitrary scale. The limits for intermediate frequencies lie on straight lines drawn between the given values on a linear (decibel) – logarithmic (hertz) scale.		

8.1.4.2 Test

1) The test set-up 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 to determine the sending sensitivity.
3) The sending sensitivity is determined in one-third octave bands, as given by [IEC 61260-1] for frequencies between 100 Hz and 8 kHz inclusive, measured at the POI. In each one-third octave band, the level of the measured signal takes as reference 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 (headset)

8.1.5.1 Requirements

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

The measured frequency response shall be within the limits specified in Table 8-2a. If supporting SWB or FB speech, the measured frequency response shall also comply with the limits specified in Table 8-2b.

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

Frequency (Hz)	**Upper limit**	**Lower limit**
100	12	−∞
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	−∞
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.		

Table 8-2b – Tolerance mask for the super-wideband and fullband receiving frequency response

Frequency (Hz)	**Upper limit**	**Lower limit**
100	12	
200	10	–10
300	9	–6
1 000	6	–6
2 000	8	–6
5 000	8	–6
12 500	8	–12
20 000	8	

8.1.5.2 Test

1) The test set-up 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 WB, the receiving sensitivity is determined in one-third octave bands as given by [IEC 61260-1] for frequencies between 100 Hz and 10 kHz inclusive, measured at the headset interface. If supporting SWB or FB speech transmission, it is determined between 100 Hz and 20 kHz inclusive. In each one-third octave band, the level of the measured signal takes as reference 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 (headset)

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).

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

8.1.6.2 Test

1) The test set-up 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 that potentially contribute to noise should be considered. Interference from radio frequencies is not accurately covered by an interface specification, as the complete terminal or 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 idle channel noise is measured at the output in the frequency range between 100 Hz and 8 kHz. If supporting SWB or FB speech, it is determined between 100 Hz and 20 kHz inclusive. The length of the time window is 1 s, which is the averaging time for the idle channel noise. The test laboratory 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 an FFT of size 8 192 samples for a 48 kHz sampling rate or equivalent for other sampling rates. A Hann window with 75% overlap is used.

5) The idle channel noise level is determined by applying A-weighting [IEC 61672-1] and then referred to the reference speech signal level as determined in step 2). Spectral peaks are measured in the frequency domain. The average noise spectrum is used as a threshold to determine spectral peaks and is calculated as the arithmetic mean of the noise spectrum values in dBV.

8.1.7 Distortion in Send (headset)

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 masks listed 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 (decibel) – logarithmic (hertz) scale.	

8.1.7.2 Test

1) The test set-up 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 are 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 the 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. If supporting SWB or FB speech, it is measured selectively up to 20 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 (headset)

8.1.8.1 Requirements

The weighted headset terminal coupling loss (HTCLw; for NB speech transmission, headset terminal coupling loss (HTCL) for WB, SWB and FB) is measured at the headset reference interface, from receiving to sending.

The HTCLw provided by the headset shall be ≥40 dB.

8.1.8.2 Test

1) The test set-up 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.

3) 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°.

4) For the measurement, a time window has to be applied that is adapted to the duration of the actual test signal (250 ms).

5) For NB mode, HTCLw is calculated according to clause B.4 of [ITU-T G.122] (trapezoidal rule). In WB, SWB and FB mode, the HTCL is calculated as the EL from 100 Hz to 8 kHz. For the calculation, the averaged test signal level at each frequency band takes as reference the averaged measured echo signal level in each frequency band. The EL is calculated by the following equations. The form of the first is generalized from the equation specified in clause B.4 of [ITU-T G.122] to calculate EL based on tabulated data, which allows the calculation of EL within any frequency range between f_0 and f_N.

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

$$\text{where} \quad C = 10\log_{10}[2\,(\log_{10} f_N - \log_{10} f_0)],$$

where

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

A_1 is the output/input power ratio at frequency f_i;

A_N is the output/input power ratio at frequency $f_N = 8\,000$ Hz.

8.2 Multimedia playback mode (headset)

8.2.1 Test set-up (headset)

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

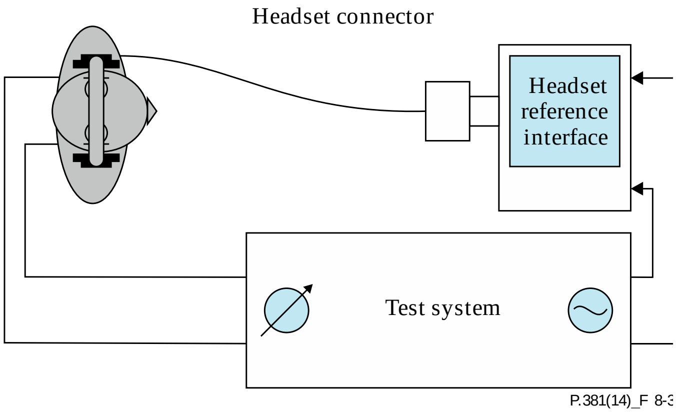


Figure 8-3 –Set-up for testing the headset

The same requirements for HATS and artificial ears as specified in clause 8.1.1 apply.

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

The output impedance shall be <2 Ω. The maximum RMS output voltage shall be 150 mV ± 1 mV when loaded with a 32 Ω resistor. The common ground impedance (between sending and receiving sides) for the test system shall be ≤0.05 Ω.

8.2.1.2 Test signals and test signal levels

The programme simulation noise specified in [EN 50332-1] is used for the measurements.

Artificial test signals, which are used in Receive, have to be band limited. The band limitation is achieved by low-pass filtering in the frequency up to 22 kHz that provides at least 24 dB/octave roll off. The programme simulation noise according to [EN 50332-1] is band limited by design and requires no filtering.

Unless described otherwise, all test signal levels are averaged over the complete test sequence length.

The nominal average signal level for the measurements is −32 dBV.

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

8.2.1.3 Positioning of the headsets

The same guidelines and requirements for headset positioning apply as in clause 8.1.1.3.

8.2.1.4 Position and calibration of HATS

The same requirements on position and calibration of HATS apply as specified in clause 8.1.1.4.

8.2.2 Sensitivity for the multimedia playback mode (headset)

8.2.2.1 Requirements

The receiving sensitivity is measured from the receiving output of the headset reference interface to the DRP of HATS and then corrected to the diffuse field 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 SPCV. This is the voltage of a specific programme simulation noise signal required to produce 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 mV ≤ SPCV ≤ 300 mV when measured according to [EN 50332-2].

8.2.2.2 Test

1) The test set-up 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 for the multimedia playback mode (headset)

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 masks listed 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 set-up 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 are 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 (headset)

8.2.4.1 Requirements

The receiving crosstalk is measured as that L-R and R-L 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.

Over a period of 0 s to 5 s, the attenuation, measured at the right ear and referenced to the level at the left ear, shall be above 20 dB. Over a period of 5 s to 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 ≥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

Period (s)	**Level**	
	Left channel (dBV)	**Right channel** (dBV)
0-5	−22	0
5-10	0	−22

8.2.4.2 Test

1) The test set-up 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 listed in Table 8-5.
3) The test signal is output to the headset, the crosstalk is determined by analysing the measured signal at the output of artificial ear. Over a period of 0 s to 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. Over a period of 5 s to 10 s, the measured level at the left ear is referenced to that at the right, and the attenuation is R-L crosstalk.
4) The crosstalk is determined as 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 or four poles.
4) Some headsets have function or /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 as listed in Table 9-1.

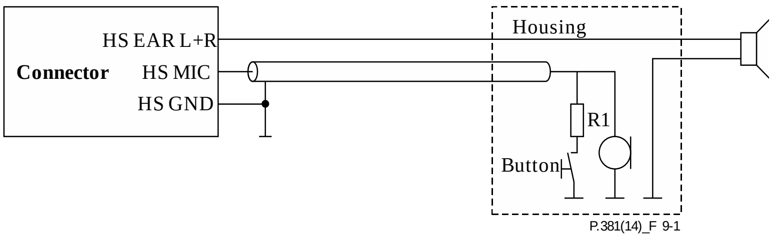


Figure 9-1 – Button-control headset function

NOTE – For harmonization with some commercially available headsets and 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 or end button detection by the terminal and recommended values for the headset

Resistance range	Min threshold for detection by the terminal [Ω]	Headset typical [Ω]	Max threshold for detection by the terminal [Ω]
End/play/pause	0	0	70
Voice control	110	135	180
Vol+	210	240	290
Vol−	360	470	680
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 Ω bias network).			

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 one-12th octave bands, linear interpolation on a log scale from one-12th octave band interpolated values in Table 14b of [ITU-T P.58] shall be used.

Annex B

Recording procedure using a representative headset

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

Headset used for recording

Figure B.1 provides the sending frequency response of the representative (or reference) headset used for noisy speech recordings. In addition, the lower and upper limits of clause 8.1.4 (headset Send sensitivity) are provided. Note that the transfer function complies with the tolerance required in clause 8.1.4.

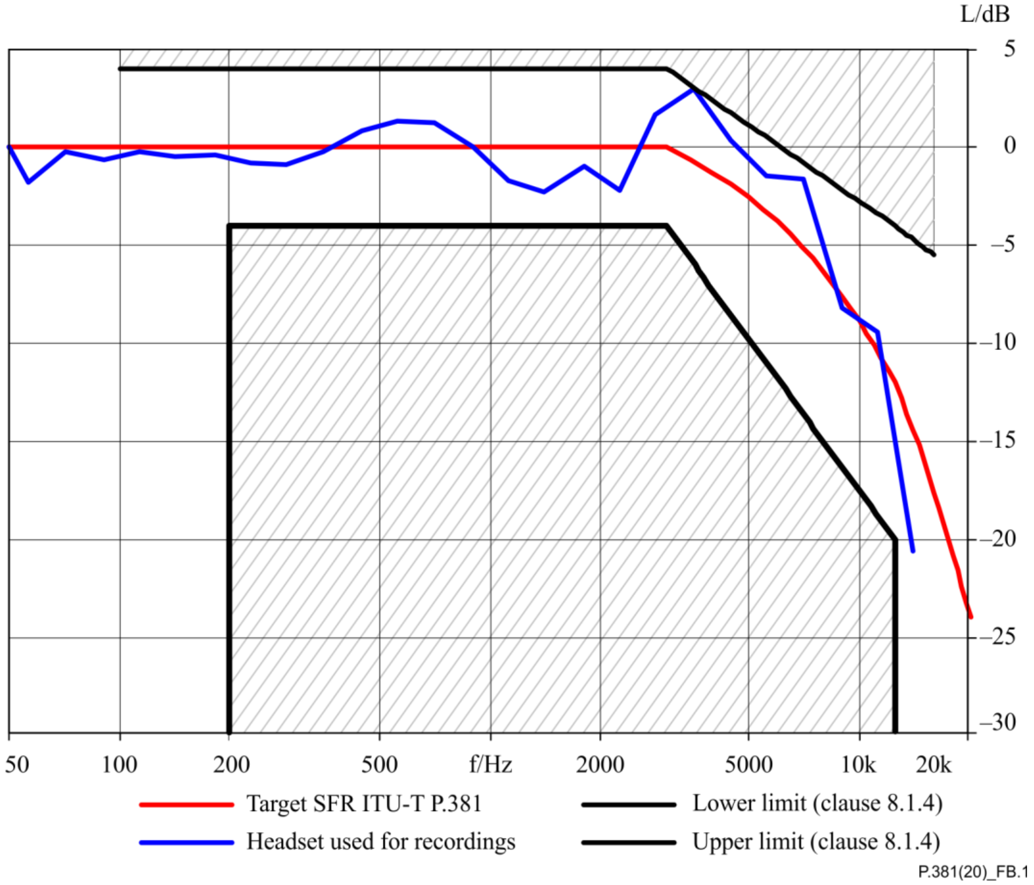


Figure B.1 – Sending frequency response of reference headset versus target curve (clause 8.1.4)

It should be noted that the target SFR is extrapolated via the tolerances given in Table 8-1b.

Background noise

Table B.1 lists the eight noise types used for testing according to clause 7.1.15 (Table 7-8).

Table B.1 – Noise types used in clause 7.1.15

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

Each noise was generated according to [b-ETSI EG 202 396-1]. The speech test sequence required for calculating S-MOS, N-MOS and G-MOS according to Annex C of [b-ETSI TS 103 106] was used with an active speech level at MRP of −1.7 dBPa for noisy and −4.7 dBPa for silent conditions. The speech signal was produced simultaneously with the noises by the artificial head and recorded at the output of the reference headset. In order to comply with the nominal level described in clause 7.1.1.2, the sensitivity was adjusted in such a way that −4.7 dBPa at the MRP corresponded to −60 dBV at the electrical interface.

For the calculation of S-MOS, N-MOS and G-MOS according to [b-ETSI TS 103 106], the unprocessed reference speech signal is also required. These signals were recorded with a measurement MIC close to the input MIC of the headset.

The recordings are exported to wave files and are provided as an electronic attachment to this Recommendation, available from https://www.itu.int/net/itu-t/sigdb/genaudio/AudioForm-g.aspx?val=10000381. Table B.2 provides the filenames and the A-weighted level for each signal. Here the identifier (ID) column corresponds to Table B.1.

For the calculation of S-MOS, N-MOS and G-MOS according to [ETSI TS 103 281], additional recordings under silent conditions are required. These signals are provided with an ID of 0 in Table B.2.

Table B.2 – Noise types used in clause 7.1.15

ID	Filename for electrical insertion	Level [dB V(A)]	Filename for unprocessed reference	Level [dB SPL(A)]
0	prerecorded_silent.wav	–65.9	unprocessed_silent.wav	75.2
1	prerecorded_pub.wav	–64.1	unprocessed_pub.wav	77.1
2	prerecorded_road.wav	–64.0	unprocessed_road.wav	77.2
3	prerecorded_xroad.wav	–65.3	unprocessed_xroad.wav	75.8
4	prerecorded_train.wav	–65.2	unprocessed_train.wav	75.9
5	prerecorded_car.wav	–65.1	unprocessed_car.wav	76.1
6	prerecorded_cafecounter.wav	–65.3	unprocessed_cafecounter.wav	76.0
7	prerecorded_mensa.wav	–65.7	unprocessed_mensa.wav	75.5
8	prerecorded_callcenter.wav	–65.9	unprocessed_callcenter.wav	75.3

Since the physical domain of these two types of necessary signals are quite different, the wave files are calibrated as follows:

– pre-recorded source signals for electrical insertions: −26 dBov of wave file corresponds to 60 dBV;

– unprocessed reference signals (close to the MIC of the DUT): −26 dBov of wave file corresponds to −15 dBPa.

NOTE – An electronic attachment holding the signal that fulfils the requirements as illustrated in Figure B.1 is available as a part of this Recommendation, and is available at https://www.itu.int/net/itu-t/sigdb/genaudio/AudioForm-g.aspx?val=10000381.

Appendix I

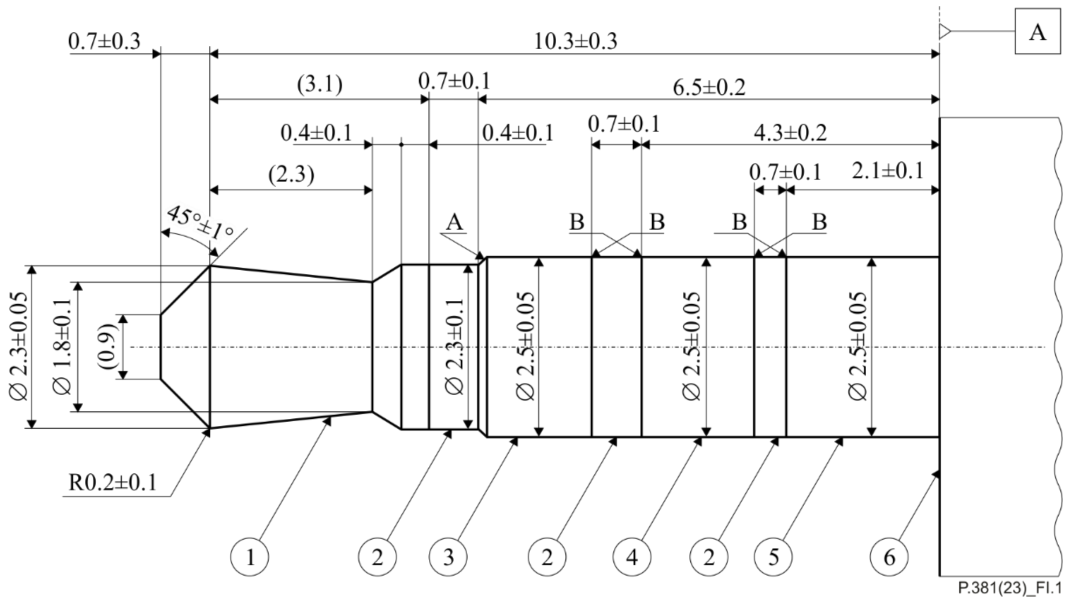
Audio connectivity for sockets with four contact points

(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 flash.



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 – "Flash" 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.

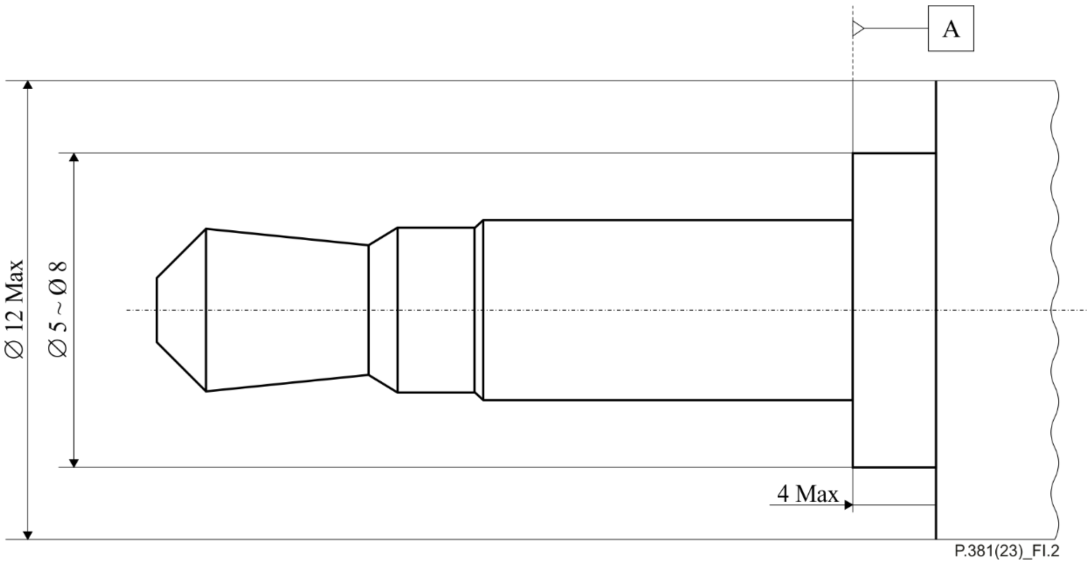


Figure I.2 – Dimensions of the 2.5 mm diameter plug connector grip

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 the 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.

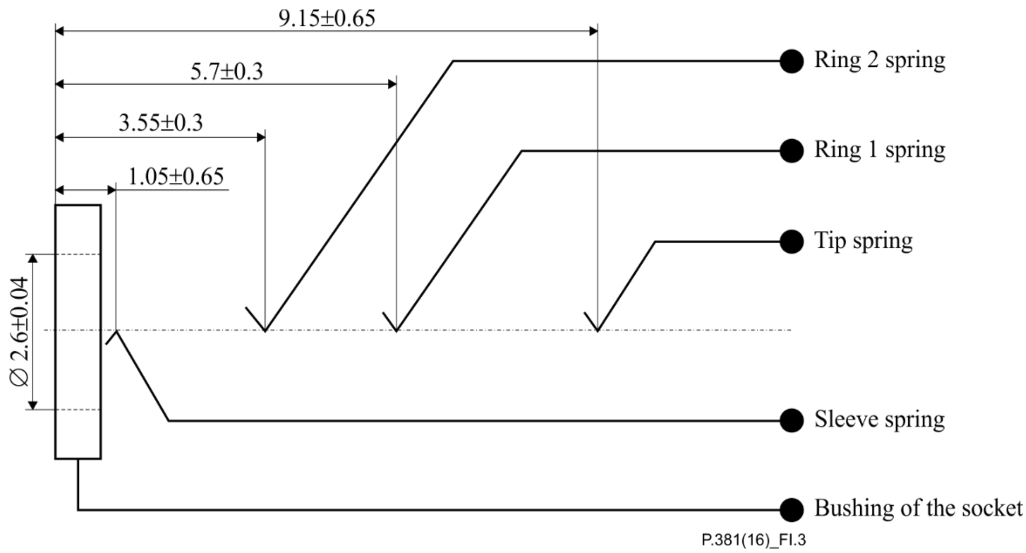
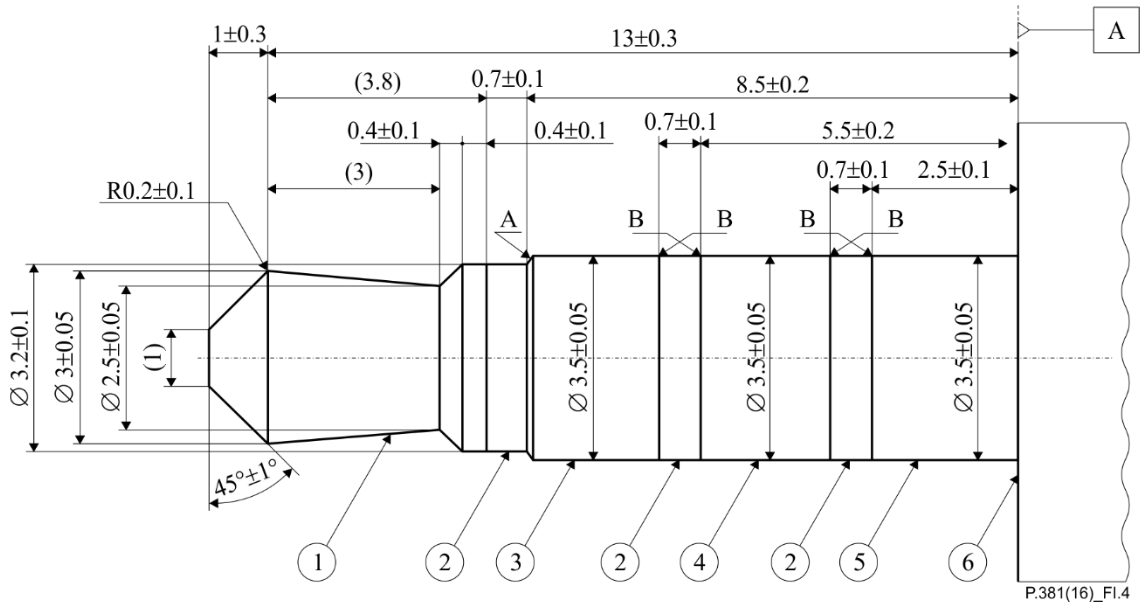


Figure I.3 – Dimensions of the 2.5 mm diameter socket with four contact points and positioning of each contact spring

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 flash.



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 – "Flash" 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.

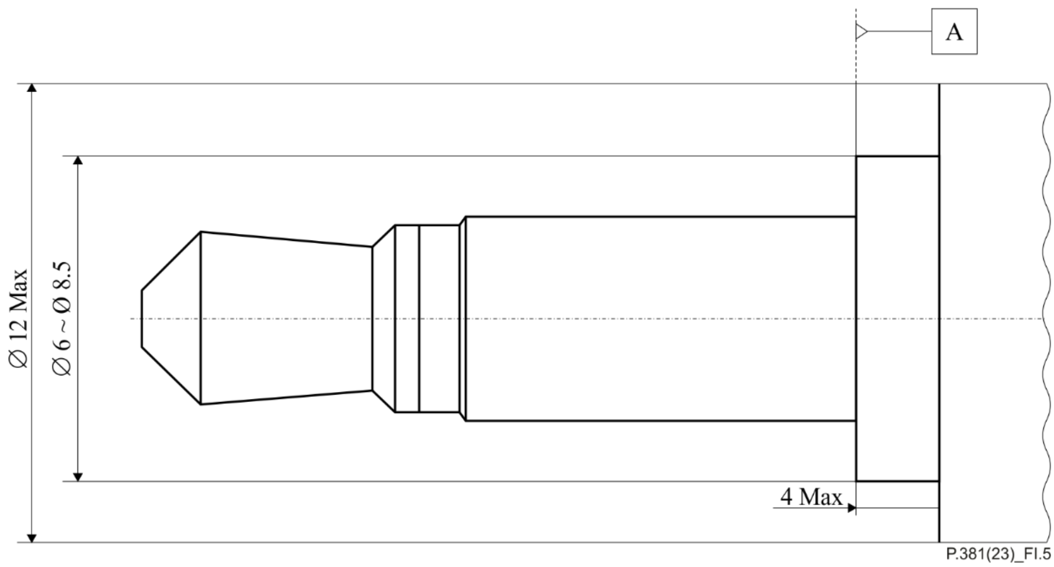


Figure I.5 – Dimensions of the 2.5 mm diameter plug connector grip

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 the 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.

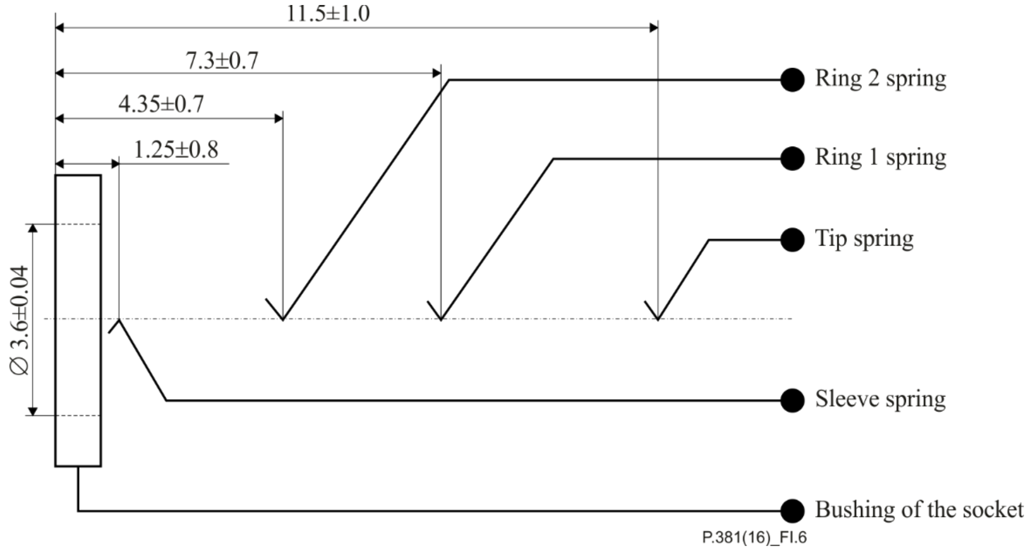


Figure I.6 – Dimensions of the 3.5 mm diameter socket with four contact points and positioning of each contact spring

Appendix II

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

This appendix has intentionally been left blank.

Appendix III

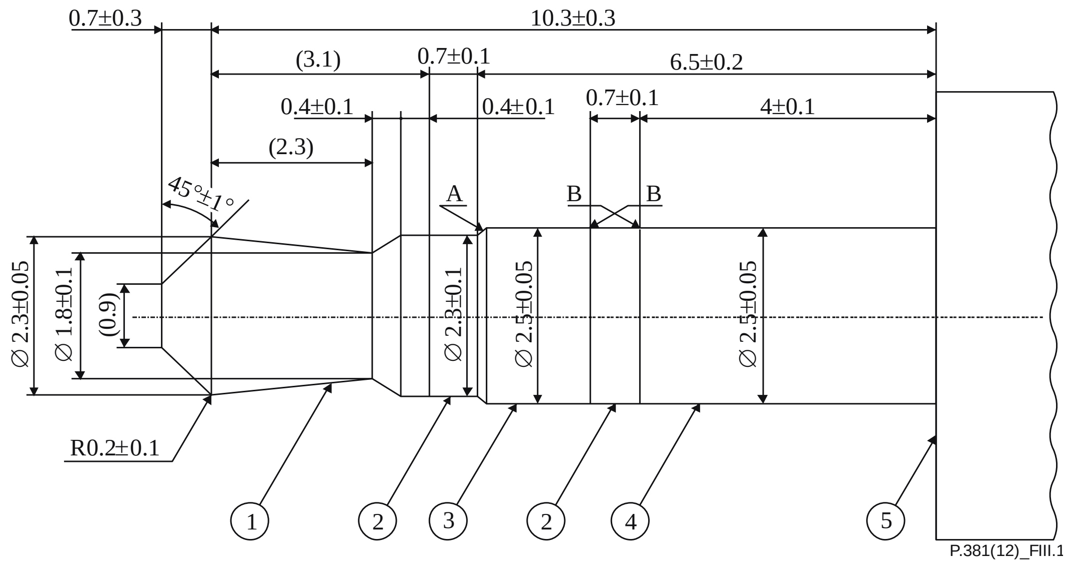
Audio connectivity for sockets with three contact points

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

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

III.1 2.5 mm Diameter plug connector with three poles

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



NOTE 1 – 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.

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

Figure III.1 – Shape and dimensions of the 2.5 mm diameter plug connector with three poles

III.2 2.5 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.2. To achieve better compatibility with various HPs or headsets, especially for those used in digital mobile terminals, sockets with three contact points should be in line with this requirement. If the 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.2, so that bushing of the socket should not be longer than 3.9 mm.

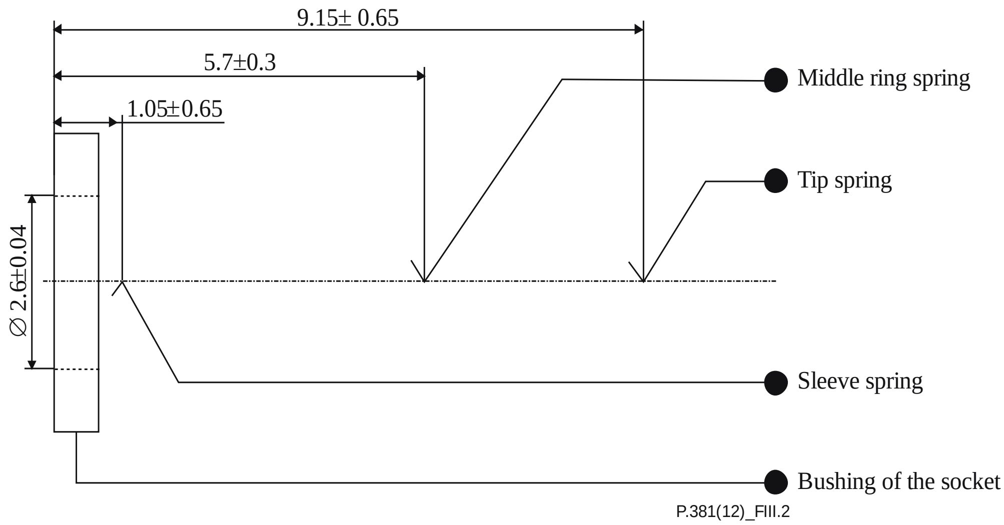
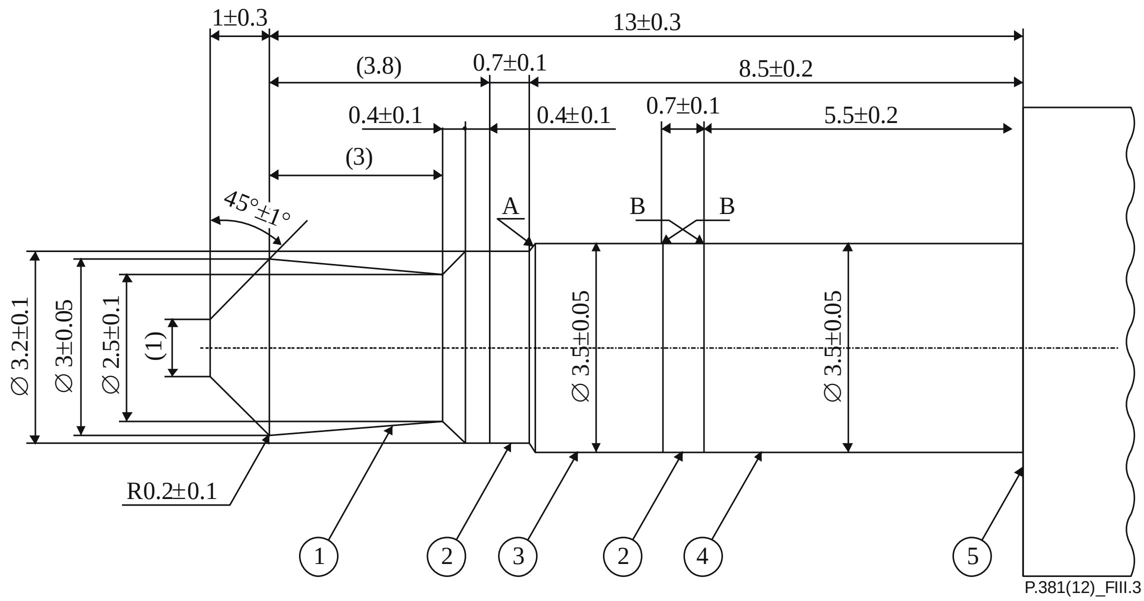


Figure III.2 – Dimensions of the 2.5 mm diameter socket with three contact points and positioning of each contact spring

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 flash.



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 HPs or headsets, the 3.5 mm diameter sockets with three contact points should be in line with this requirement. If the 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.

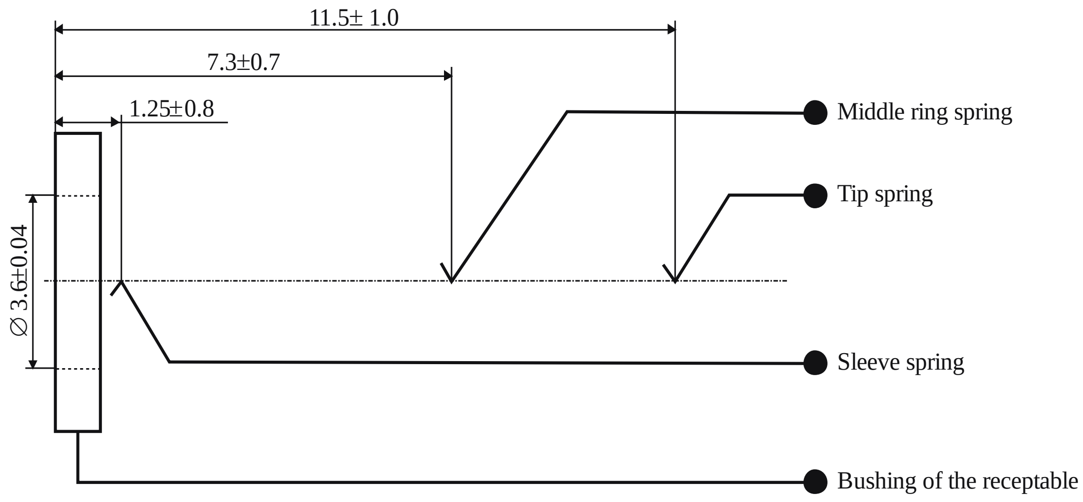


Figure III.4 – Dimensions of the 3.5 mm diameter socket with three contact points and positioning of each contact spring

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 MIC in order that the headset shall not generate any appreciable non-linear distortion as a result of interfering radio frequency (RF) signals. This is covered by audio breakthrough testing (ABT) specifications.

IV.2 Electrostatic discharge

It is recommended that the MIC be protected using an ESD diode. A low value can result in poor electromagnetic interference performance (this can result in ABT failures) and a high value will result in voltage spikes as seen by the MIC itself.

IV.3 Microphone basics – background

The standard electret MICs 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, under 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 kΩ and the current can, for example, be 220-250 µA.

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 MIC 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 kΩ to 6 kΩ. However, in some cases a lower value is desired, for example, if you want to detect multiple passive button values.

Traditionally, 2.2 kΩ has been used in the past 30 years by electret manufacturers. The value comes from a max current of 500 µA from the old JFET design. As JFET technology has improved, manufacturers have increased the DC resistance value. Although conservative, a typical value seen is 2.6 kΩ.

IV.4 V_{cc} 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 V to 3.6 V can be used, but the range 2.7 V to 3.0 V is preferred, for compatibility considerations with all headsets (see Note).

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

NOTE – The preferred range of 2.7 V to 3.0 V for MIC 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 MIC bias will also be representative of performance when using a slightly higher MIC bias in the range 2.7 V to 3.0 V. However, the reverse may not be true, because 2.6 V is near the edge of what is needed for some MIC headsets to operate properly. Therefore, testing with a slightly lower MIC bias voltage may catch interoperability issues with existing sockets not designed for the preferred 2.7 V to 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 MIC current should be in the range 120 µA to 300 µA (i.e., the DC resistance should be such that the resulting MIC 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.10] Recommendation ITU-T P.10/G.100 (2017), *Vocabulary for performance, quality of service and quality of experience.*

[IEC 60268-7] International Standard IEC 60268-7:2010, *Sound system equipment – Part 7: Headphones and earphones.*

[b-3GPP TS 26.131] Technical Specification 3GPP TS 26.131 V17.3.0 (2023), *Universal mobile telecommunications system (UMTS); LTE; 5G; Terminal acoustic characteristics for telephony; Requirements*.

[b-3GPP TS 26.071] Technical Specification 3GPP TS 26.071 V17.0.0 (2022), *Digital cellular telecommunications system (Phase 2+) (GSM); Universal mobile telecommunications system (UMTS); LTE; 5G; Mandatory speech CODEC speech processing functions; AMR speech codec; General description*.

[b-3GPP TS 26.441] Technical Specification 3GPP TS 26.441 V17.3.0 (2023), *Universal mobile telecommunications system (UMTS); LTE; 5G; Codec for enhanced voice services (EVS): General overview*.

[b-ETSI TS 103 106] Technical Specification ETSI TS 103 106 V1.5.1 (2018), *Speech and multimedia transmission quality (STQ); Speech quality performance in the presence of background noise: Background noise transmission for mobile terminals – objective test methods*.

[b-ETSI TS 103 281] Technical Specification ETSI TS 103 281 V1.1.1 (2017), *Speech and multimedia transmission quality (STQ); Speech quality in the presence of background noise: Objective test methods for super-wideband and fullband terminals*.

[b-ETSI EG 202 396-1] ETSI Guide ETSI EG 202 396-1 V1.9.1 (2023), *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*.

SERIES OF ITU-T RECOMMENDATIONS

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

Printed in Switzerland
Geneva, 2023