International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.313

(06/2015)

SERIES P: TERMINALS AND SUBJECTIVE AND OBJECTIVE ASSESSMENT METHODS

Voice terminal characteristics

Transmission characteristics for cordless and mobile digital terminals

Recommendation ITU-T P.313

ITU-T

International Telecommunication Union

ITU-T P-SERIES RECOMMENDATIONS

TERMINALS AND SUBJECTIVE AND OBJECTIVE ASSESSMENT METHODS

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

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

Recommendation ITU-T P.313

Transmission characteristics for cordless and mobile digital terminals

Summary

Recommendation ITU-T P.313 provides audio performance requirements for portable digital cordless and mobile handsets, headsets and speakerphone sets. The requirements apply to narrowband systems (3.1 kHz) regardless of the coding algorithm used in the terminal. Associated test methods are also given.

Requirements are specified for the major electro-acoustic performance parameters affecting audio quality, including sending and receiving levels, frequency responses, noise, sidetone, stability, echo path and delay. The requirements given in this Recommendation should ensure satisfactory service quality in a high percentage of installations under normal conditions.

Major changes over the previous version of this Recommendation (2007) are as follows:

- Specifications for wideband handset, wideband headset and wideband hands-free terminals have been added;
- The use of type 1 and type 3.2 artificial ears are removed from this Recommendation;
- For handset terminals the test of speech quality in the presence of ambient noise has been added;
- For most of the test point receive is now for the diffuse-field and no longer for ear reference point (ERP).
- Requirements for some tests such as frequency response, distortion and linearity have been updated.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T P.313	1999-09-30	12	11.1002/1000/4752
2.0	ITU-T P.313	2004-05-14	12	11.1002/1000/7301
3.0	ITU-T P.313	2007-03-16	12	11.1002/1000/9064
4.0	ITU-T P.313	2015-06-29	12	11.1002/1000/12514

Keywords

Cordless, electro-acoustic measurements, hands-free, mobile, performance specification, speakerphone, terminals, wireless.

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

FOREWORD

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

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

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

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

NOTE

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

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

INTELLECTUAL PROPERTY RIGHTS

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

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

Table of Contents

Page

Recommendation ITU-T P.313

Transmission characteristics for cordless and mobile digital terminals

1 Scope

This Recommendation deals with electro-acoustic performance parameters of portable digital cordless and mobile terminals. The requirements given below should ensure satisfactory voice service in a high percentage of installations under normal conditions, but other factors impacting the performance, such as the radio link, are not included.

This Recommendation could be used as the baseline text of requirements and test methods for mobile or cordless telephony, e.g., when developing new telephony systems. It is not intended to override other acoustic specifications that have been developed for specific telephony systems.

This Recommendation includes specifications for handsets, headsets and speakerphone sets of mobile terminals. These specifications may be applicable to digital cordless handsets, headsets and speakerphone sets. The requirements contained in this Recommendation apply only to narrowband systems (3.1 kHz) and wideband systems (7 kHz) regardless of the coding algorithm used.

Specifications for speakerphone sets are recommended for handheld terminals and desktops.

Specifications for car mounted hands-free terminals will be included in a separate Recommendation.

Requirements are given for handsets, headsets and speakerphone sets in conjunction with a 0 dBr 4-wire reference (see Figure 1) base station having an appropriate air interface, and are specified irrespective of the particular technology and air interface.

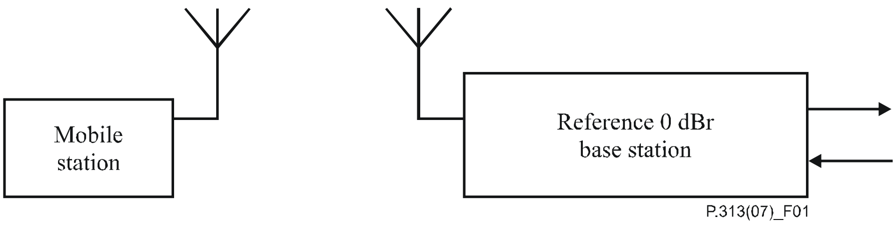

Figure 1 – Reference configuration

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 G.131] Recommendation ITU-T G.131 (2003), *Talker echo and its control*.

[ITU-T G.168] Recommendation ITU-T G.168 (2014), *Digital network echo cancellers.*

[ITU-T O.41] Recommendation ITU-T O.41 (1994), *Psophometer for use on telephone-type circuits*.

[ITU-T P.10] Recommendation ITU-T P.10/G.100 (2006), *Vocabulary for performance and quality of service.*

[ITU-T P.50] Recommendation ITU-T P.50 (1999), *Artificial voices*.

[ITU-T P.51] Recommendation ITU-T P.51 (1996), *Artificial mouth.*

[ITU-T P.57] Recommendation ITU-T P.57 (2011, *Artificial ears*.

[ITU-T P.58] Recommendation ITU-T P.58 (2013), *Head and torso simulator for telephonometry*.

[ITU-T P.64] Recommendation ITU-T P.64 (2007), *Determination of sensitivity/frequency characteristics of local telephone systems.*

[ITU-T P.79] Recommendation ITU-T P.79 (2007), *Calculation of loudness ratings for telephone sets*.

[ITU-T P.310] Recommendation ITU-T P.310 (2009), *Transmission characteristics for narrow-band digital handset and headset telephones.*

[ITU-T P.330] Recommendation ITU-T P.330 (2003), *Speech processing devices for acoustic enhancement.*

[ITU-T P.340] Recommendation ITU-T P.340 (2000), *Transmission characteristics and speech quality parameters of hands-free terminals.*

[ITU-T P.342] Recommendation ITU-T P.342 (2009), *Transmission characteristics for narrow-band digital loudspeaking and hands-free telephony terminals.*

[ITU-T P.360] Recommendation ITU-T P.360 (2006), *Efficiency of devices for preventing the occurrence of excessive acoustic pressure by telephone receivers and assessment of daily noise exposure of telephone users.*

[ITU-T P.380] Recommendation ITU-T P.380 (2003), *Electro-acoustic measurements on headsets.*

[ITU-T P.501] Recommendation ITU-T P.501 (2012), *Test signals for use in telephonometry*.

[ITU-T P.502] Recommendation ITU-T P.502 (2000), *Objective test methods for speech communication systems using complex test signals.*

[ITU-T P.581] Recommendation ITU-T P.581 (2014), *Use of head and torso simulator for hands-free terminal testing.*

[ITU-T P.832] Recommendation ITU-T P.832 (2000), *Subjective performance evaluation of hands-free terminals.*

[IEC 61260] IEC 61260:1995, *Electroacoustics – Octave-band and fractional octave-band filters*.

[ETSI ES 103 106] ETSI TS 103 106 V1.2.1 (2013-03), *Speech & multimedia Transmission Quality (STQ); Speech quality performance in the presence of background noise: Background noise transmission for mobile terminals-objective test methods.*

[ETSI ES 202 396-1] ETSI ES 202 396-1 V1.4.1 (2012-10), *Speech quality performance in the presence of background noise; Part 1: Background noise simulation technique and background noise database*.

NOTE – The latest versions of the annexes to these Recommendations shall apply.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 hands-free reference point (HFRP) [ITU-T P.10]: (see [ITU-T P.340], [ITU-T P.341] and [ITU-T P.342]): A point located on the axis of the artificial mouth, at 50 cm from the outer plane of the lip ring, where the level calibration is made, under free-field conditions. It corresponds to the measurement point 11, as defined in [ITU-T P.51].

3.1.2 hands-free terminal [ITU-T P.10]: A telephone set that does not require the use of hands during the communications session; examples are headset, speakerphone and group-audio terminal.

3.1.3 handset [ITU-T P.10]: A device which includes telephone receiver and transmitter which is typically coupled to the ear by hand.

3.1.4 handset telephone [ITU-T P.10]: A telephone set equipped with a handset.

3.1.5 head and torso simulator (HATS) [ITU-T P.10]: (see [ITU-T P.58]): Manikin extending downward from the top of the head to the waist, designed to simulate the sound pick-up characteristics and the acoustic diffraction produced by a median human adult and to reproduce the acoustic field generated by the human mouth.

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

3.1.7 recommended test position (RTP) [ITU-T P.310]: Corresponds to the position in which the headset should be placed on HATS, e.g., as instructed by the manufacturer. In all cases, the RTP should resemble the RWP on humans.

3.1.8 recommended wearing position (RWP) [ITU-T P.310]: Corresponds to the position in which a headset should be placed on humans according to the intended use (e.g., as instructed by the manufacturer in the user manual, etc.).

3.1.9 speakerphone set [ITU-T P.10]: A *telephone set* using a loudspeaker as a telephone receiver with or without an embedded microphone as a telephone transmitter; it may be used without the handset.

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AEC	Acoustic Echo Controller
Ardt	Attenuation range in receiving direction during double talk
Asdt	Attenuation range in sending direction during double talk
CL	Centre of Lips
CSS	Composite Source Signal
DRP	Eardrum Reference Point
ERP	Ear Reference Point
HATS	Head And Torso Simulator
HFRP	Hands-Free Reference Point

LR	Loudness Rating
MRP	Mouth Reference Point
NR	Noise Reduction
PN	Pseudorandom Noise
POI	Point Of Interconnection
RLR	Receiving Loudness Rating
RTP	Recommended Test Position
RWP	Recommended Wearing Position
SLR	Sending Loudness Rating
SPDA	Speech Processing Devices for Acoustic Enhancement
SS	System Simulator
STMR	Sidetone Masking Rating
TCLw	weighted Terminal Coupling Loss
TCLwdt	weighted Terminal Coupling Loss during double talk
TCLwst	weighted Terminal Coupling Loss during single talk
TELR	Talker Echo Loudness Rating
TELRdt	Talker Echo Loudness Rating during double talk
TRst-r	build-up Time, single talk, receive signal
TRst-s	build-up Time, single talk, send signal
UE	User Equipment
VAD	Voice Activity Detector

5 Conventions

None.

6 Test configuration

The general access to terminals is described in Figure 2. This can be made by using a head and torso simulator (HATS), with appropriate artificial ear and appropriate mountings for handset terminal in a realistic but reproducible way.

HATS is described in [ITU-T P.58]; appropriate artificial ears are described in [ITU-T P.57] (type 3.3 and type 3.4 ears); a proper positioning of handset in realistic conditions can be found in [ITU-T P.64]; the test set-ups for various types of speakerphone sets can be found in [ITU-T P.581].

NOTE – The detailed description of the artificial ears and their applicability can be found in [ITU-T P.57].

The preferred way of testing is the connection of a terminal to the system simulator (SS) with exact defined settings and access points. The test sequences are fed in either electrically using a reference codec, or acoustically.

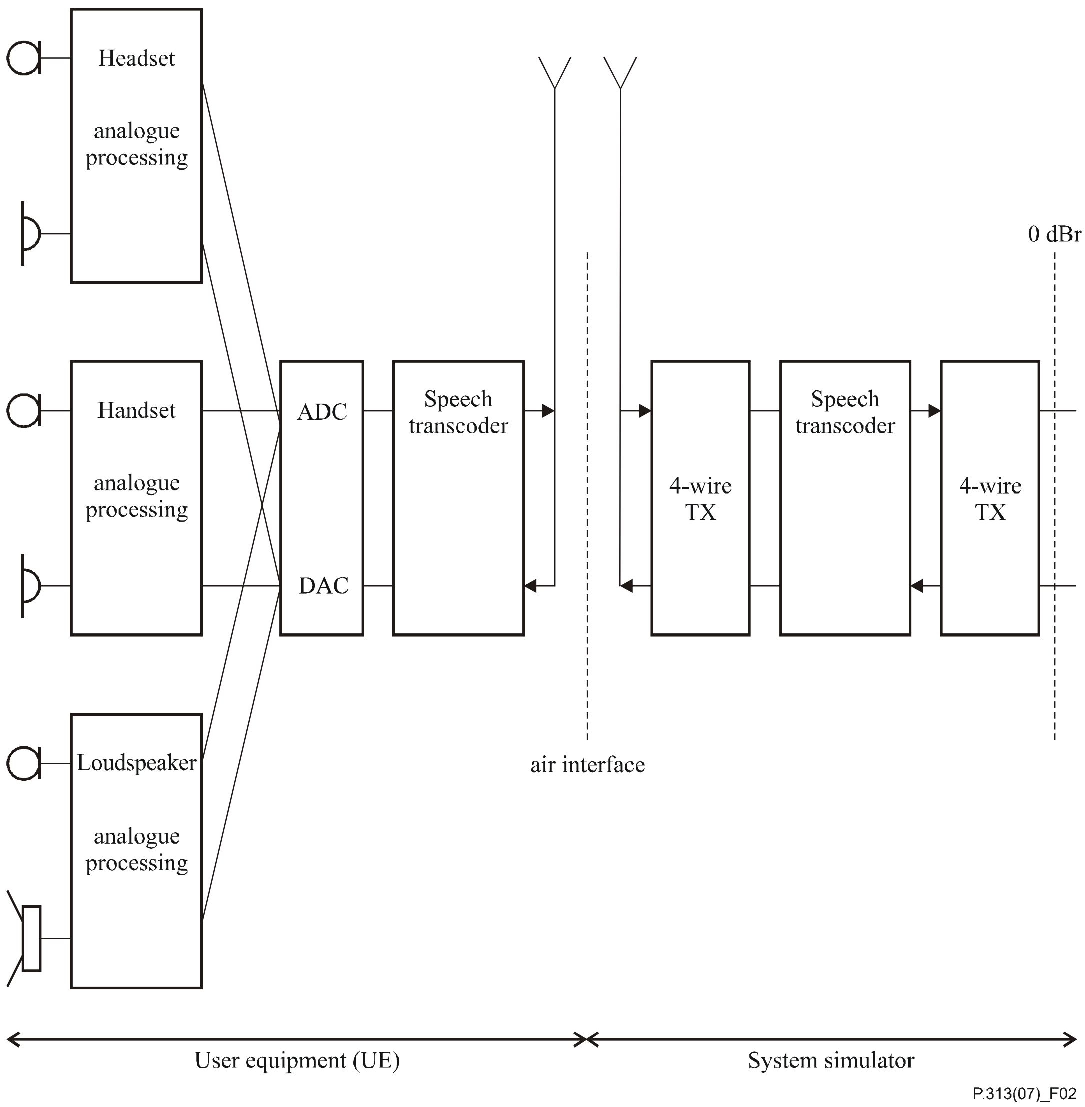

Figure 2 – Interfaces for specification and testing of terminal acoustic characteristics

6.1 Set-up for handset terminal

When using a HATS, the handset is placed according to [ITU-T P.64]. The artificial mouth shall comply with [ITU-T P.58], the artificial ear shall comply with [ITU-T P.57], type 3.3 or type 3.4 ears should be used.

It shall be indicated what application force was used. If not stated otherwise, an application force of 8 ± 2 N shall be used.

The horizontal positioning of a HATS reference plane shall be guaranteed within ± 2º.

The type of artificial ear chosen will be used for all the tests.

6.2 Set-up for headset

The test set-up for headset terminals is defined in [ITU-T P.380].

Some insert earphones might not fit properly in type 3.3 or type 3.4 ear simulators. For such insert type headsets, an [ITU-T P.57] specified type 2 ear simulator may be used in conjunction with a HATS mouth simulator. An ear canal extension representative for typical fitting of the earphone to human ears should be used.

A HATS should be equipped with two artificial ears as specified in [ITU-T P.57]. For binaural headsets and speakerphone sets, two artificial ears are required.

NOTE 1 – As indicated in [ITU-T P.380], the measurement position for headset has to be specified and fully documented along with measurement.

NOTE 2 – Once the position of the headset has been specified, it should be used for all the tests.

6.3 Set-up for speakerphone sets

6.3.1 Desktop operated speakerphone sets

HATS test equipment and set-up for speakerphone sets can be found in [ITU-T P.581].

Measurement set-up using a free-field microphone and a discrete ITU-T P.51 artificial mouth for desktop speakerphone set can be found in [ITU-T P.340].

The signal level for sending characteristics measurement shall be adjusted to –4.7 dBpa at mouth reference point (MRP_ or –28.7 dBpa at hands-free reference point (HFRP) or HATS HFRP.

The signal level for receiving characteristics measurement shall be adjusted to –16 dBm0.

6.3.2 Handheld speakerphone set

If HATS measurement equipment is used, it should be configured to the handheld speakerphone set according to Figure 3. The HATS should be positioned so that the HATS reference point is at a distance d_{HF} from the centre point of the visual display of the mobile station. The distance d_{HF} is specified by the manufacturer. A vertical angle θ_{HF} may be specified by the manufacturer.

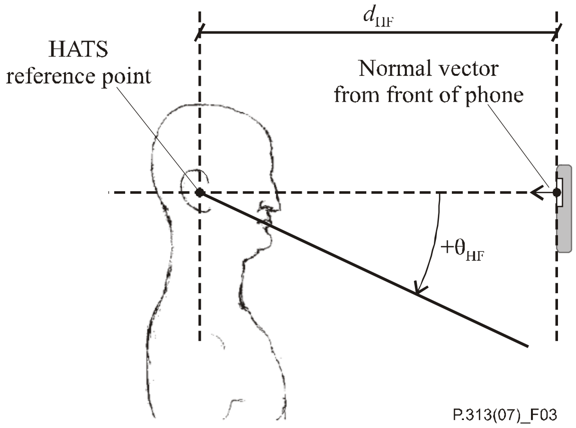

Figure 3 – Configuration of handheld loudspeaker user equipment (UE) relative to the HATS

If a free-field microphone together with a discrete ITU-T P.51 mouth is used, they should be configured to the handheld speakerphone set as per Figure 4 for receiving measurements and Figure 5 for sending measurements. The measurement instrument should be located at a distance d_{HF} from the centre of the visual display of the mobile station. The distance d_{HF} is specified by the manufacturer, and $d_{HFR} = d_{HF}$, $d_{HFS} = d_{HF}-d_{EM}$, where d_{HFR} is the distance for receiving measurement, d_{HFS} is the distance for sending measurement, and d_{EM} is the distance from the ear reference point (ERP) to MRP.

The signal level for sending characteristics measurement shall be adjusted to –4.7 dBPa at the MRP.

The signal level for receiving characteristics measurement shall be adjusted to –16 dBm0.

If the distance d_{HF} is not defined by the manufacturer, the default position as described below is used. The default position is 30 cm for d_{HFS}, which seems a typical value for a handheld phone.

With this value of d_{HFS}, we would obtain $d_{HF} = d_{HFR} = d_{HFS} + d_{EM} = 42$ cm.

$$d_{EM} = d_{EEP\text{-}CL} \times \cos(\theta_{HF})$$

with $d_{EEP\text{-}CL}$ = distance between EEP and centre of lips = 13 cm nominal

$$\theta_{HF} = 24° \text{ nominal}$$

Resulting in $d_{EM} = 41.9$ cm.

The calibration for this default position is as follows:

– calibration and equalization at the MRP: –4.7 dBPa;

– calibration at 30 cm: –24.3 dBPa (attenuation is supposed to be 24 dB from the MRP to HFRP located at 50 cm; so the attenuation at 30 cm is 4.4 dB less than for 50 cm).

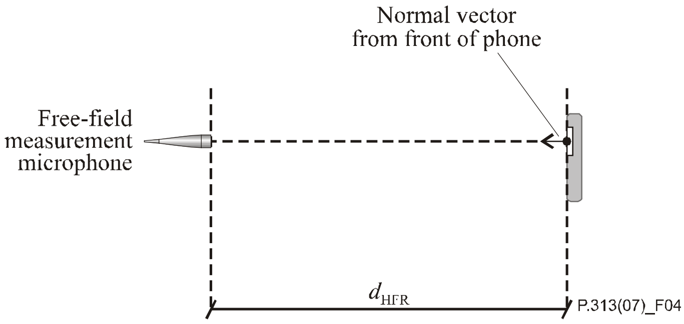

Figure 4 – Configuration of handheld speakerphone set, free-field microphone for receiving measurements

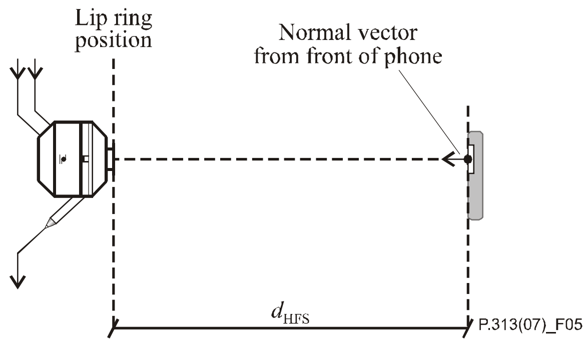

Figure 5 – Configuration of handheld speakerphone set, discrete ITU-T P.51 artificial mouth for sending measurements

7 Test conditions

7.1 Test signal

In general, the test signals used should be the real speech signals defined in clause 7.3 of [ITU-T P.501]. The use of sine wave signals and other signals such as composite source signals (CSSs) are only applicable when specifically stated. The type of test signal used shall be stated in the test report.

For testing the narrowband transmission characteristics of the terminal, the test signal used shall be band limited between 100 Hz and 4 kHz with a band-pass filter providing a minimum of 24 dB/octave filter roll-off when feeding into the receiving direction.

For testing the wideband transmission characteristics of the terminal, the test signal used shall be band limited between 100 Hz and 8 kHz with a band-pass filter providing a minimum of 24 dB/octave filter roll-off when feeding into the receiving direction.

If not stated otherwise, the signal level for sending characteristics measurement shall be adjusted to –4.7 dBPa at the MRP. The signal level for receiving characteristics measurement shall be adjusted to –16 dBm0. The test signal levels are referred to the average level of the (band limited in receiving direction) test signal, averaged over the complete test sequence, unless specified otherwise. For real speech, the test signal levels are referred to the ITU-T P.56 active speech level of the (band limited in receiving direction) test signal, calculated over the complete test sequence.

The spectrum of the acoustic signal produced by the artificial mouth is calibrated under free-field conditions at the MRP.

When using the real speech signals defined in [ITU-T P.501], care should be taken not to overload the system under test due to the high crest factor of the real speech signals.

7.2 Test environment

It is the responsibility of the test laboratory to ensure that the measurements done in the test room give identical results to those obtained in a free-field environment (e.g., an anechoic room).

7.2.1 Handset and headset terminal

The environmental conditions for testing handset and headset terminals are specified as follows:

For handset and headset measurements, the test room shall be practically free-field down to a lowest frequency of 275 Hz; the handset or the headset coupled with the HATS shall lie totally within this free-field volume. This is met if deviations from the ideal free-field conditions are less than ±1 dB.

The ambient noise level shall be less than –64 dBPa(A) for all the tests, except for tests requesting simulated ambient noise.

7.2.2 Speakerphone sets

When testing the acoustic performance of a speakerphone set, care must be taken that, for example, noise levels are sufficiently low in order not to interfere with the measurements.

The broadband noise level shall not exceed –70 dBPa(A) for all tests, with the exception of those tests requesting simulated ambient noise.

The octave band noise level shall not exceed the values specified in Table 1.

Table 1 – Noise level

Centre frequency (Hz)	Octave band pressure level (dBPa)
63	–45
125	–60
250	–65
500	–65
1 k	–65
2 k	–65
4 k	–65
8 k	–65

7.3 Test set-up for simulated ambient noise in laboratory

The set-up for simulating realistic ambient noises and the positioning of a HATS in a lab-type environment is described in [ETSI ES 202 396-1].

[ETSI ES 202 396-1] contains a description of the recording arrangement for realistic ambient noises, a description of the set-up for a loudspeaker arrangement suitable to simulate an ambient noise field in a lab-type environment, and a database of realistic ambient noises, part of which is used for testing the terminal performance in a variety of conditions.

The equalization and calibration procedures for the test set-up are given in detail in [ETSI ES 202 396-1].

8 Narrowband handset technical requirements

Unless otherwise specified, measurements defined in clause 8 shall be carried at the setting of the volume control where the receiving loudness rating (RLR) is as close as possible to nominal value.

8.1 Sending characteristics

8.1.1 Sending loudness rating (SLR)

In view of the following considerations:

– wireless terminals provide connectivity to the existing wireline-based international telecommunication networks;

– digital wireless access networks should provide the same signal levels as the digital access wireline networks.

The following loudness rating (LR) value is recommended:

– a nominal value of SLR = 8 ± 3 dB.

NOTE – Several reports have indicated that sending speech levels in mobile networks are 5 dB higher on average than speech levels from fixed networks. The most likely cause is that mobile terminals are frequently used in noisy environments where users tend to talk louder. Manufacturers of mobile handsets should be aware that this may lead to saturation, clipping and distortion of the signal and reduced speech quality. Changing SLR at this time is not recommended due to concerns of reduced speech quality when used at normal talker levels. However, the manufacturers of devices intended for use in noisy conditions can consider adopting a higher SLR.

8.1.1.1 Measurement method

The real speech signal as defined in [ITU-T P.501] shall be used. The test signal used shall be specified in the test report. The test signal level shall be –4.7 dBPa, measured at the MRP. The test signal level is averaged over the complete test signal sequence.

The handset shall be mounted according to [ITU-T P.64].

The sending sensitivity shall be calculated from each band of the 14 frequencies given in bands 4 to 17 in Table 1 of [ITU-T P.79]. For the calculation, the averaged measured level at the electrical reference point for each frequency band is referred to the averaged test signal level measured in each frequency band at the MRP.

The sensitivity is expressed in terms of dBV/Pa and the SLR shall be calculated according to formula (A-23b) in [ITU-T P.79], over bands 4 to 17, using m = 0.175 and the sending weighting factors from Table 1 of [ITU-T P.79].

8.1.2 Sending frequency responses

In view of the following considerations:

– the compatibility with wireline digital telephones and the mixed analogue/digital telephone network;

– the compatibility with most existing wireless systems;

– the aim to achieve the best possible overall quality with the cordless and mobile terminals, sending nominal sensitivity/frequency response within the limits given in Table 2 is recommended.

Table 2 – Sending

Frequency (Hz)	Upper limit (dB)	Lower limit (dB)
100	–12	–∞
200	0	–∞
300	0	–12
1 000	0	–6
2 000	4	–6
3 000	4	–6
3 400	4	–9
4 000	0	–

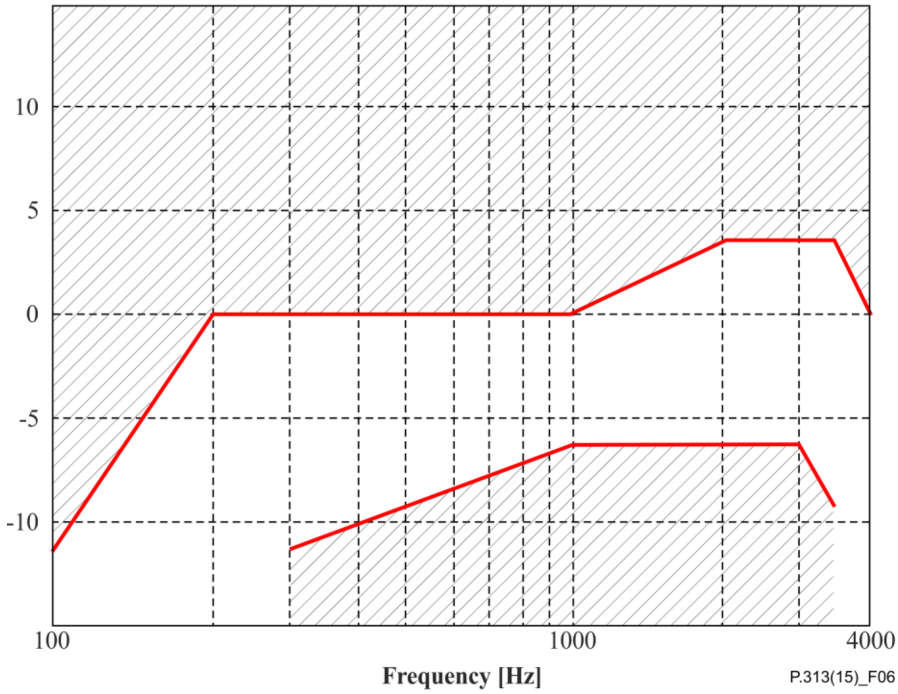

Figure 6 – Sending mask

NOTE –The basis for the target frequency responses in send and receive is the orthotelephonic reference response measured between 2 subjects 1 m apart under free-field conditions and assumes an ideal receive characteristic. Under these conditions, the overall frequency response shows a rising slope. The present Recommendation now uses the diffuse-field as a reference point for receive and no longer uses the ERP. With the concept of diffuse-field based receive measurements, a rising slope for the overall frequency response is achieved by a flat target frequency response in send, and a flat diffuse-field based receive frequency response.

8.1.2.1 Measurement method

The sending frequency response is measured according to [ITU-T P.64], using the measurement set-up shown in Figure 7.

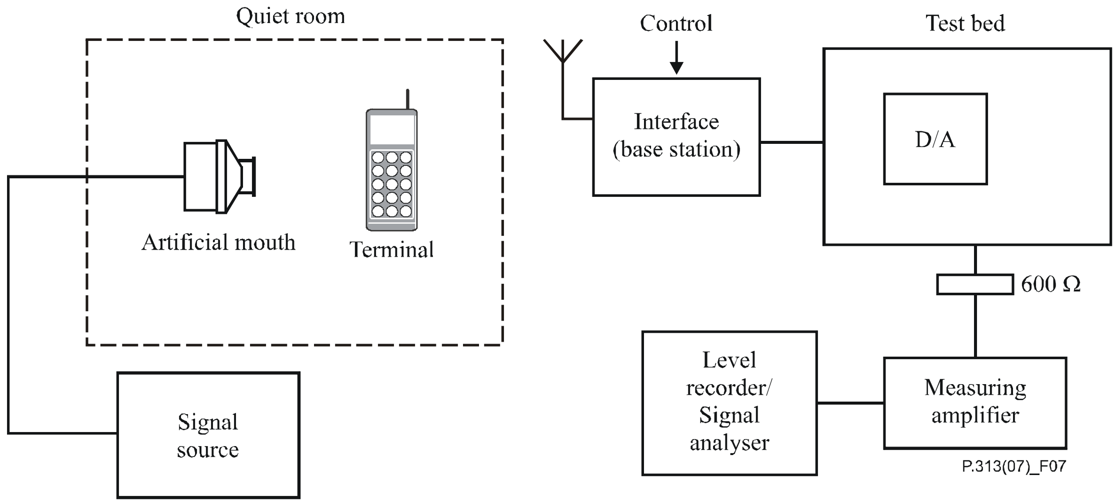

Figure 7 – Sending frequency response measurement method

The real speech signal as defined in [ITU-T P.501] and a spectrum analyser should be used. The test signal used shall be specified in the test report. The test signal level shall be –4.7 dBPa, measured at the MRP. The test signal level is averaged over the complete test signal sequence.

The handset is mounted according to [ITU-T P.64].

Measurements shall be made at 1/12-octave intervals as given by the [IEC 61260] for frequency bands from 100 Hz to 4 kHz inclusive. For the calculation, the averaged measured level at the electrical reference point for each frequency band is referred to the averaged test signal level measured in each frequency band at the MRP.

The sensitivity is expressed in terms of dBV/Pa.

8.1.3 Idle channel noise

The following limit is recommended:

– sending noise level maximum –64 dBm0p.

NOTE – This figure applies to the total noise level with psophometric weighting. It is recommended that the level of single frequency disturbances should be ≤ –74 dBm0p in the frequency range from 300 Hz to 3.4 kHz.

8.1.3.1 Measurement method

The handset is mounted according to [ITU-T P.64] in a quiet environment, as defined in clause 7.2, the sending noise level at the SS audio output is measured with apparatus including psophometric weighting described in [ITU-T O.41].

8.1.4 Distortion

The sending distortion includes non-linear distortion and quantizing distortion. The sending part shall meet the following distortion requirements:

Limits for intermediate levels are found by drawing straight lines between the breaking points in Table 3 on a linear (dB signal level) – linear (dB ratio) scale.

Table 3 – Limits for signal-to-total distortion ratio

Sending level (dBPa at the MRP)	Sending ratio (dB)
5	30
0	35
–4.7	35
–10	33
–15	30
–20	27

The sending distortion shall be measured between the MRP and the SS audio output (output of the reference speech decoder of the SS). The ratio of signal-to-total distortion power measured shall be above the limits given in Table 3.

8.1.4.1 Measurement method

The handset is mounted according to [ITU-T P.64] in a quiet environment as defined in clause 7.2.

The signal used is an activation signal followed by a sine wave signal with a frequency of 1 020 Hz which is the actual test signal. The signal is applied at the MRP. Appropriate signals for activation and signal combinations can be found in [ITU-T P.501].

The sine wave signal level shall be calibrated to the following RMS levels at the MRP: 5, 0, −4.7, −10, −15, −20 dBPa. The test signals have to be applied in this sequence, i.e., from high levels down to low levels.

The total distortion power of the signal output of the SS shall be measured with the psophometric noise weighting described in [ITU-T O.41].

8.1.5 Linearity

NOTE – Linearity tests as described below assume linear and time invariant behaviour of the UE. If this cannot be assured, alternative test methods based on more complex test signals may be needed for these tests. The principles are described in [ITU-T P.501] and [ITU-T P.502].

The following characteristic is recommended:

– the linearity relative to the gain for the input level of –4.7dBPa should remain within the limits given in Table 4.

Table 4 – Linearity

Sending gain relative to −4.7dBPa (dB)	Upper limit (dB)	Lower limit (dB)
+4	1	–2
–10	1	–2

The limits for intermediate levels can be found by drawing straight lines between the break points in the table on a linear (dB signal level) – linear (dB gain) scale.

8.1.5.1 Measurement method

The handset is mounted according to [ITU-T P.64] using a suitable artificial ear as recommended in [ITU-T P.57].

A sine wave signal with a frequency in the range of 1004 Hz to 1025 Hz should be applied at the MRP. The level of the signal at the MRP is –4.7dBPa.

The test signal should be applied at the following gains relative to the input level of –4.7dBPa: –10, 4 dB.

The linearity relative to the input level of –4.7dBPa should be measured.

Appropriate signals for activation and signal combinations can be found in [ITU-T P.501].

NOTE 1 – In general, care must be taken in the case of time variant and/or non-linear terminals. In such cases, a sine wave may not be appropriate as the test signal; a speech-like test signal should be chosen as described in [ITU-T P.501] and [ITU-T P.50]. If the sine wave techniques cannot be used (which is likely if the terminal has a noise reduction (NR) device), another appropriate technique should be applied. For example, an artificial speech generator (e.g., as specified in [ITU-T P.50] and [ITU-T P.501]) and a spectrum analyser can be used. Additional test methods may be found in [ITU-T P.502]. The test signal used shall be specified in the test report.

NOTE 2 – Selective measurement may be used to avoid the effects of ambient noise.

8.1.6 Out-of-band signals

With any sine wave signal above 4.6 kHz and up to 8 kHz applied at the MRP at a level of –4.7dBPa, the level of any image frequency produced at SS audio output shall be below a reference level obtained at 1 kHz (–4.7 dBPa at MRP) by at least the amount (in dB) specified in Table 5.

Table 5 – Discrimination levels

Applied sine wave frequency	Limit (minimum)
4.6 kHz	30 dB
8 kHz	40 dB

The limit at intermediate frequencies lies on a straight line drawn between the given values on a log (frequency) – linear (dB) scale.

8.1.6.1 Measurement method

The handset is mounted according to [ITU-T P.64].

The signal used is an activation signal followed by a sine wave with the frequencies of 4.65 kHz, 5 kHz, 6 kHz, 6.5 kHz, 7 kHz and 7.5 kHz at the level of –4.7 dBPa which are the actual test signals. The test signal is applied at the MRP. Appropriate activation signals and signal combinations can be found in [ITU-T P.501].

For input signals at frequencies of 4.65 kHz, 5 kHz, 6 kHz, 6.5 kHz, 7 kHz and 7.5 kHz at the level of –4.7 dBPa, the level of any image frequencies at the SS audio output shall be measured.

8.2 Receiving characteristics

An application force of 8 ± 2 N should be used when measuring with type 3.3 or type 3.4 artificial ears. This force is considered to be a realistic average value.

8.2.1 RLR

In view of the following considerations:

– wireless terminals provide connectivity to the existing wireline based international telecommunication networks;

– digital wireless terminals should be compatible with the digital access wireline networks.

The following LR value is recommended:

– a nominal value of RLR = 2 ± 3 dB.

8.2.1.1 Measurement method

The real speech signal as defined in [ITU-T P.501] shall be used. The test signal used shall be specified in the test report. The test signal level shall be −16 dBm0 measured at the digital reference point or the equivalent analogue point. The test signal level is averaged over the complete test signal sequence.

The handset should be mounted according to [ITU-T P.64].

The receiving sensitivity shall be calculated from each band of the 14 frequencies given in bands 4 to 17 in Table 1 of [ITU-T P.79]. For the calculation, the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dBPa/V and the RLR shall be calculated according to formula (A-23c) of [ITU-T P.79], over bands 4 to 17, using m = 0.175 and the receiving weighting factors from Table 1 of [ITU-T P.79].

Eardrum reference point (DRP)-ERP correction is used. No leakage correction shall be applied.

8.2.2 Volume control

In view of the following considerations of mobile handsets:

– the mobile terminals are used often in noisy environments;

– the need to provide service for people with hearing impairments. For these applications, [ITU-T P.370] applies.

Where a user-controlled receiving volume control is provided, the RLR shall meet the nominal value for at least one setting of the control. When the control is set to maximum, the RLR shall not be ≤ (equal or louder than) −13 dB.

With the volume control set to the minimum position the RLR shall not be ≥ (equal or quieter than) 18 dB.

8.2.3 Receiving frequency responses

In view of the following considerations:

– the compatibility with wireline digital telephones and the mixed analogue/digital telephone network;

– the compatibility with most existing wireless systems;

– the aim to achieve the best possible overall quality with the cordless and mobile terminals.

The receiving sensitivity/frequency response within the limits given in Table 6 is recommended.

Table 6 – Receiving sensitivity

Frequency (Hz)	Upper limit (dB)	Lower limit (dB)
100	6	–∞
200	6	–∞
300	6	–6
1 000	6	–6
3 400	6	–6
4 000	6	–∞

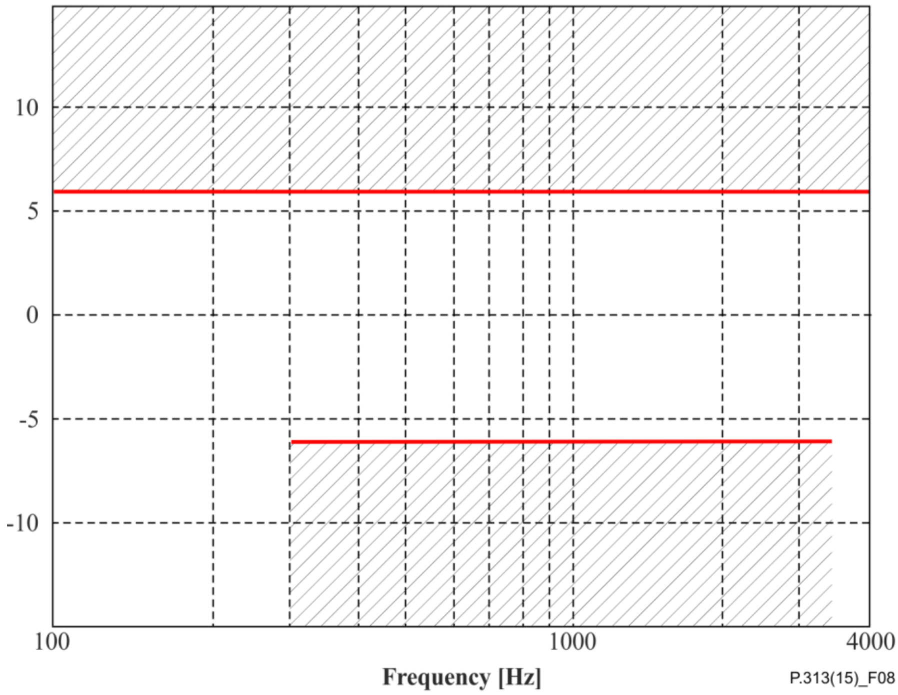

Figure 8 – Receiving mask

NOTE – The following rules may be additionally applied:

In general, the frequency response should not introduce a strong rolloff at lower frequencies, regardless of which coupler is used for measurements. Signal levels at frequencies down to 300 Hz should not be attenuated more than 5 dB as compared to the level measured at 1 kHz. Too much emphasis at high frequencies should also be avoided. Compared to the level measured at 1 kHz, the emphasis introduced up to 3.4 kHz should not be more than 5 dB.

8.2.3.1 Measurement method

The handset should be mounted according to [ITU-T P.64], using a suitable artificial ear as recommended in [ITU-T P.57].

The receiving frequency response is measured using the measurement set-up shown in Figure 9. Terminals with adjustable receive levels shall be adjusted so that their RLR is as close as possible to the nominal value for this test.

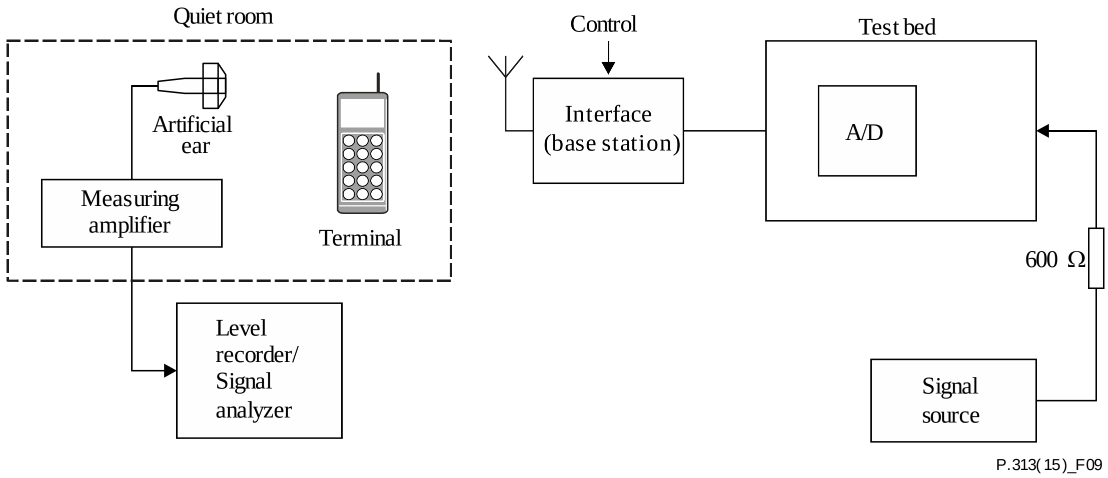

Figure 9 – Receive frequency response measurement method

The real speech signal (as specified in [ITU-T P.501]) shall be used. The test signal used shall be specified in the test report. The test signal level shall be –16.0 dBm0, measured at the digital reference point or the equivalent analogue point. The test signal level is averaged over the complete test signal sequence.

Measurements shall be made at 1/12-octave intervals as given by [IEC 61260] for frequency bands from 100 Hz to 4 kHz inclusive. For the calculation, the averaged level at each frequency band measured at the HATS DRP with diffused-field correction is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dBPa/V. Information about correction factors is available in [ITU-T P.57].

8.2.4 Idle channel noise

The maximum (acoustic) noise level at the handset terminal when no signal is transmitted to the input of the SS shall be as follows:

– if no user-controlled receiving volume control is provided, or, if it is provided, at the setting of the user-controlled receiving volume control at which the RLR is equal to the nominal value, the noise measured at the DRP with diffuse-field correction contributed by the receiving equipment alone shall be ≤ −57 dBPa(A);

– where a volume control is provided, the measured noise shall also not exceed –54 dBPa(A) at the maximum setting of the volume control.

For the nominal volume control setting, the level of single frequency disturbances should be ≤ −60 dBPa(A) in the frequency range from 100 Hz to 10 kHz. As a performance objective it is recommended that the level should be ≤ −64 dBPa(A).

8.2.4.1 Measurement method

With the handset mounted according to [ITU-T P.64], using a suitable artificial ear as recommended in [ITU-T P.57], placed in a quiet environment, as defined in clause 7.2, the receiving noise level at the DRP with diffuse-field correction is measured with A-weighting.

8.2.5 Distortion

The receiving distortion includes non-linear distortion and quantizing distortion.

The receiving part between the SS audio input (input of the reference speech encoder of the SS) and DRP with diffuse-field correction, shall meet the requirements at the nominal setting of the volume control.

The ratio of signal-to-total distortion power measured shall be above the limits given in Table 7. For sound pressures exceeding +10 dBPa at the DRP with diffuse-field correction, there is no distortion requirement.

Table 7 – Limits for signal-to-total distortion ratio

Frequency (Hz)	Receiving level at the digital interface (dBm0)	Receiving ratio at nominal volume setting (dB)	Receiving ratio at maximum volume setting (dB)
315	–16	20	–
408	–16	28	–
510	–16	28	–
816	–16	28	–
1 020	0	25.5	–
	–3	31.2	–
	–10	33.5	–
	–16	33.5	–
	–20	33	–
	–30	30.5	–
	–40 (Note)	22.5 (Note)	–
	–45 (Note)	17.5 (Note)	–
NOTE – For levels –40 and –45 dBm0 a lower signal-to-total distortion ratio may not be possible, and hence would not be regarded as a failing result. However, the obtained results would be reported.			

Limits for intermediate levels are found by drawing straight lines between the breaking points in the table on a linear (dB signal level) – linear (dB ratio) scale.

8.2.5.1 Measurement method

The handset should be mounted according to [ITU-T P.64], using a suitable artificial ear as recommended in [ITU-T P.57].

The signal used is an activation signal followed by a sine wave signal with a frequency specified in clause 8.2.5. The signal level shall be −16 dBm0, except for the sine wave signal with a frequency 1 020 Hz that shall be applied at the signal input of the SS at the following levels: 0, –3, –10, –16, –20, –30, –40, –45 dBm0. The test signals have to be applied in this sequence, i.e., from high levels down to low levels. Appropriate signals for activation and signal combinations can be found in [ITU-T P.501].

The total distortion power shall be measured at the DRP with diffuse-field correction and calculated with the psophometric noise weighting as specified in [ITU-T O.41].

8.2.6 Linearity

If the system is intended to operate in a linear mode, the following is recommended:

– the linearity relative to the gain at an input level of −16 dBm0 should be within the limits given in Table 8.

Table 8 – Variation of gain with input level, receiving

Receiving gain at the digital interface (dB)	Upper limit (dB)	Lower limit (dB)
+4 –10	1 1	–2 –2

The limits for intermediate levels can be found by drawing straight lines between the break points in the table on a linear (dB signal level) – linear (dB gain) scale.

8.2.6.1 Measurement method

The handset should be mounted according to [ITU-T P.64] using a suitable artificial ear as recommended in [ITU-T P.57].

A digitally simulated sine wave signal with a frequency in the range of 1004 Hz to 1025 Hz should be applied at the digital interface at the following gains relative to the input level of –16 dBm0: –10, 4 dB.

NOTE 1 – In general, care must be taken in the case of time variant and/or non-linear terminals. In such cases, a sine wave may not be the appropriate test signal; a more speech-like test signal should be chosen as described in [ITU-T P.501] and [ITU-T P.50]. The test signal used shall be specified in the test report.

The linearity relative to the gain at an input level of –16 dBm0 should be measured using the artificial ear.

NOTE 2 – Selective measurement may be used to avoid the effects of ambient noise.

8.2.7 Spurious out-of-band signals

The level of out-of-band signals at the DRP shall meet the following requirements:

With a digitally-simulated sine wave signal in the frequency range of 300 Hz-3.4 kHz and at a level of –5 dBm0 applied at the digital interface, the level of spurious out-of-band image signals in the frequency range of 4.6-8 kHz measured selectively at the DRP shall be at least lower than the in-band acoustic level produced by a digital signal at 1 kHz set at the level specified in Table 9.

Table 9 – Discrimination levels

Image signal frequency	Equivalent input signal level
4.6 kHz	–35 dBm0
8 kHz	–45 dBm0

The limit at intermediate frequencies lies on a straight line drawn between the given values on a log (frequency) – linear (dB) scale.

8.2.7.1 Measurement method

The handset should be mounted according to [ITU-T P.64], using a suitable artificial ear as recommended in [ITU-T P.57].

The signal used is an activation signal followed by a sine wave signal at the frequencies 500, 1 000, 2 000 and 3 150 Hz, which is the actual test signal. The signal is applied at the input of the SS at the level of –5 dBm. Appropriate signals for activation and signal combinations can be found in [ITU-T P.501]. For test signals at frequencies 500, 1 000, 2 000 and 3 150 Hz, the level of spurious out-of-band image signals at frequencies of up to 8 kHz shall be measured selectively at the DRP.

8.3 Sidetone characteristics

8.3.1 Sidetone masking rating (STMR)

In view of the following considerations:

– the optimum STMR for conditions free from echo;

– the difficulties of high ambient noise conditions.

The following is recommended:

– the value of the STMR shall be ≥ 15 dB and should be ≤ 23 dB for the nominal setting of the volume control;

– for all other positions of the volume control, the STMR shall be ≥ 10 dB.

In case the STMR is below the lower limit, and also when the electrical sidetone path has been disabled, the result shall not be regarded as a failure.

NOTE 1 – The STMR calculation algorithm being used was developed for quantifying the audibility of the electrical sidetone path using a sealed coupler. The air-conducted path was not intended to be included in the test set-up. A lower STMR limit was specified to avoid annoying effects (e.g., howling, increase of ambient noise level in the ear) of an excessive electrical sidetone. In HATS-based measurements, the air-conducted path cannot be avoided in the test set-up. With some terminal form factors, especially large size terminals, the air-conducted path can be substantial resulting in low STMR figures, and also when there are no annoying effects from any excessive electrical sidetone. See [ITU-T P.76] for definitions of sidetone paths.

8.3.1.1 Measurement methods

A real speech test signal in [ITU-T P.501] with the level of –4.7 dBPa shall be applied at the MRP. For each frequency given in Table 3 of [ITU-T P.79], bands 1 to 20, the sound pressure (at the ERP) shall be measured.

The handset should be mounted according to [ITU-T P.64], using a suitable artificial ear as recommended in [ITU-T P.57].

The test set-up shown in Figure 10 is used to measure the sidetone frequency response. The sidetone path loss L_{meST} and the STMR shall be calculated according to [ITU-T P.79], using equation 4-1 (m = 0.225) and the weighting factors in Table 3 of [ITU-T P.79].

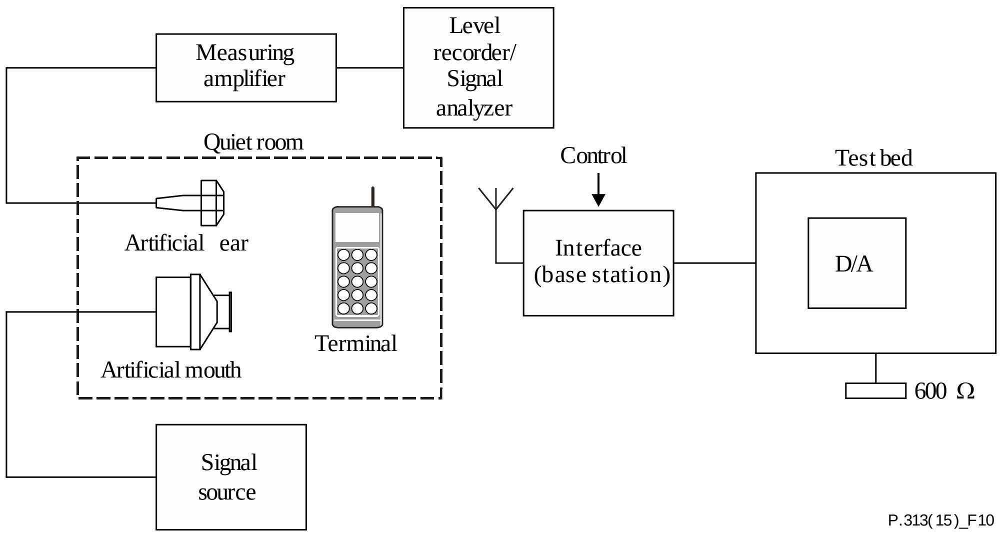

Figure 10 – Sidetone measurement method

8.4 Noise contrast and comfort noise

In some circumstances, such as application of voice-operated devices, the continuous background noise present, regardless of whether the users are talking or not, may be interrupted. This switching on and off is annoying to the users and may in fact degrade speech intelligibility. To reduce this effect, noise contrast should be minimized by increasing signal-to-noise ratio.

Comfort noise may be injected during silent periods to reduce the impairments created by the noise contrast. This may create undesirable performance degradation by itself if not done properly, due to the level or spectrum content differences between the injected and the transmitted noise. Efforts should be made to match the characteristics of the injected comfort noise to the transmitted noise to reduce any perceptible contrast between them.

8.4.1 Measurement method

For further study.

8.5 Weighted terminal coupling loss (TCLw)

In view of the following considerations:

– the aim to achieve as high an acoustic coupling loss as possible to minimize degradation caused by echo;

– that the far-end talker echo should be controlled under all volume control settings, and for the range of transducer sensitivities as long as the handset is properly used;

– the far-end terminal may be connected via a mobile or an IP network which introduces long talker echo path delay;

– what is practically obtainable in real use where the user himself chooses the way to hold his handset.

The following limit is recommended:

– in order to meet the ITU-T G.131 talker echo objective requirements, for handsets fitted with a volume control, the TCLw shall be greater than 46 dB for any setting of the volume control. The TCLw should be greater than 55 dB at the nominal setting of the volume control;

– it is recommended that the volume control should be set back to nominal after each call unless TCLw $\geq$ 55 dB can also be maintained with the maximum volume setting. Depending on the UE idle channel noise in the sending direction, it may not always be possible to measure an echo loss greater than 55 dB.

NOTE – In consideration of the increasing delays introduced by modern networks, a higher TCLw value than that specified here may be necessary for proper operation with these networks.

8.5.1 Measurement method

The handset should be mounted according to [ITU-T P.64]; noise and reflections in the test space must not influence the measurement. The test should be performed in the environment as defined in clause 7.2.

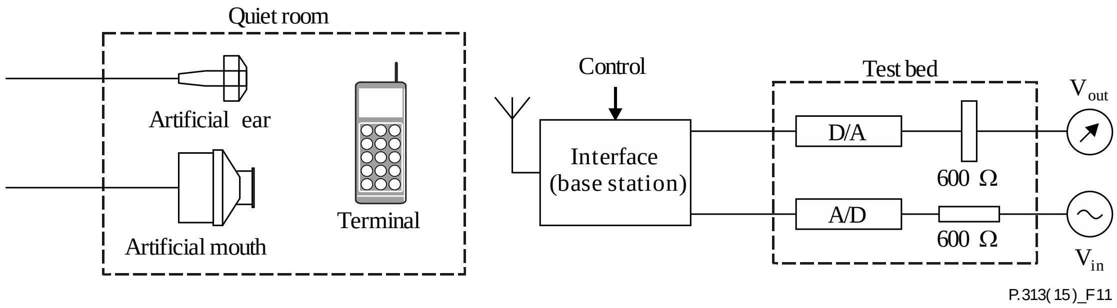

Figure 11 – Terminal coupling loss measurement method

The attenuation from digital input to digital output is measured at one-twelfth octave frequencies as given by [IEC 61260] for frequency bands from 300 to 3 350 Hz, using the measurement arrangement shown in Figure 11.

The test signal used is the compressed real speech signal described in clause 7.3.3 of [ITU-T P.501]. The test signal level shall be −10 dBm0. Terminals with adjustable receive levels shall be tested at the nominal setting. For the nominal setting, adjust the level so that the RLR is as close as possible to the nominal RLR value.

The TCLw is calculated according to clause B.4 of [ITU-T G.122] (trapezoidal rule). For the calculation, the averaged measured echo level at each frequency band is referred to the averaged test signal level measured in each frequency band. The first 17.0 s of the test signal (6 sentences) are discarded from the analysis to allow for convergence of the acoustic echo canceller. The analysis is performed over the remaining length of the test sequence (last 6 sentences).

NOTE – There might be problems measuring 46 dB TCL in the case where sophisticated coding with limited dynamic range is used. In such cases, typically speech or speech-like test signals need to be used that themselves have crest factors in the range of 15 dB which reduces the measurement dynamic by the same amount. In such cases, the signal measured in the sending direction should be evaluated more carefully in order to find whether an echo signal is present, or whether this signal is completely masked by the noise signal introduced by the codec. If the signal measured in the sending direction is completely masked by the noise, the requirement can be considered to be fulfilled. If this is not the case, more sophisticated measurement procedures such as time averaging (in order to improve the signal-to-noise ratio) need to be applied in order to achieve reliable measurement results.

8.6 Stability loss

The minimum stability loss at any volume control setting should be at least 6 dB.

8.6.1 Measurement method

Before the actual test a training sequence consisting of the real speech signal described in [ITU-T P.501] is applied. The training sequence level shall be −16 dBm0 in order to not overload the codec.

The test signal is a PN-sequence complying with [ITU-T P.501] with a length of 4 096 points (for a 48 kHz sampling rate system) and a crest factor of 6 dB instead of 11 dB. The PN-sequence is generated as described in [ITU-T P.501] with W(k) constant within the frequency range 200-4 000 Hz and zero outside this range. The duration of the test signal is 250 ms with an input signal of −3 dBm0.

With the handset and transmission circuit fully active, measure the attenuation from the digital input to the digital output using the following method.

Place the handset in the reference corner, as shown in Figure 12, with the earcap and mouthpiece facing a hard, smooth surface with at least 400 mm free space in all direction.

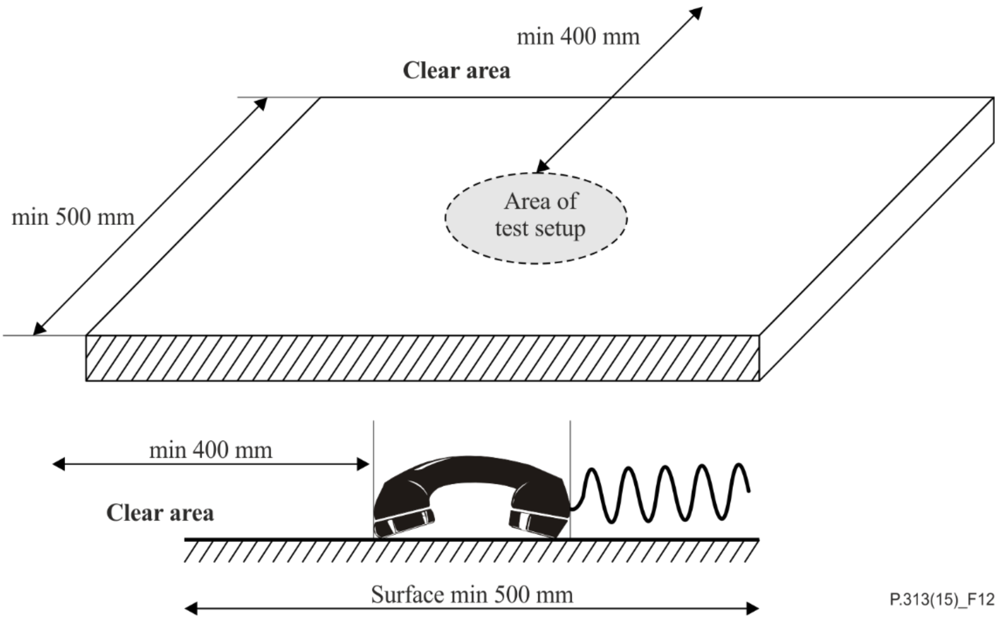

Figure 12 – Reference corner

8.7 Delay

In view of the following considerations:

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

– the amount of delay introduced by wireless systems depends on specific technology and may be inherent to the adopted coding technique.

The following is recommended:

– delay added by the terminal equipment should be minimized in accordance with the guidelines provided in [ITU-T G.114] even with the use of echo control;

– the terminal specific implementation dependent delay， including both the delay in sending direction and the delay in receiving direction, should be less than 70 ms.

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

8.7.1 Measurement method

The delay in the sending direction is obtained by measuring the delay between MRP and the electrical access point of the test equipment and subtracting the delays introduced by the test equipment from the measured value. The delay in the receiving direction is obtained by measuring the delay between the electrical access point of the test equipment and the DRP and subtracting the delays introduced by the test equipment from the measured value.

For the measurements, a CSS according to [ITU-T P.501] is used. The pseudorandom noise (PN) part of the CSS has to be longer than the maximum expected delay. It is recommended to use a PN sequence of 16 k samples (with 48 kHz sampling rate or equivalent).

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

The delay in the sending direction is determined by the cross-correlation analysis between the measured signal at the electrical access point and the original signal at MRP. The delay in the receiving direction is determined by cross-correlation analysis between the measured signal at the DRP and the original signal at the electrical access point. The measurement is corrected by delays which are caused by the test equipment.

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

8.8 Speech clipping

In view of the following considerations:

– that wireless systems may employ a variety of speech interpolation techniques as well as being susceptible to bursts of errors in the radio channel;

– excessive loss of speech signal may affect quality of a connection;

– subjective impact of clipping depends upon duration of clipping, percentage of speech clipped, frequency of clipping and overall speech activity.

The following is recommended for speech clipping, i.e., loss of speech:

– no speech loss occurrences longer than 64 ms should be present;

– speech loss periods shorter than 64 ms should be kept below 0.2 percent of active speech.

NOTE – Percent of clipped speech is 100 times the product of the frequency of speech clipping times clipping duration, divided by the speech activity factor.

8.8.1 Measurement methods

For further study.

8.9 Acoustic safety of telephone user

In order to ensure safety and to minimize annoyance to the user, the terminal shall meet limits specified in [ITU-T P.360].

8.9.1 Measurement method

The maximum steady state acoustic pressure is measured by applying maximum positive digital code to the receive input defined for the handset under test. The test procedure is in accordance with clause 8.2.3.1 for the receive characteristics except that the acoustic pressure in the artificial ear is measured with a sound level meter.

The measurement shall refer to [ITU-T P.360].

Terminals with adjustable receive levels shall be tested at the maximum setting.

8.10 Speech quality in presence of ambient noise

In view of the following considerations:

– mobile sets are being used often in noisy environments;

– the difficulties of high ambient noise conditions.

The handset is recommended to reduce the ambient noise picked up by the microphone(s) without significantly degrading the quality of the speech signal, and comply with the following requirements:

– S-MOS-LQOn:

The average of S-MOS-LQOn scores across all test conditions shall be ≥ 3.0, and as a performance objective, the average of the S-MOS-LQOn scores across all test conditions should be ≥ 3.5;

– N-MOS-LQOn:

The average of the N-MOS-LQOn scores across all test conditions shall be ≥ 2.3, and as a performance objective, the average of N-MOS-LQOn scores across all test conditions should be ≥ 3.0;

– G-MOS-LQOn:

No requirements for G-MOS-LQOn.

8.10.1 Measurement method

The speech quality in sending for narrowband systems is tested based on [ETSI TS 103 106]. The measurement is conducted for 8 noise conditions as described in Table 10. The measurements should be made in the same unique and dedicated call. The noise types shall be presented according to the order specified in Table 10.

Table 10 – Noise conditions used for ambient noise simulation

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

Before starting the measurements a proper conditioning sequence shall be used. The conditioning sequence shall be comprised of the four additional sentences 1-4 described in [ETSI TS 103 106], applied to the beginning of the 16-sentence test sequence. The conditioning signal level is −1.7 dBPa at the MRP, measured as the active speech level according to [ITU-T P.56].

NOTE – The sequence of speech samples concatenated for the test signal, consisting of alternating talkers in the sending direction, reduces the overall test time but may represent an unrealistic behaviour for certain voice enhancement technologies. Alternative concatenations are for further study.

The send speech signal consists of the 16 sentences of speech as described in [ETSI TS 103 106]. The test signal level is −1.7 dBPa at the MRP, measured as the active speech level according to [ITU-T P.56]. Three signals are required for the tests:

– the clean speech signal is used as the undisturbed reference (see [ETSI TS 103 106]);
– the speech plus undisturbed background noise signal is recorded at the terminal's microphone position using an omnidirectional measurement microphone with a linear frequency response between 50 Hz and 12 kHz;
– the send signal is recorded at the point of interconnection (POI).

N-MOS-LQOn, S-MOS-LQOn and G-MOS-LQOn are calculated as described in [ETSI TS 103 106] on a per sentence basis and averaged over all 16 sentences. The results shall be reported as average and standard deviation.

The measurement is repeated for each ambient noise condition described in Table 10, and the average of the results derived from all ambient noise types is calculated.

9 Narrowband headset technical requirements

Unless otherwise specified, measurements defined in clause 9 will be carried at the setting of the volume control where the RLR is as close as possible to nominal value.

9.1 Headset sending characteristics

9.1.1 SLR

The nominal values of SLR shall be: SLR = 8 ± 3 dB.

9.1.1.1 Measurement method

The measurement method narrowband headset is the same as for narrowband handset.

9.1.2 Sending frequency response

The sending sensitivity/frequency characteristics shall be as follows:

The sending sensitivity frequency response, measured from the MRP to the SS audio output (digital output of the reference speech decoder of the SS) with microphone at the recommended test position (RTP), shall be within a mask, which can be drawn between the points given in Table 11. The mask is drawn with straight lines between the breaking points in Table 11 on a logarithmic (frequency) – linear (dB sensitivity) scale.

Table 11 – Sending sensitivity/frequency mask-headset

Frequency (Hz)	Upper limit	Lower limit
100	–12	–
200	0	–
300	0	–12
1 000	0	–6
2 000	4	–6
3 000	4	–6
3 400	4	–9
4 000	0	–
NOTE – All sensitivity values are expressed in dB on an arbitrary scale.		

9.1.2.1 Measurement method

Measurement method for narrowband headset is the same as for narrowband handset, except that the headset should be mounted at the test position using a HATS according to [ITU-T P.380].

9.1.3 Idle channel noise

The maximum noise level produced by the apparatus at the output of the SS under silent conditions in the sending direction shall not exceed –64 dBm0p.

NOTE – This figure applies to the total noise level with psophometric weighting. It is recommended that the level of single frequency disturbances should be ≤ –74 dBm0p in the frequency range from 300 Hz-3.4 kHz.

9.1.3.1 Measurement method

The headset should be mounted according to [ITU-T P.380] using a HATS in a quiet environment, as defined in clause 7.2, the sending noise level at the SS audio output is measured with apparatus including psophometric weighting described in [ITU-T O.41].

9.1.4 Distortion

The sending part shall meet the following distortion requirements:

The sending distortion shall be measured between MRP and the SS audio output (output of the reference speech decoder of the SS). The ratio of signal-to-total distortion power measured shall be above the limits given in Table 12.

Table 12 – Limits for signal-to-total distortion ratio

Sending level (dBPa at the MRP)	Sending ratio (dB)
5	30
0	35
–4.7	35
–10	33
–15	30
–20	27

Limits for intermediate levels are found by drawing straight lines between the breaking points in the table on a linear (dB signal level) – linear (dB ratio) scale.

9.1.4.1 Measurement method

The measurement method for a narrowband headset is the same as for a narrowband handset, except that the headset should be mounted at the test position according to [ITU-T P.380] using a HATS.

9.2 Headset receiving characteristics

9.2.1 RLR

The nominal values of RLR shall be: RLR = 2 ± 3dB.

RLR (binaural headset) = 8 ± 3 dB for each earphone.

Where a user-controlled receiving volume control is provided, the RLR shall meet the nominal value for at least one setting of the control. When the control is set to maximum, the RLR shall not be ≤ (equal or louder than) −13 dB.

With the volume control set to the minimum position, the RLR shall not be ≥ (equal or quieter than) 18 dB and shall not be ≥ (equal or quieter than) 24 dB for a binaural headset.

9.2.1.1 Measurement method

Measurement method for a narrowband headset is the same as for a narrowband handset.

For binaural earphones, the receiving sensitivity equals the power sum of that measured with the left ear and right ear individually.

9.2.2 Receiving frequency response

The receiving sensitivity/frequency characteristics shall be as follows:

The receiving sensitivity frequency response, measured from the SS audio input (analogue or digital input of the reference speech encoder of the SS) to the ERP, shall be within a mask, which can be drawn with straight lines between the breaking points in Table 13 on a logarithmic (frequency) – linear (dB sensitivity) scale.

For binaural earphones, the left earphone and right earphone should meet the requirement specified in Table 13.

Table 13 – Receiving sensitivity/frequency mask-headset

Frequency (Hz)	Upper limit 8 ± 2 N	Lower limit 8 ± 2 N
100	6	–
300	6	–6
3 400	6	–6
4 000	6	–

NOTE 1 – All sensitivity values are expressed in dB on an arbitrary scale.

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

NOTE 3 – The basis for the target frequency responses in send and receive is the orthotelephonic reference response measured between 2 subjects 1 m apart under free-field conditions and assumes an ideal receive characteristic. Under these conditions the overall frequency response shows a rising slope. The present Recommendation now uses the diffuse-field as a reference point for receive and no longer uses the ERP. With the concept of diffuse-field based receive measurements, a rising slope for the overall frequency response is achieved by a flat target frequency response in send and a flat diffuse-field based receive frequency response.

9.2.2.1 Measurement method

Measurement method for a narrowband headset is the same as for a narrowband handset, except that the headset should be mounted at the test position using a HATS as specified in [ITU-T P.380].

9.2.3 Idle channel noise

The maximum (acoustic) noise level at the headset UE when no signal is applied to the input of the SS shall be as follows:

– if no user-controlled receiving volume control is provided, or, if it is provided at the setting of the user-controlled receiving volume control at which the RLR is equal to the nominal value, the noise measured at the ERP contributed by the receiving part alone shall not exceed –57 dBPa(A).

For binaural earphones, each receiver shall not exceed –60 dBPa(A):

– where a volume control is provided, the measured noise shall also not exceed –54 dBPa(A) at the maximum setting of the volume control.

For the nominal volume control setting, the level of single frequency disturbances should be ≤ 60 dBPa(A) in the frequency range from 100 Hz-10 kHz. As a performance objective it is recommended that the level should be ≤ 64 dBPa(A). For binaural earphones, each receiver shall not exceed –57 dBPa(A).

9.2.3.1 Measurement method

The headset should be mounted at test position using a HATS as specified in [ITU-T P.380] in a quiet environment, as defined in clause 7.2, the receiving noise level at the DRP with diffuse-field correction is measured with A-weighting.

9.2.4 Distortion

The receiving part between the SS audio input (input of the reference speech encoder of the SS) and DRP shall meet the requirements in this clause at the nominal setting of the volume control:

– The ratio of signal-to-total distortion power measured shall be above the limits given in Table 14 when the sound pressure at DRP with diffuse-field correction is up to +10 dBPa. For sound pressures exceeding +10 dBPa at the DRP with diffused correction, there is no distortion requirement.

Table 14 – Limits for signal-to-total distortion ratio

Frequency (Hz)	Receiving level at the digital interface (dBm0)	Receiving ratio at nominal volume setting (dB)	Receiving ratio at maximum volume setting (dB)
315	–16	20	–
408	–16	28	–
510	–16	28	–
816	–16	28	–
1 020	0	25.5	–
	–3	31.2	–
	–10	33.5	–
	–16	33.5	–
	–20	33	–
	–30	30.5	–
	–40 (Note)	22.5 (Note)	–
	–45 (Note)	17.5 (Note)	–
NOTE – For levels −40 and −45 dBm0 a lower signal-to-total distortion ratio may not be possible, and hence would not be regarded as a failing result. However, the obtained results would be reported.			

Limits for intermediate levels are found by drawing straight lines between the breaking points in the table on a linear (dB signal level) – linear (dB ratio) scale.

9.2.4.1 Measurement method

The measurement method for a narrowband headset is the same as for a narrowband handset, except that the headset should be mounted at the test position according to [ITU-T P.380] using a HATS.

9.3 Sidetone characteristics

The talker STMR shall be ≥ 15 dB and should be ≤ 23 dB for the nominal setting of the volume control. For all other positions of the volume control, the STMR shall be ≥ 10 dB.

9.3.1 Measurement method

The headset should be mounted at the test position according to [ITU-T P.380] using a HATS.

A real speech test signal, described in [ITU-T P.501] with a level of –4.7 dBPa shall be applied at the MRP. For each frequency given in Table 3 of [ITU-T P.79], bands 1 to 20, the sound pressure (at the ERP) shall be measured.

The test set-up shown in Figure 11 is used to measure the sidetone frequency response. The sidetone path loss L_{meST} and the STMR shall be calculated according to [ITU-T P.79], using equation 4-1 (m = 0.225) and the weighting factors in Table 3 of [ITU-T P.79].

9.4 TCLw

The TCLw for headset UE shall be ≥ 46 dB for any setting of the volume control.

The TCLw for headset UE should be ≥ 55 dB at the nominal setting of the volume control.

NOTE – It is recommended that the volume control should be set back to nominal after each call unless TCLw ≥ 55 dB can also be maintained with the maximum volume setting. Depending on the UE idle channel noise in the sending direction, it may not always be possible to measure an echo loss ≥ 55 dB.

9.4.1 Measurement method

The measurement method for a narrowband headset is the same as for a narrowband handset, except that the headset should be mounted at the test position according to [ITU-T P.380] using a HATS.

9.5 Acoustic safety of telephone user

In order to ensure safety and to minimize annoyance to the user, the terminal shall meet limits specified in [ITU-T P.360].

9.5.1 Measurement method

The maximum steady state acoustic pressure is measured by applying maximum positive digital code to the receive input defined for the handset under test. The test procedure is in accordance with clause 9.2.2.1 for the receive characteristics except that the acoustic pressure in the artificial ear is measured with a sound level meter.

The measurement shall refer to [ITU-T P.360].

Terminals with adjustable receive levels shall be tested at the maximum setting.

9.6 Delay

In view of the following considerations:

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

– the amount of delay introduced by wireless systems depends on specific technology and may be inherent to the adopted coding technique.

The following is recommended:

– delay added by the terminal equipment should be minimized in accordance with the guidelines provided in [ITU-T G.114] even with the use of echo control;

– The terminal specific implementation dependent processing delay， including both the delay in sending direction and the delay in receiving direction, should be less than 70ms.

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

9.6.1 Measurement method

The measurement method for a narrowband headset is the same as for a narrowband handset, except that the headset should be mounted at the test position according to [ITU-T P.380] using a HATS.

10 Narrowband speakerphone sets technical requirements

The requirements and test methods in this Recommendation are only for desktop operated and handheld loudspeaker mobile terminals. The transmission characteristics and test methods specified in [ITU-T P.340] and [ITU-T P.342] can be used for the two kinds of speakerphone sets.

Speakerphone sets use speech processing devices for acoustic enhancement (SPDA) to control acoustic echo, reduce background noise transmission, etc. These parameters are defined in [ITU-T P.330].

Appropriate test signals are described in [ITU-T P.50] and [ITU-T P.501]. Test methods appropriate for parameters defined in this Recommendation may be found in [ITU-T P.502]. Proper use of HATS testing may be found in [ITU-T P.581].

Unless otherwise specified, measurements defined in clause 10 will be carried at the setting of the volume control where the RLR is as close as possible to nominal value.

10.1 Loudspeaker sending characteristics

10.1.1 SLR

According to [ITU-T P.340], the SLR of a loudspeaker telephone should be about 5 dB higher than the SLR of the corresponding handset telephone.

Therefore, the nominal value of SLR for handheld and desktop speakerphone sets shall be 13 ± 4 dB.

10.1.1.1 Measurement method

The real speech signal as defined in [ITU-T P.501] shall be used. The test signal used shall be specified in the test report. The test signal level shall be –4.7 dBPa, measured at the MRP. The test signal level is averaged over the complete test signal sequence. The broadband signal level is then adjusted to −28.7 dBPa at the HFRP or the HATS HFRP as defined in [ITU-T P.581] and the spectrum is not altered.

The spectrum at the MRP and the actual level at the MRP (measured in 1/3-octaves) are used as references.

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The SLR shall be calculated according to [ITU-T P.79].

The sending sensitivity shall be calculated from each band of the 14 frequencies given in Table 1 of [ITU-T P.79], bands 4 to 17. For the calculation, the averaged measured level at the electrical reference point for each frequency band is referred to the averaged test signal level measured in each frequency band at the MRP.

The sensitivity is expressed in terms of dBV/Pa and the SLR shall be calculated according to formula (A-23b) of [ITU-T P.79], over bands 4 to 17, using m = 0.175 and the sending weighting factors from table 1 of [ITU-T P.79].

10.1.2 Sending frequency response

In view of the following considerations:

– ideally, the frequency response should be flat for optimal voice quality;

– many manufacturers use microphones and analogue filters to reduce noise;

the sending tolerance mask of handheld and desktop speakerphone sets for the nominal sensitivity/frequency response mask shown in Table 15 is recommended. The mask is drawn as straight lines between the breaking points in the table on a logarithmic (frequency) – linear (dB sensitivity) scale.

Table 15 – Speakerphone sending frequency response

Frequency (Hz)	Upper limit (dB)	Lower limit (dB)
100	–12	–
200	0	–
300	0	–12
1 000	0	–6
2 000	4	–6
3 000	4	–6
3 400	4	–9
4 000	0	–
NOTE – As stated in [ITU-T P.340], the interval between 200-300 Hz makes a significant contribution to the naturalness of the transmitted speech. However, the noise energy in a noisy environment in that frequency range might be significantly higher than the speech energy. Therefore, the manufacturers should consider this trade-off between naturalness of the transmitted speech and background noise transmitted when determining the optimal frequency response in the sending direction.		

10.1.2.1 Measurement methods

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The real speech signal as described in [ITU-T P.501] shall be used for the test. The type of test signal used shall be stated in the test report. The spectrum of acoustic signal produced by the artificial mouth is calibrated under free-field conditions at the MRP. The test signal level shall be –4.7 dBPa, measured at the MRP. The test signal level is averaged over the complete test signal sequence. The broadband signal level then is adjusted to the required value and the spectrum is not altered.

The spectrum at the MRP and the actual level at the MRP (measured in third octaves) are used as a reference to determine the sending sensitivity S_{mJ}.

Measurements shall be made at 1/3-octave intervals as given by [IEC 61260] for frequency bands from 100 Hz to 4 kHz inclusive. For the calculation, the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dBV/Pa.

10.1.3 Idle channel noise

The following limit is recommended:

– send noise level maximum –64 dBm0p.

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

10.1.3.1 Measurement methods

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The environment shall comply with the conditions described in clause 7.2.

The psophometric noise level at the output of the SS is measured. The psophometric filter is described in [ITU-T O.41].

10.1.4 Distortion

The ratio of signal to total distortion shall be above the mask defined in Table 16.

Table 16 – Signal to distortion ratio limit, sending

Sending level (dBPa at the MRP)	Sending ratio (dB)
5	30
0	35
–4.7	35
–10	33
–15	30
–20	27

10.1.4.1 Measurement methods

The measurement method for a narrowband speakerphone is the same as for a narrowband handset, except for the test set-up.

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

10.1.5 Out-of-band signals in sending direction

With any signal above 4.6 kHz and up to 8 kHz applied at the MRP at a level of –4.7 dBPa, the level of any image frequency shall be below the level obtained for the reference signal by at least the amount (in dB) specified in Table 17.

Table 17 – Out-of-band signal limit, sending

Frequency (kHz)	Signal limit (dB)
4.6	30
8	40
NOTE – The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) – logarithmic (kHz) scale.	

10.1.5.1 Measurement methods

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The signal used is an activation signal followed by a sine wave with the frequencies of 4.65 kHz, 5 kHz, 6 kHz, 6.5 kHz, 7 kHz and 7.5 kHz at the level of –4.7 dBPa which are the actual test signals. The test signal is applied at the MRP. Appropriate activation signals and signal combinations can be found in [ITU-T P.501].

10.2 Speakerphone set receiving characteristics

10.2.1 RLR

For a desktop speakerphone set, the nominal value of RLR shall be 5 ± 4 dB.

Where a user-controlled volume control is provided, the RLR shall meet the nominal value for at least one setting of the control. It is recommended that a volume control giving at least 15 dB increase from the nominal RLR (louder) is provided for hands-free units. This is to allow for the increased acoustic noise level in the usage environment.

RLR at the maximum volume control setting should be ≤ (equal or louder than) 1 dB.

For a handheld speakerphone set, the nominal value of RLR shall be 9 +9 / −7 dB.

As a performance objective it is recommended that the RLR at the maximum volume control setting is ≤ (equal to or louder than) 2 dB.

Where a user-controlled volume control is provided, the RLR shall meet the nominal value for at least one setting of the control. It is recommended that a volume control range of ≥ 15 dB be provided.

10.2.1.1 Measurement methods

The real speech signal as defined in [ITU-T P.501] shall be used. The test signal used shall be specified in the test report. The test signal level shall be −16 dBm0 measured at the digital reference point or the equivalent analogue point. The test signal level is averaged over the complete test signal sequence.

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The receiving sensitivity shall be calculated from each band of the 14 frequencies given in Table 1 of [ITU-T P.79], bands 4 to 17. For the calculation, the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dBPa/V and the RLR shall be calculated according to formula (A-23c) of [ITU-T P.79], over bands 4 to 17, using m = 0.175 and the receiving weighting factors from Table 1 of [ITU-T P.79].

Free-field correction is used. No leakage correction shall be applied.

10.2.2 Receiving frequency response

In view of the following considerations:

– ideally, the frequency response should be flat for optimal voice quality;

the receiving tolerance mask for a handheld speakerphone set shown in Table 18 and for a desktop speakerphone set shown in Table 19 are recommended. The mask is drawn as straight lines between the breaking points in the table on a logarithmic (frequency) – linear (dB sensitivity) scale.

Table 18 – Receiving frequency response-handheld

Frequency (Hz)	Upper limit	Lower limit
200	6	
500	6	–9 (Note 2)
630	6	–6 (Note 2)
800	6	–6
3 100	6	–6
4 000	6	

NOTE 1 – All sensitivity values are expressed in dB on an arbitrary scale.
NOTE 2 – The values stated in the Table 6 for 500 and 630 Hz are listed for performance objective purposes (not mandatory).

Table 19 – Receiving frequency response-desktop

Frequency (Hz)	Upper limit	Lower limit
200	6	
315	6	–9
400	6	–6
3 100	6	–6
4 000	6	
NOTE – All sensitivity values are expressed in dB on an arbitrary scale.		

10.2.2.1 Measurement method

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The real speech test signal as described in [ITU-T P.501] shall be used for the test. The type of test signal used shall be stated in the test report. The test signal level shall be –16 dBm0, measured at the digital reference point or the equivalent analogue point. The test signal level is averaged over the complete test signal sequence.

If a HATS is used, then it is free-field equalized as described in [ITU-T P.581]. The equalized output signal of each artificial ear is power-averaged over the total duration of the analysis; the right and left artificial ear signals are voltage-summed for each 1/3-octave band frequency band; these 1/3-octave band data are considered as the input signal to be used for calculations or measurements. Measurements shall be made at 1/3-octave intervals as given by [IEC 61260] for frequency bands from 100 Hz-4 kHz inclusive. For the calculation, the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dBPa/V.

10.2.3 Idle channel noise

The following limit is recommended at the nominal RLR setting:

– receive noise level maximum –53 dBPa(A).

Spectral peaks shall be less than 10 dB above the average noise spectrum. A narrowband FFT analysis with a frequency resolution of 32 Hz or less is recommended.

The noise produced by the speakerphone terminal providing volume control shall be less than –45 dBPa(A) with the volume control set to its maximum level (minimum RLR setting).

10.2.3.1 Measurement method

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The environment shall comply with the conditions described in clause 7.2.

Under quiet conditions, as defined in clause 7.2, the receiving noise level at the DRP with free-field correction is measured with A-weighting.

10.2.4 Distortion

The receiving part between the SS audio input (input of the reference speech encoder of the SS) and the DRP with free-field correction shall meet the requirements in this subclause at the nominal setting of the volume control:

– The ratio of signal to total distortion power measured shall be above the limits given in Table 20 when the sound pressure at the DRP with free-field correction is up to 10 dBPa. For a sound pressure ≥ 10 dBPa at the DRP with free-field correction, there is no distortion requirement.

Table 20 – Signal to distortion ratio limit, receiving

Frequency (Hz)	Receiving level at the digital interface (dBm0)	Receiving ratio at nominal volume setting (dB)	Receiving ratio at maximum volume setting (dB)
1 020	0	25.5	–
	–3	31.2	–
	–10	33.5	–
	–16	33.5	–
	–20	33	–
	–30	30.5	–
	–40 (Note)	22.5 (Note)	–
	–45 (Note)	17.5 (Note)	–
NOTE –For levels −40 and −45 dBm0 a lower signal-to-total distortion ratio may not be possible, and hence would not be regarded as a failing result. However, the obtained results would be reported.			

10.2.4.1 Measurement method

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The signal used is an activation signal followed by a series of sine wave signals with a frequency of 1 020 Hz, which is the actual test signal. Appropriate signals for activation and signal combinations can be found in [ITU-T P.501]. The signal level shall be calibrated to –16 dBm0.

The signal level shall be applied at the signal input of the SS at the following levels: 0, –3, –10, –16, –20, –30, –40, –45 dBm0. The test signals have to be applied in this sequence, i.e., from high levels down to low levels.

The total distortion power shall be measured at the DRP with free-field correction and calculated with the psophometric noise weighting as specified in [ITU-T O.41].

10.2.5 Linearity in receiving direction

The audio components of the speakerphone terminal should operate in a linear fashion. This is not only for optimal voice quality, but also for optimal performance of the acoustic echo controller (AEC) within the speakerphone terminal. Specific requirements are for further study.

10.2.6 Spurious out-of-band signals

The test signal is a digitally-simulated sine wave signal in the frequency range of 300 Hz to 3.4 kHz and at a level of –5 dBm0 applied at the digital interface.

The level of spurious out-of-band image signals in the frequency range of 4.6 to 8 kHz measured selectively at the artificial ear shall be at least lower than the in-band acoustic level produced by a digital signal at 1 kHz set at the level specified in Table 21.

Table 21 – Discrimination levels

Image signal frequency (kHz)	Equivalent input signal level (dBm0)
4.6	–35
8	–45
NOTE – The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) – logarithmic (kHz) scale.	

10.2.6.1 Measurement method

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The signal used is an activation signal followed by a sine wave signal at the frequencies 500, 1 000, 2 000 and 3 150 Hz which is the actual test signal. The signal is applied at the input of the SS at the level of –5 dBm. Appropriate signals for activation and signal combinations can be found in [ITU-T P.501]. For the test signals at the frequencies 500, 1 000, 2 000 and 3 150 Hz, the level of spurious out-of-band image signals at frequencies of up to 8 kHz shall be measured selectively at the artificial ear.

10.3 TCLw

In order to meet the ITU-T G.131 talker echo objective requirements, the weighted terminal coupling loss during single talk (TCLwst) should be greater than 46 dB.

For terminals fitted with a volume control, the TCLwst shall be not less than 40 dB for any setting of the volume control.

The limits for weighted terminal coupling loss during double talk (TCLwdt) are defined in clause 10.5.

10.3.1 Measurement method

The speakerphone is set up in a room with acoustic properties similar to a typical "office-type" room, and the ambient noise level should ≤ 70 dBPa(A).

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The attenuation from reference point input to reference point output shall be measured using the compressed real speech signal described in clause 7.3.3 of [ITU-T P.501]. The test signal level shall be −10 dBm0.

The TCLw is calculated according to clause B.4 (trapezoidal rule) of [ITU-T G.122], using the frequency range 300 to 3 350Hz. For the calculation, the averaged measured echo level at each frequency band is referred to the averaged test signal level measured in each frequency band. The first 17.0 s of the test signal (6 sentences) are discarded from the analysis to allow for convergence of the acoustic echo canceller. The analysis is performed over the remaining length of the test sequence (last 6 sentences).

10.4 Switching characteristics

Many mobile speakerphone terminals use a voice activity detector (VAD) and a loss controller to control acoustic echo. The VAD distinguishes silent periods (no active speech signal), single-talk periods (near-end speech periods or far-end speech periods) and double-talk periods (near-end and far-end speech signals active at the same time). According to the state determined by the VAD, the loss controller reduces the acoustic echo level by inserting variable losses on the received and/or transmitted audio signals. (The VAD and loss controller may also be used to control electrical echo from the network if the mobile terminal operates on an analogue wireless network without echo control.) An AEC and/or NLP may also be used to control acoustic echo. However, most devices with an AEC and/or NLP also require a loss controller to control residual acoustic echo. Although no VAD is perfect and instantaneous, errors by the VAD may lead to speech clipping and acoustic echo.

The switching parameters and switching functions are described in detail in [ITU-T P.330] and [ITU-T P.340].

The minimum activation level in the sending direction shall be ≤ –20 dBPa. The build-up time for activation in the sending direction shall be:

$$\text{TRst-s} \leq 50 \text{ ms}$$

The minimum activation level in the receiving direction shall be ≤ –35.7 dBm0. The build-up time for activation in the receiving direction shall be:

$$\text{TRst-r} \leq 50 \text{ ms}$$

Requirements for the attenuation and build-up time for receive-to-send states (and send-to-receive states) are for further study.

10.4.1 Measurement method

The set-up for a desktop speakerphone set is described in [ITU-T P.581].

The set-up for a handheld speakerphone set is described in Figure 3.

The build-up time, single talk, send signal (TRst-s) and build-up time, single talk, receive signal (TRst-r) shall be measured according to [ITU-T P.340].

The echo loss measurement during double talk should be carried out according to [ITU-T P.502].

10.5 Double talk performance

The speech quality during double talk is mainly determined by two parameters: talker echo loudness rating (TELR) (impairment caused by echo during double talk, and related to TCLwdt) and level variation between single and double talk (attenuation range).

In order to guarantee sufficient quality under double talk conditions, the TELR should be high and the attenuation inserted should be as low as possible. Terminals that do not allow double talk should provide a good echo attenuation realized by high attenuation.

The most important parameters to determine the quality during double talk are as follows:

– attenuation range in sending direction during double talk (Asdt);

– attenuation range in receiving direction during double talk, (Ardt);

– echo attenuation during double talk.

The requirements for Asdt and Ardt at nominal RLR for each category of speakerphone terminals are provided in Table 4 of [ITU-T P.340]. Higher values for Asdt and Ardt are permitted at lower RLR (due to increasing the received volume level) without affecting the behaviour characteristics. However, the talker echo loudness rating during double talk (TELRdt) must maintain the requirements in [ITU-T P.340] for all received volume settings.

Subjective evaluation methods for double talk quality, as well as single talk quality, are described in [ITU-T P.832].

NOTE – Acoustic echo control is more easily and effectively done in terminals. However, network equipment, as described in [ITU-T G.168], may also include acoustic echo control processing. This leads to tandeming issues of acoustic echo control functionalities. As specified in [ITU-T G.168], the added network component shall prevent any degradation of the overall perceived quality. In practice, however, double-talk capability may be reduced due to network equipment.

10.5.1 Measurement method

The set-up for a desktop speakerphone set is described in [ITU-T P.581].

The set-up for a handheld speakerphone set is described in Figure 3.

The attenuation range in sending and receiving direction during double talk, Asdt and Ardt, shall be measured according to [ITU-T P.340].

10.6 Background noise transmission and comfort noise injection

Mobile speakerphone terminals are typically used in noisy environments.

Most speakerphone terminals have some type of NR capability. The main purpose of a NR system in a device is to reduce the annoying and fatiguing effects of the transmitted background noise. The techniques used to reduce background noise may be classified as analogue only, digital only, and combined analogue and digital techniques. These techniques are described in [ITU-T P.330].

Some parameters describing NR and some requirements are also given in [ITU-T P.330]. There is usually a trade-off between the level of digital NR and the voice quality. Too much NR may reduce the transmitted voice level and distort the speech signal.

Without comfort noise, the acoustic echo control system (loss controller, VAD, NLP and AEC) will reduce and/or distort the background noise transmitted during receive and double talk states. This can be annoying to the far-end talker. Some speakerphone terminals inject comfort noise in order to maintain a constant level of background noise transmitted. Some parameters for comfort noise are described in [ITU-T P.330].

The level of comfort noise should be within a range of +2 and –5 dB from the original (transmitted) background noise.

10.6.1 Measurement methods

For further study.

11 Wideband handset technical requirements

Unless otherwise specified, measurements defined in clause 8 shall be carried at the setting of the volume control where the RLR is as close as possible to nominal value.

11.1 Sending characteristics

11.1.1 SLR

In view of the following considerations:

– wireless terminals provide connectivity to the existing wireline-based international telecommunication networks;

– digital wireless access networks should provide the same signal levels as the digital access wireline networks.

The following LR value is recommended:

– a nominal value of SLR = 8 ± 3 dB.

NOTE – Several reports have indicated that sending speech levels in mobile networks are 5 dB higher on average than speech levels from fixed networks. The most likely cause is that mobile terminals are frequently used in noisy environments where users tend to talk louder. Manufacturers of mobile handsets should be aware that this may lead to saturation, clipping and distortion of the signal and reduced speech quality. Changing

SLR at this time is not recommended due to concerns of reduced speech quality when used at normal talker levels. However, the manufacturers of devices intended for use in noisy conditions can consider adopting a higher SLR.

11.1.1.1 Measurement method

The real speech signal as defined in [ITU-T P.501] shall be used. The test signal used shall be specified in the test report. The test signal level shall be –4.7 dBPa, measured at the MRP. The test signal level is averaged over the complete test signal sequence.

The handset is mounted according to [ITU-T P.64].

The sending sensitivity shall be calculated from each band of the 20 frequencies given in Table A.1 of [ITU-T P.79], bands 1 to 20. For the calculation, the averaged measured level at the electrical reference point for each frequency band is referred to the averaged test signal level measured in each frequency band at the MRP.

The sensitivity is expressed in terms of dBV/Pa and the SLR shall be calculated according to formula (A-23b) of [ITU-T P.79], over bands 1 to 20, using m = 0.175 and the sending weighting factors from Table A.2 of [ITU-T P.79].

11.1.2 Sending frequency responses

In view of the following considerations:

– the compatibility with wireline digital telephones and the mixed analogue/digital telephone network;

– the compatibility with most existing wireless systems;

– the aim to achieve the best possible overall quality with the cordless and mobile terminals sending nominal sensitivity/frequency response within the limits given in Table 22 is recommended.

Table 22 – Sending

Frequency (Hz)	Upper limit	Lower limit
100	0	
200	5	–5
5 000	5	–5
6 300	5	–10
8 000	5	
NOTE – All sensitivity values are expressed in dB on an arbitrary scale.		

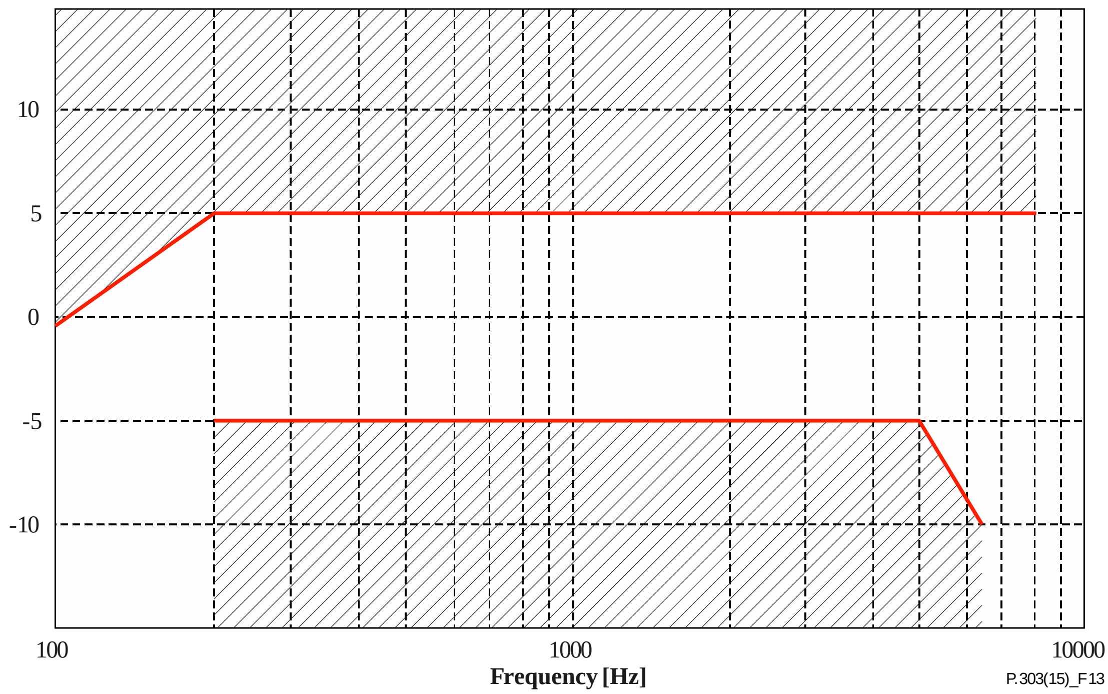

Figure 13 – Sending mask

NOTE – The basis for the target frequency responses in send and receive is the orthotelephonic reference response measured between 2 subjects 1 m apart under free-field conditions and assumes an ideal receive characteristic. Under these conditions the overall frequency response shows a rising slope. The present Recommendation now uses the diffuse-field as the reference point for receive and no longer uses the ERP. With the concept of diffuse-field based receive measurements a rising slope for the overall frequency response is achieved by a flat target frequency response in send and a flat diffuse-field based receive frequency response.

11.1.2.1 Measurement method

The sending frequency response is measured according to [ITU-T P.64], using the measurement set-up shown in Figure 14.

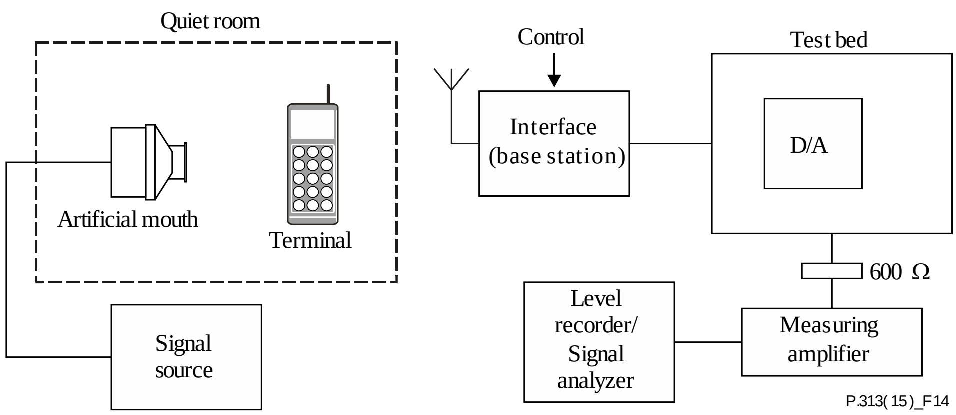

Figure 14 – Sending frequency response measurement method

The real speech signal as defined in [ITU-T P.501] shall be used. The test signal used shall be specified in the test report. The test signal level shall be –4.7 dBPa, measured at the MRP. The test signal level is averaged over the complete test signal sequence.

The handset is mounted according to [ITU-T P.64].

Measurements shall be made at 1/12-octave intervals as given by [IEC 61260] for frequency bands from 100 Hz to 8 kHz inclusive. For the calculation, the averaged measured level at the electrical reference point for each frequency band is referred to the averaged test signal level measured in each frequency band at the MRP.

The sensitivity is expressed in terms of dBV/Pa.

11.1.3 Idle channel noise

The maximum noise level produced by the apparatus at the output of the SS under silent conditions in the sending direction shall not exceed –64 dBm0(A).

NOTE – This figure applies to the total noise level with A-weighting. It is recommended that the level of single frequency disturbances should be ≤ –74 dBm0(A) in the frequency range from 100 Hz to 8 kHz.

11.1.3.1 Measurement method

The handset is mounted according to [ITU-T P.64] in a quiet environment, as defined in clause 7.2, the sending noise level at the SS audio output is measured with A-weighting.

11.1.4 Distortion

The sending distortion includes non-linear distortion and quantizing distortion. The sending part shall meet the following distortion requirements:

Limits for intermediate levels are found by drawing straight lines between the breaking points in the Table 23 on a linear (dB signal level) – linear (dB ratio) scale.

Table 23 – Limits for signal-to-total distortion ratio

Frequency	Sending level (dBPa at the MRP)	Sending ratio (dB)
315	–4.7	28
408	–4.7	32
510	–4.7	32
816	–4.7	32
1 020	5	30
	0	35
	–4.7	35
	–10	33
	–15	30
	–20	27

11.1.4.1 Measurement method

The handset is mounted according to [ITU-T P.64] in a quiet environment as defined in clause 7.2.

The signal used is an activation signal followed by a sine wave signal with a frequency of 315 Hz, 408 Hz, 510 Hz, 816 Hz and 1 020 Hz. The sine wave signal level shall be calibrated to –4.7 dBPa at the MRP for all frequencies, except for the sine wave with a frequency 1 020 Hz which shall be applied at the following levels at the MRP: 5, 0, –4.7, –10, –15, –20 dBPa. The test signals have to be applied in this sequence, i.e., from high levels down to low levels. Appropriate signals for activation and signal combinations can be found in [ITU-T P.501].

The total distortion power of the signal output of the SS shall be measured with the psophometric noise weighting described in [ITU-T O.41].

11.2 Receiving characteristics

An application force of 8 ± 2N should be used when measuring with type 3.3 or 3.4 artificial ears. This force is considered to be a realistic average value.

11.2.1 RLR

In view of the following considerations:

– wireless terminals provide connectivity to the existing wireline based international telecommunication networks;

– digital wireless terminals should be compatible with the digital access wireline networks.

The following LR value is recommended:

– a nominal value of RLR = 2 ± 3 dB.

11.2.1.1 Measurement method

The real speech signal as defined in [ITU-T P.501] shall be used. The test signal used shall be specified in the test report. The test signal level shall be −16 dBm0 measured at the digital reference point or the equivalent analogue point. The test signal level is averaged over the complete test signal sequence.

The handset is mounted according to [ITU-T P.64].

The receiving sensitivity shall be calculated from each band of the 20 frequencies given in Table A.2 of [ITU-T P.79], bands 1 to 20. For the calculation, the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dBPa/V and the RLR shall be calculated according to formula (A-23c) of [ITU-T P.79], over bands 1 to 20, using m = 0.175 and the receiving weighting factors from Table A.2 of [ITU-T P.79] .

DRP-ERP correction is applied. No leakage correction shall be applied.

11.2.2 Volume control

In view of the following considerations of mobile handsets:

– the mobile terminals are used often in noisy environment;

– the need to provide service for people with hearing impairments. For these applications, [ITU-T P.370] applies.

Where a user-controlled receiving volume control is provided, the RLR shall meet the nominal value for at least one setting of the control. When the control is set to maximum, the RLR shall not be ≤ (equal or louder than) −13 dB.

With the volume control set to the minimum position the RLR shall not be ≥ (equal or quieter than) 18 dB.

11.2.3 Receiving frequency responses

In view of the following considerations:

– the compatibility with wireline digital telephones and the mixed analogue/digital telephone network;

– the compatibility with most existing wireless systems;

– the aim to achieve the best possible overall quality with the cordless and mobile terminals.

The sensitivity/frequency characteristics shall be as follows:

The receiving sensitivity frequency response, measured either from the digital interface to the DRP with diffuse-field correction, or from the SS audio input (analogue or digital input of the reference speech encoder of the SS) to the DRP with diffuse-field correction, shall be within a mask, which can be drawn with straight lines between the breaking points in Table 24 on a logarithmic (frequency) − linear (dB sensitivity) scale.

Table 24 – Receiving sensitivity

Frequency (Hz)	Upper limit (dB)	Lower limit (dB)
100	6	
200	6	–10
300	6	–6
1 000	6	–6
2 000	8	–6
5 000	8	–6
6 300	8	–12
8 000	8	

11.2.3.1 Measurement method

The handset should be mounted according to [ITU-T P.64], using a suitable artificial ear as recommended in [ITU-T P.57].

The receiving frequency response is measured, using the measurement set-up shown in Figure 15. Terminals with adjustable receive levels shall be adjusted so that their RLR is as close as possible to the nominal value for this test.

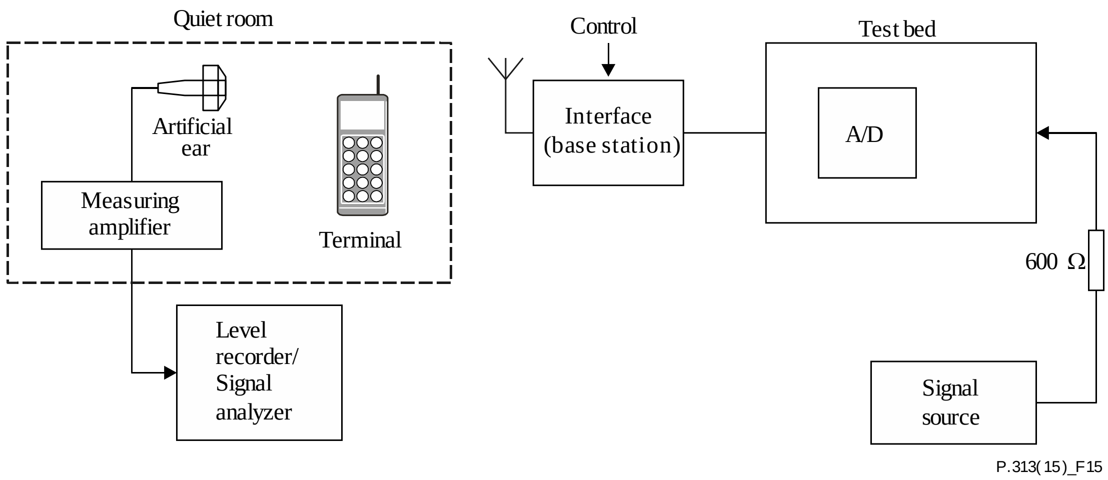

Figure 15 – Receive frequency response measurement method

The real speech signal (as specified [ITU-T P.501]) and a spectrum analyser can be used. The test signal used shall be specified in the test report. The test signal level shall be –16.0 dBm0, measured at the digital reference point or the equivalent analogue point. The test signal level is averaged over the complete test signal sequence.

Measurements shall be made at 1/12-octave intervals as given by [IEC 61260] for frequency bands from 100 Hz-8 kHz inclusive. For the calculation, the averaged level at each frequency band measured at HATS DRP with diffused-field correction is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dBPa/V. Information about correction factors is available in [ITU-T P.57].

11.2.4 Idle channel noise

The maximum (acoustic) noise level at the handset terminal when no signal is transmitted to the input of the SS shall be as follows:

– if no user-controlled receiving volume control is provided, or, if it is provided, at the setting of the user-controlled receiving volume control at which the RLR is equal to the nominal value, the noise measured at the DRP with diffuse-field correction contributed by the receiving equipment alone shall be ≤ –57 dBPa(A);

– where a volume control is provided, the measured noise shall also not exceed –54 dBPa(A) at the maximum setting of the volume control.

For the nominal volume control setting, the level of single frequency disturbances should be ≤ −60 dBPa(A) in the frequency range from 100 Hz to 10 kHz. As a performance objective it is recommended that the level should be ≤ −64 dBPa(A).

NOTE – In a connection with the PSTN, noise conditions as described in [ITU-T G.103] can be expected at the POI of the 3G or LTE network. The characteristics of this noise may be influenced by the speech transcoding process (for further study).

11.2.4.1 Measurement method

With the handset mounted according to [ITU-T P.64], using a suitable artificial ear as recommended in [ITU-T P.57], placed in a quiet environment, as defined in clause 7.2, the receiving noise level at DRP with diffuse-field correction is measured with A-weighting.

11.2.5 Distortion

The receiving distortion includes non-linear distortion and quantizing distortion.

The receiving part between the SS audio input (input of the reference speech encoder of the SS) and DRP with diffuse-field correction shall meet the requirements given in Table 25 at the nominal setting of the volume control. For sound pressures exceeding +10 dBPa at the DRP with diffuse-field correction, there is no distortion requirement.

Table 25 – Limits for signal-to-total distortion ratio

Frequency (Hz)	Receiving level at the digital interface (dBm0)	Receiving ratio at nominal volume setting (dB)	Receiving ratio at maximum volume setting (dB)
315	–16	20	–
408	–16	28	–
510	–16	28	–
816	–16	28	–
1 020	0	25.5	–
	–3	31.2	–
	–10	33.5	–
	–16	33.5	–
	–20	33	–
	–30	30.5	–

Table 25 – Limits for signal-to-total distortion ratio

Frequency (Hz)	Receiving level at the digital interface (dBm0)	Receiving ratio at nominal volume setting (dB)	Receiving ratio at maximum volume setting (dB)
	–40 (Note)	22.5 (Note)	–
	–45 (Note)	17.5 (Note)	–
NOTE – For levels −40 and −45 dBm0 a lower signal-to-total distortion ratio may not be possible, and hence would not be regarded as a failing result. However, the obtained results would be reported.			

Limits for intermediate levels are found by drawing straight lines between the breaking points in the table on a linear (dB signal level) – linear (dB ratio) scale.

11.2.5.1 Measurement method

The handset should be mounted according to [ITU-T P.64], using a suitable artificial ear as recommended in [ITU-T P.57].

The signal used is an activation signal followed by a sine wave signal with frequencies of 315 Hz, 408 Hz, 510 Hz, 816 Hz and 1 020 Hz. The signal level shall be −16 dBm0, except for the sine wave signal with a frequency 1 020 Hz that shall be applied at the signal input of the SS at the following levels: 0, –3, –10, –16, –20, –30, –40, –45 dBm0. The test signals have to be applied in this sequence, i.e., from high levels down to low levels. Appropriate signals for activation and signal combinations can be found in [ITU-T P.501].

The total distortion power shall be measured at the DRP with diffuse-field correction and calculated with the psophometric noise weighting as specified in [ITU-T O.41].

11.3 Sidetone characteristics

11.3.1 Sidetone masking rating (STMR)

In view of the following considerations:

– the optimum STMR for conditions free from echo;

– the difficulties of high ambient noise conditions.

The following is recommended:

– the value of the STMR shall be ≥ 15 dB and should be ≤ 23 dB for the nominal setting of the volume control;

– for all other positions of the volume control, the STMR shall be ≥ 10 dB.

In case the STMR is below the lower limit also when the electrical sidetone path has been disabled, the result shall not be regarded as a failure.

11.3.1.1 Measurement methods

A real speech test signal in [ITU-T P.501] with level of –4.7 dBPa shall be applied at the MRP. For each frequency given in Table 3 of [ITU-T P.79], bands 1 to 20, the sound pressure (at the ERP) shall be measured.

The handset should be mounted according to [ITU-T P.64], using a suitable artificial ear as recommended in [ITU-T P.57].

The test set-up shown in Figure 16 is used to measure the sidetone frequency response. The sidetone path loss L_{meST} and the STMR shall be calculated according to [ITU-T P.79], using Equation 4-1 (m = 0.225) and the weighting factors in Table 3 of [ITU-T P.79].

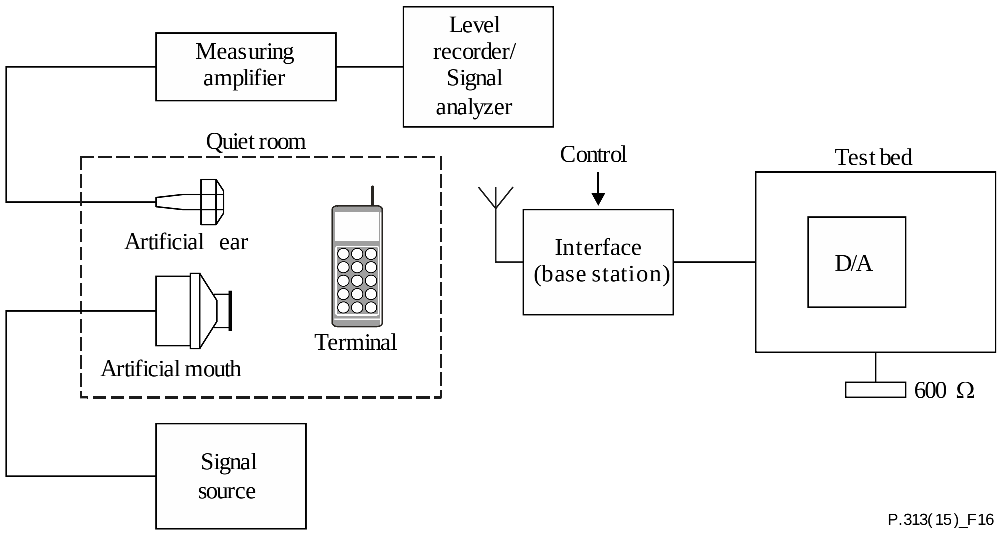

Figure 16 – Sidetone measurement method

11.4 Noise contrast and comfort noise

In some circumstances, such as application of voice-operated devices, the continuous background noise present, regardless of whether the users are talking or not, may be interrupted. This switching on and off is annoying to the users and may in fact degrade speech intelligibility. To reduce this effect, noise contrast should be minimized by increasing signal-to-noise ratio.

Comfort noise may be injected during silent periods to reduce the impairments created by the noise contrast. This may create undesirable performance degradation by itself if not done properly, due to the level or spectrum contents differences between the injected and the transmitted noise. Efforts should be made to match the characteristics of the injected comfort noise to the transmitted noise to reduce any perceptible contrast between them.

11.4.1 Measurement method

For further study.

11.5 TCLw

In view of the following considerations:

– the aim to achieve as high an acoustic coupling loss as possible to minimize degradation caused by echo;

– that the far-end talker echo should be controlled under all volume control settings, and for the range of transducers sensitivities as long as the handset is properly used;

– the far-end terminal may be connected via a mobile or an IP network which introduces long talker echo path delay;

– what is practically obtainable in real use where the user himself chooses the way to hold his handset.

The following limit is recommended:

– in order to meet the ITU-T G.131 talker echo objective requirements, for handsets fitted with a volume control, the TCLw should be greater than 46 dB for any setting of the volume control. The TCLw should be greater than 55 dB at the nominal setting of the volume control;

– it is recommended that the volume control should be set back to nominal after each call unless TCLw ≥ 55 dB can also be maintained with the maximum volume setting. Depending on the UE idle channel noise in the sending direction, it may not always be possible to measure an echo loss greater than 55 dB.

NOTE – In consideration of the increasing delays introduced by modern networks, a higher TCLw value than that specified here may be necessary for proper operation with these networks.

11.5.1 Measurement method

The handset should be mounted according to [ITU-T P.64].

Noise and reflections in the test space must not influence the measurement. The test should be performed in the environment as defined in clause 7.2.

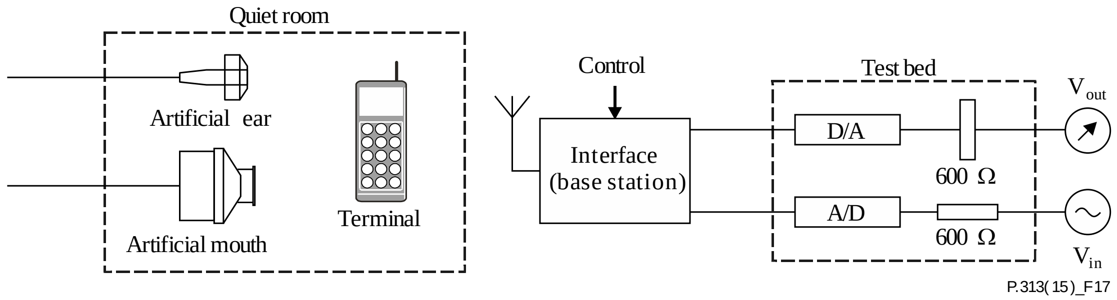

Figure 17 – Terminal coupling loss measurement method

The attenuation from digital input to digital output is measured at one-twelfth octave frequencies as given by [IEC 61260] for frequency bands from 300 to 6 700 Hz, using the measurement arrangement shown in Figure 17.

The test signal used is the compressed real speech signal described in clause 7.3.3 of [ITU-T P.501]. The test signal level shall be −10 dBm0. Terminals with adjustable receive levels shall be tested at the nominal setting. For the nominal setting, adjust the level so that the RLR is as close as possible to the nominal RLR value.

The TCLw is calculated according to clause B.4 (trapezoidal rule) of [ITU-T G.122]. For the calculation, the averaged measured echo level at each frequency band is referred to the averaged test signal level measured in each frequency band. The first 17.0 s of the test signal (6 sentences) are discarded from the analysis to allow for convergence of the acoustic echo canceller. The analysis is performed over the remaining length of the test sequence (last 6 sentences).

NOTE – There might be problems measuring 46 dB TCL in the case where sophisticated coding with limited dynamic range is used. In such cases, typically speech or speech-like test signals need to be used that themselves have crest factors in the range of 15 dB which reduces the measurement dynamic by the same amount. In such cases, the signal measured in the sending direction should be evaluated more carefully in order to find whether an echo signal is present, or whether this signal is completely masked by the noise signal introduced by the codec. If the signal measured in the sending direction is completely masked by the noise, the requirement can be considered to be fulfilled. If this is not the case, more sophisticated measurement procedures such as time averaging (in order to improve the signal-to-noise ratio) need to be applied in order to achieve reliable measurement results.

11.6 Stability loss

The minimum stability loss at any volume control setting should be at least 6 dB.

11.6.1 Measurement method

Before the actual test a training sequence consisting of the real speech signal described in [ITU-T P.501] is applied. The training sequence level shall be –16 dBm0 in order to not overload the codec.

The test signal is a PN-sequence complying with [ITU-T P.501] with a length of 4 096 points (for a 48 kHz sampling rate system) and a crest factor of 6 dB instead of 11 dB. The PN-sequence is generated as described in [ITU-T P.501] with W(k) constant within the frequency range 100 to 8 000 Hz and zero outside this range. The duration of the test signal is 250 ms with an input signal of −3 dBm0.

With the handset and transmission circuit fully active, measure the attenuation from the digital input to the digital output using the following method.

Place the handset in the reference corner, as shown in Figure 18, with the earcap and mouthpiece facing a hard, smooth surface with at least 400 mm free space in all directions.

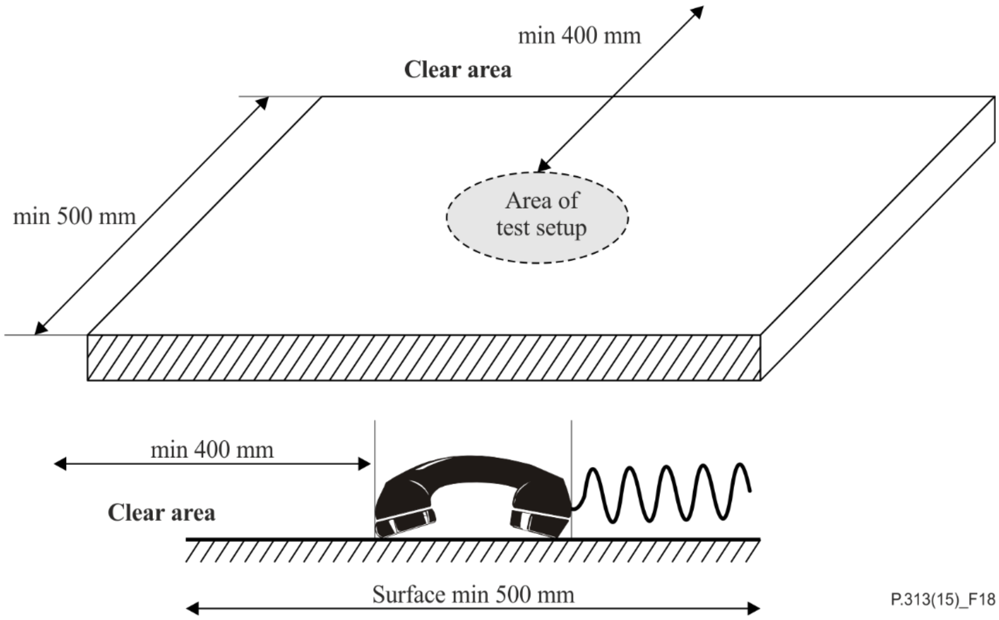

Figure 18 – Reference corner

11.7 Delay

In view of the following considerations:

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

– the amount of delay introduced by wireless systems depends on specific technology and may be inherent to the adopted coding technique.

The following is recommended:

– delay added by the terminal equipment should be minimized in accordance with the guidelines provided in [ITU-T G.114] even with the use of echo control;

– the terminal specific implementation dependent processing delay, including both the delay in sending direction and the delay in receiving direction, should be less than 70 ms.

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

11.7.1 Measurement method

The delay in the sending direction is obtained by measuring the delay between MRP and the electrical access point of the test equipment and subtracting the delays introduced by the test equipment from the measured value. The delay in the receiving direction is obtained by measuring the delay between the electrical access point of the test equipment and the DRP and subtracting the delays introduced by the test equipment from the measured value.

For the measurements, a CSS according to [ITU-T P.501] is used. The PN part of the CSS has to be longer than the maximum expected delay. It is recommended to use a PN sequence of 16 k samples (with 48 kHz sampling rate or equivalent).

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

The delay in the sending direction is determined by the cross-correlation analysis between the measured signal at the electrical access point and the original signal at MRP. The delay in the receiving direction is determined by cross-correlation analysis between the measured signal at the DRP and the original signal at the electrical access point. The measurement is corrected by delays which are caused by the test equipment.

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

11.8 Speech clipping

In view of the following considerations:

– that wireless systems may employ a variety of speech interpolation techniques as well as being susceptible to bursts of errors in the radio channel;

– excessive loss of speech signal may affect quality of a connection;

– subjective impact of clipping depends upon duration of clipping, percentage of speech clipped, frequency of clipping and overall speech activity.

The following is recommended for speech clipping, i.e., loss of speech:

– no speech loss occurrences longer than 64 ms should be present;

– speech loss periods shorter than 64 ms should be kept below 0.2 percent of active speech.

NOTE – Percent of clipped speech is 100 times the product of the frequency of speech clipping times clipping duration, divided by the speech activity factor.

11.8.1 Measurement methods

For further study.

11.9 Acoustic safety of telephone user

In order to ensure safety and to minimize annoyance to the user, the terminal shall meet the limits specified in [ITU-T P.360].

11.9.1 Measurement method

The maximum steady state acoustic pressure is measured by applying maximum positive digital code to the receive input defined for the handset under test. The test procedure is in accordance with clause 11.2.3.1 for the receive characteristics except that the acoustic pressure in the artificial ear is measured with a sound level meter.

The measurement shall refer to [ITU-T P.360].

Terminals with adjustable receive levels shall be tested at the maximum setting.

11.10 Speech quality in the presence of ambient noise

In view of the following considerations:

– mobile sets are being used often in noisy environments;

– the difficulties of high ambient noise conditions.

The handset is recommended to reduce the ambient noise picked up by the microphone(s) without significantly degrading the quality of the speech signal, and comply with the following requirements:

– S-MOS-LQOw:

The average of S-MOS-LQOw scores across all test conditions shall be ≥ 3.0, and as a performance objective, the average of the S-MOS-LQOw scores across all test conditions should be ≥ 3.5.

– N-MOS-LQOw:

The average of the N-MOS-LQOw scores across all test conditions shall be ≥ 2.3, and as a performance objective, the average of N-MOS-LQOw scores across all test conditions should be ≥ 3.0.

– G-MOS-LQOw:

No requirements for G-MOS-LQOw.

11.10.1 Measurement method

The speech quality in sending for narrowband systems is tested based on [ETSI TS 103 106]. The measurement is conducted for eight noise conditions as described in Table 26. The measurements should be made in the same unique and dedicated call. The noise types shall be presented according to the order specified in Table 26.

Table 26 – Noise conditions used for ambient noise simulation

Description	File name	Duration	Level	Type
Recording in pub	Pub_Noise_binaural_V2	30 s	L: 75.0 dB(A) R: 73.0 dB(A)	Binaural
Recording at pavement	Outside_Traffic_Road_binaural	30 s	L: 74.9 dB(A) R: 73.9 dB(A)	Binaural
Recording at pavement	Outside_Traffic_Crossroads_binaural	20 s	L: 69.1 dB(A) R: 69.6 dB(A)	Binaural
Recording at departure platform	Train_Station_binaural	30 s	L: 68.2 dB(A) R: 69.8 dB(A)	Binaural
Recording at the drivers position	Fullsize_Car1_130Kmh_binaural	30 s	L: 69.1 dB(A) R: 68.1 dB(A)	Binaural
Recording at sales counter	Cafeteria_Noise_binaural	30 s	L: 68.4 dB(A) R: 67.3 dB(A)	Binaural
Recording in a cafeteria	Mensa_binaural	22 s	L: 63.4 dB(A) R: 61.9 dB(A)	Binaural
Recording in business office	Work_Noise_Office_Callcenter_binaural	30 s	L: 56.6 dB(A) R: 57.8 dB(A)	Binaural

Before starting the measurements a proper conditioning sequence shall be used. The conditioning sequence shall be comprised of the four additional sentences 1 to 4 described in [ETSI TS 103 106], applied to the beginning of the 16-sentence test sequence. The conditioning signal level is –1.7 dBPa at the MRP, measured as the active speech level according to [ITU-T P.56].

NOTE – The sequence of speech samples concatenated for the test signal, consisting of alternating talkers in the sending direction, reduces the overall test time but may represent an unrealistic behaviour for certain voice enhancement technologies. Alternative concatenations are for further study.

The send speech signal consists of the 16 sentences of speech as described in [ETSI TS 103 106]. The test signal level is −1.7 dBPa at the MRP, measured as the active speech level according to [ITU-T P.56].

Three signals are required for the tests:

– the clean speech signal is used as the undisturbed reference (see [ETSI TS 103 106]);
– the speech plus undisturbed background noise signal is recorded at the terminal's microphone position using an omnidirectional measurement microphone with a linear frequency response between 50 Hz and 12 kHz;
– the send signal is recorded at the POI.

N-MOS-LQOw, S-MOS-LQOw and G-MOS-LQOw are calculated as described in [ETSI TS 103 106] on a per sentence basis and averaged over all 16 sentences. The results shall be reported as average and standard deviation.

The measurement is repeated for each ambient noise condition described in Table 26, and the average of the results derived from all ambient noise types is calculated.

12 Wideband headset technical requirements

Unless otherwise specified, measurements defined in clause 9 will be carried at the setting of the volume control where the RLR is as close as possible to nominal value.

12.1 Headset sending characteristics

12.1.1 SLR

The nominal values of SLR shall be: SLR = 8 ± 3 dB.

12.1.1.1 Measurement method

Measurement method for a wideband headset is the same as for a wideband handset.

12.1.2 Sending frequency response

The sending sensitivity/frequency characteristics shall be as follows:

The sending sensitivity frequency response, measured from the MRP to the SS audio output (digital output of the reference speech decoder of the SS) with microphone at RTP position, shall be within a mask, which can be drawn between the points given in Table 11. The mask is drawn with straight lines between the breaking points in Table 27 on a logarithmic (frequency) – linear (dB sensitivity) scale.

Table 27 – Sending sensitivity/frequency mask-headset

Frequency (Hz)	Upper limit	Lower limit
100	0	
200	5	–5
5 000	5	–5
6 300	5	–10
8 000	5	
NOTE – All sensitivity values are expressed in dB on an arbitrary scale.		

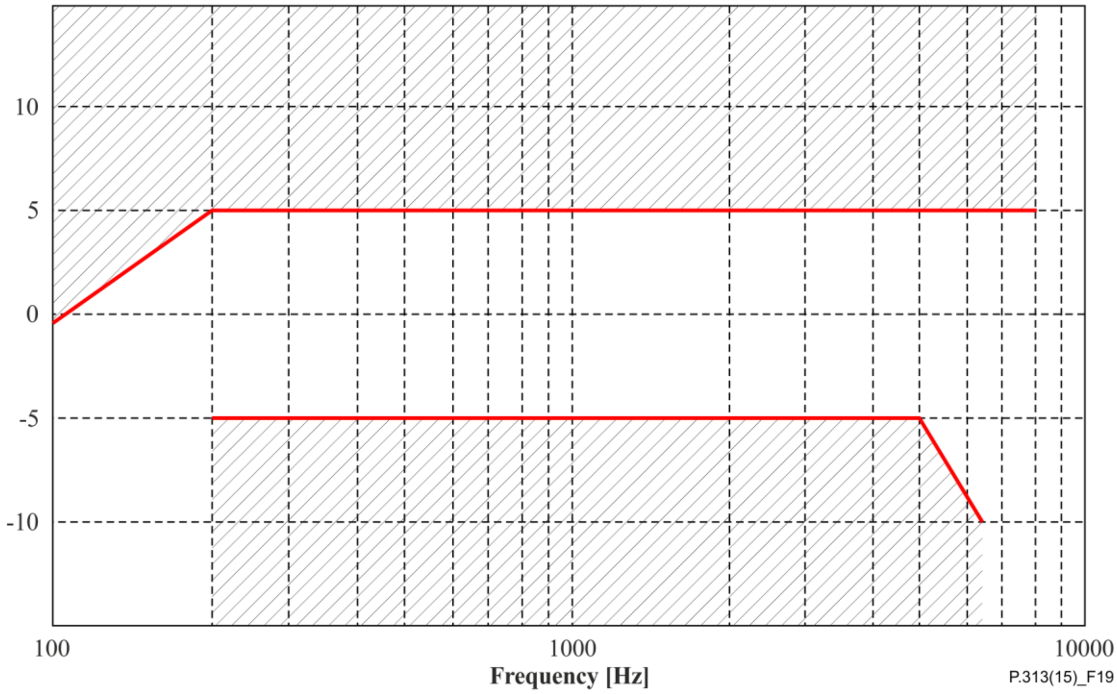

Figure 19 – Sending mask

12.1.2.1 Measurement method

The measurement method for a wideband headset is the same as for a wideband handset, except that the headset should be mounted at the test position according to [ITU-T P.380] using a HATS.

12.1.3 Idle channel noise

The maximum noise level produced by the apparatus at the output of the SS under silent conditions in the sending direction shall not exceed –64 dBm0(A).

NOTE – This figure applies to the total noise level with A-weighting. It is recommended that the level of single frequency disturbances should be ≤ –74 dBm0(A) in the frequency range from 100 Hz to 8 kHz.

12.1.3.1 Measurement method

The headset should be mounted according to [ITU-T P.380] using a HATS in a quiet environment, as defined in clause 7.2, the sending noise level at the SS audio output is measured with A-weighting.

12.1.4 Distortion

The sending part shall meet the following distortion requirements:

The sending distortion shall be measured between MRP and the SS audio output (output of the reference speech decoder of the SS). The ratio of signal-to-total distortion power measured with a psophometric noise weighting described in [ITU-T G.223] shall be above the limits given in Table 28.

Table 28 – Limits for signal-to-total distortion ratio

Frequency	Sending level (dBPa at the MRP)	Sending ratio (dB)
315	–4.7	28
408	–4.7	32
510	–4.7	32
816	–4.7	32
1 020	5	30
	0	35
	–4.7	35
	–10	33
	–15	30
	–20	27

Limits for intermediate levels are found by drawing straight lines between the breaking points in the table on a linear (dB signal level) – linear (dB ratio) scale.

12.1.4.1 Measurement method

The measurement method for a wideband headset is the same as for a wideband handset, except that the headset should be mounted at the test position according to [ITU-T P.380] using a HATS.

12.2 Headset receiving characteristics

12.2.1 RLR

The nominal values of RLR shall be: RLR = 2 ± 3dB.

RLR (binaural headset) = 8 ± 3 dB for each earphone.

Where a user-controlled receiving volume control is provided, the RLR shall meet the nominal value for at least one setting of the control. When the control is set to maximum, the RLR shall not be ≤ (equal or louder than) –13 dB.

With the volume control set to the minimum position, the RLR shall not be ≥ (equal or quieter than) 18 dB and shall not be ≥ (equal or quieter than) 24 dB for a binaural headset.

12.2.1.1 Measurement method

Measurement method for a wideband headset is the same as for a wideband handset.

For binaural earphones, the receiving sensitivity equals the power sum of that measured with the left ear and right ear individually.

12.2.2 Receiving frequency response

The receiving sensitivity/frequency characteristics shall be as follows:

The receiving sensitivity frequency response, measured from the SS audio input (analogue or digital input of the reference speech encoder of the SS) to the DRP, shall be within a mask, which can be drawn with straight lines between the breaking points in Table 13 on a logarithmic (frequency) – linear (dB sensitivity) scale.

For binaural earphones, the left earphone and right earphone should meet the requirement specified in Table 29.

Table 29 – Receiving sensitivity/frequency mask-headset

Frequency (Hz)	Upper limit 8 ± 2 N	Lower limit 8 ± 2 N
100	6	
200	6	–10
300	6	–6
1 000	6	–6
2 000	8	–6
5 000	8	–6
6 300	8	–12
8 000	8	

NOTE 1 – All sensitivity values are expressed in dB on an arbitrary scale.

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

NOTE 3 – The limits in the table above are enforced but are under evaluation. The values are expected to be modified taking into account that the change from ERP to diffuse-field correction is reflected in the table.

12.2.2.1 Measurement method

The measurement method for a wideband headset is the same as for a wideband handset, except that the headset should be mounted at the test position according to [ITU-T P.380] using a HATS.

12.2.3 Idle channel noise

The maximum (acoustic) noise level at the headset UE when no signal is applied to the input of the SS shall be as follows:

– if no user-controlled receiving volume control is provided, or, if it is provided, at the setting of the user-controlled receiving volume control at which the RLR is equal to the nominal value, the noise measured at the DRP with diffuse-field correction contributed by the receiving part alone shall not exceed –57 dBPa(A).

For binaural earphones, each receiver shall not exceed –60 dBPa(A).

– where a volume control is provided, the measured noise shall also not exceed –54 dBPa(A) at the maximum setting of the volume control.

For the nominal volume control setting, the level of single frequency disturbances should be ≤ 60 dBPa(A) in the frequency range from 100 Hz to 10 kHz. As a performance objective it is recommended that the level should be ≤ 64 dBPa(A). For binaural earphones, each receiver shall not exceed –57 dBPa(A).

12.2.3.1 Measurement method

The headset should be mounted at the test position using a HATS as specified in [ITU-T P.380] in a quiet environment, as defined in clause 7.2, the receiving noise level at DRP with diffuse-field correction is measured with A-weighting.

12.2.4 Distortion

The receiving part between the SS audio input (input of the reference speech encoder of the SS) and DRP with diffuse-field correction shall meet the requirements in this clause at the nominal setting of the volume control:

The ratio of signal-to-total distortion shall be above the limits given in Table 30 when the sound pressure at DRP with diffuse-field correction is up to +10 dBPa. For sound pressures exceeding +10 dBPa at the DRP with diffuse-field correction, there is no distortion requirement.

Table 30 – Limits for signal-to-total distortion ratio

Frequency (Hz)	Receiving level at the digital interface (dBm0)	Receiving ratio at nominal volume setting (dB)	Receiving ratio at maximum volume setting (dB)
315	–16	20	–
408	–16	28	–
510	–16	28	–
816	–16	28	–
1 020	0	25.5	–
	–3	31.2	–
	–10	33.5	–
	–16	33.5	–
	–20	33	–
	–30	30.5	–
	–40 (Note)	22.5 (Note)	–
	–45 (Note)	17.5 (Note)	–
NOTE –For levels –40 and –45 dBm0 a lower signal-to-total distortion ratio may not be possible, and hence would not be regarded as a failing result. However, the obtained results would be reported.			

Limits for intermediate levels are found by drawing straight lines between the breaking points in the table on a linear (dB signal level) – linear (dB ratio) scale.

12.2.4.1 Measurement method

The measurement method for a wideband headset is the same as for a wideband handset, except that the headset should be mounted at the test position according to [ITU-T P.380] using a HATS.

12.3 Sidetone characteristics

The talker STMR shall be ≥ 15 dB and shall be ≤ 23 dB for the nominal setting of the volume control. For all other positions of the volume control, the STMR shall be ≥ 10 dB.

In case the STMR is below the lower limit also when the electrical sidetone path has been disabled, the result shall not be regarded as a failure.

NOTE – The STMR calculation algorithm being used was developed for quantifying the audibility of the electrical sidetone path using a sealed coupler. The air-conducted path was not intended to be included in the test set-up. A lower STMR limit was specified to avoid annoying effects (e.g., howling, increase of ambient noise level in the ear) of an excessive electrical sidetone. In HATS-based measurements, the air-conducted path cannot be avoided in the test set-up. With some terminal form factors especially large size terminals the air-conducted path can be substantial resulting in low STMR figures also when there are no annoying effects from any excessive electrical sidetone. See [ITU-T P.76] for definitions of sidetone paths.

12.3.1 Measurement method

The headset should be mounted at test position according to [ITU-T P.380] using a HATS.

A real speech test signal in [ITU-T P.501] with a level of –4.7, dBPa shall be applied at the MRP. For each frequency given in Table 3 of [ITU-T P.79], bands 1 to 20, the sound pressure (at the ERP) shall be measured.

The test set-up shown in Figure 11 is used to measure the sidetone frequency response. The sidetone path loss L_{meST} and the STMR shall be calculated according to [ITU-T P.79], using Equation 4-1 ($m = 0.225$) and the weighting factors in Table 3 of [ITU-T P.79].

12.4 TCLw

The TCLw for a headset UE shall be ≥ 46 dB for any setting of the volume control.

The TCLw for a headset UE should be ≥ 55 dB at the nominal setting of the volume control.

NOTE – It is recommended that the volume control should be set back to nominal after each call unless TCLw ≥ 55 dB can also be maintained with the maximum volume setting. Depending on the UE idle channel noise in the sending direction, it may not always be possible to measure an echo loss ≥ 55 dB.

12.4.1 Measurement method

The headset should be mounted at test position according to [ITU-T P.380] using a HATS.

The attenuation from digital input to digital output is measured at one-twelfth octave frequencies as given by [IEC 61260] for frequency bands from 300 to 6700 Hz, using the measurement arrangement shown in Figure 12.

The test signal used is the compressed real speech signal described in clause 7.3.3 of [ITU-T P.501].

The TCLw is calculated according to clause B.4 (trapezoidal rule) of [ITU-T G.122]. For the calculation, the averaged measured echo level at each frequency band is referred to the averaged test signal level measured in each frequency band. The first 17.0 s of the test signal (6 sentences) are discarded from the analysis to allow for convergence of the acoustic echo canceller. The analysis is performed over the remaining length of the test sequence (last 6 sentences).

The test signal level shall be –10 dBm0. Terminals with adjustable receive levels shall be tested at the nominal setting. For the nominal setting, adjust the level so that the RLR is as close as possible to the nominal RLR value.

12.5 Acoustic safety of telephone user

In order to ensure safety and to minimize annoyance to the user, the terminal shall meet the limits specified in [ITU-T P.360].

12.5.1 Measurement method

The maximum steady state acoustic pressure is measured by applying maximum positive digital code to the receive input defined for the handset under test. The test procedure is in accordance with clause 12.2.2.1 for the receive characteristics except that the acoustic pressure in the artificial ear is measured with a sound level meter.

The measurement shall refer to [ITU-T P.360].

Terminals with adjustable receive levels shall be tested at the maximum setting.

12.6 Delay

In view of the following considerations:

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

– the amount of delay introduced by wireless systems depends on specific technology and may be inherent to the adopted coding technique.

The following is recommended:

– delay added by the terminal equipment should be minimized in accordance with the guidelines provided in [ITU-T G.114] even with the use of echo control;

– the terminal specific implementation dependent processing delay, including both the delay in sending direction and the delay in receiving direction, should be less than 70 ms.

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

12.6.1 Measurement method

The measurement method for a wideband headset is the same as for a wideband handset, except that the headset should be mounted at the test position according to [ITU-T P.380] using a HATS.

13 Wideband speakerphone sets technical requirements

The requirements and test methods in this Recommendation are only for desktop operated and handheld loudspeaker mobile terminals. The transmission characteristics and test methods specified in [ITU-T P.340] and [ITU-T P.342] can be used for the two kinds of speakerphone sets.

Speakerphone sets use SPDA to control acoustic echo, reduce background noise transmission, etc. These parameters are defined in [ITU-T P.330].

Appropriate test signals are described in [ITU-T P.50] and [ITU-T P.501]. Test methods appropriate for parameters defined in this Recommendation may be found in [ITU-T P.502]. Proper use of HATS testing may be found in [ITU-T P.581].

Unless otherwise specified, measurements defined in clause 10 will be carried at the setting of the volume control where the RLR is as close as possible to nominal value.

13.1 Loudspeaker sending characteristics

13.1.1 SLR

According to [ITU-T P.340], the SLR of a loudspeaker telephone should be about 5 dB higher than the SLR of the corresponding handset telephone. Therefore, the nominal value of SLR for handheld and desktop speakerphone sets shall be 13 ± 4 dB.

13.1.1.1 Measurement method

The real speech signal as defined in [ITU-T P.501] shall be used. The test signal used shall be specified in the test report. The test signal level shall be –4.7 dBPa, measured at the MRP. The test signal level is averaged over the complete test signal sequence. The broadband signal level is then adjusted to –28.7 dBPa at the HFRP or the HATS HFRP as defined in [ITU-T P.581] and the spectrum is not altered.

The spectrum at the MRP and the actual level at the MRP (measured in 1/3-octaves) are used as references

The set-up for desktop speakerphone sets is described in clause 6.3.1.

The set-up for handheld speakerphone sets is described in clause 6.3.2.

The SLR shall be calculated according to [ITU-T P.79].

The sending sensitivity shall be calculated from each band of the 20 frequencies given in Table G.1 of [ITU-T P.79], bands 1 to 20. For the calculation, the averaged measured level at the electrical reference point for each frequency band is referred to the averaged test signal level measured in each frequency band at the MRP.

The sensitivity is expressed in terms of dBV/Pa and the SLR shall be calculated according to formula (A 23b) of [ITU-T P.79], over bands 1 to 20, using m = 0.175 and the sending weighting factors from Table A.2 of [ITU-T P.79].

13.1.2 Sending frequency response

In view of the following considerations:

– ideally, the frequency response should be flat for optimal voice quality;

– many manufacturers use microphones and analogue filters to reduce noise;

the sending tolerance mask of handheld and desktop speakerphone sets for the nominal sensitivity/frequency response mask shown in Table 31 is recommended. The mask is drawn as straight lines between the breaking points in the table on a logarithmic (frequency) – linear (dB sensitivity) scale.

Table 31 – Speakerphone sending frequency response

Frequency (Hz)	**Upper limit (dB)**	**Lower limit (dB)**
100	0	
200	5	–5
5 000	5	–5
6 300	5	–10
8 000	5	
NOTE – As stated in [ITU-T P.340], the interval between 200 and 300 Hz makes a significant contribution to the naturalness of the transmitted speech. However, the noise energy in a noisy environment in that frequency range might be significantly higher than the speech energy. Therefore, the manufacturers should consider this trade-off between naturalness of the transmitted speech and background noise transmitted when determining the optimal frequency response in the sending direction.		

13.1.2.1 Measurement methods

The set-up for desktop a speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The real speech signal as described in [ITU-T P.501] shall be used for the test. The type of test signal used shall be stated in the test report. The spectrum of acoustic signal produced by the artificial mouth is calibrated under free-field conditions at the MRP. The test signal level shall be –4.7 dBPa, measured at the MRP. The test signal level is averaged over the complete test signal sequence. The broadband signal level then is adjusted to the required value and the spectrum is not altered.

The spectrum at the MRP and the actual level at the MRP (measured in third octaves) are used as a reference to determine the sending sensitivity S_{mJ}.

Measurements shall be made at 1/3-octave intervals as given by [IEC 61260] for frequency bands from 100 Hz to 8kHz inclusive. For the calculation, the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dBV/Pa.

13.1.3 Idle channel noise

The following limit is recommended:

– send noise level maximum –64 dBm0p.

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

13.1.3.1 Measurement methods

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The environment shall comply with the conditions described in clause 7.2.

The psophometric noise level at the output of the SS is measured. The psophometric filter is described in [ITU-T O.41].

13.1.4 Distortion

The ratio of signal to total distortion shall be above the mask defined in Table 32.

Table 32 – Signal to distortion ratio limit, sending

Sending level (dBPa at the MRP)	Sending ratio (dB)
5	30
0	35
–4.7	35
–10	33
–15	30
–20	27

13.1.4.1 Measurement methods

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The signal used is an activation signal followed by a sine wave signal with a frequency of 1 020Hz which is the actual test signal. The signal is applied at the MRP. Appropriate signals for activation and signal combinations can be found in [ITU-T P.501].

The sine wave signal level shall be calibrated to the following RMS levels at the MRP: 5, 0, –4.7, –10, –15, –20 dBPa. The test signals have to be applied in this sequence, i.e., from high levels down to low levels.

The total distortion power of the signal output of the SS shall be measured with the psophometric noise weighting described in [ITU-T O.41].

13.2 Speakerphone set receiving characteristics

13.2.1 RLR

For desktop speakerphone set, the nominal value of RLR shall be 5 ± 4 dB.

Where a user-controlled volume control is provided, the RLR shall meet the nominal value for at least one setting of the control. It is recommended that a volume control giving at least 15 dB increase from the nominal RLR (louder) is provided for hands-free units. This is to allow for the increased acoustic noise level in the usage environment.

RLR at the maximum volume control setting should be ≤ (equal or louder than) 1 dB.

For handheld speakerphone sets, the nominal value of RLR shall be 9 +9 / –7 dB.

As a performance objective it is recommended that the RLR at the maximum volume control setting is ≤ (equal or louder than) 2 dB.

Where a user-controlled volume control is provided, the RLR shall meet the nominal value for at least one setting of the control. It is recommended that a volume control range ≥ 15 dB be provided.

13.2.1.1 Measurement methods

The real speech signal as defined in [ITU-T P.501] shall be used. The test signal used shall be specified in the test report. The test signal level shall be –16 dBm0 measured at the digital reference point or the equivalent analogue point. The test signal level is averaged over the complete test signal sequence.

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The receiving sensitivity shall be calculated from each band of the 20 frequencies given in Table A.2 of [ITU-T P.79], bands 1 to 20. For the calculation, the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dBPa/V and the RLR shall be calculated according to formula (A-23c), of [ITU-T P.79], over bands 1 to 20, using m = 0.175 and the receiving weighting factors from Table A.2 of [ITU-T P.79].

Free-field correction is used. No leakage correction shall be applied.

13.2.2 Receiving frequency response

In view of the following considerations:

– ideally, the frequency response should be flat for optimal voice quality;

The receiving tolerance mask for handheld speakerphone sets shown in Table 33 and for desktop speakerphone sets shown in Table 34 are recommended. The mask is drawn as straight lines between the breaking points in the table on a logarithmic (frequency) – linear (dB sensitivity) scale.

Table 33 – Receiving frequency response-handheld

Frequency (Hz)	Upper limit	Lower limit
315	6	
630	6	–12
800	6	–6
4 000	6	–6
6 300	6	–12
8 000	6	
NOTE – All sensitivity values are expressed in dB on an arbitrary scale.		

Table 34 – Receiving frequency response-desktop

Frequency (Hz)	Upper limit	Lower limit
125 Hz	8	
200 Hz	8	–12
250 Hz	8	–9
315 Hz	7	–6
400 Hz	6	–6
5 000 Hz	6	–6
6 300 Hz	6	–9
8 000 Hz	6	–□
NOTE – All sensitivity values are expressed in dB on an arbitrary scale.		

13.2.2.1 Measurement method

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The real speech test signal as described in [ITU-T P.501] shall be used for the test. The type of test signal used shall be stated in the test report. The test signal level shall be –16 dBm0, measured at the digital reference point or the equivalent analogue point. The test signal level is averaged over the complete test signal sequence.

If a HATS is used, then it is free-field equalized as described in [ITU-T P.581]. The equalized output signal of each artificial ear is power-averaged over the total duration of the analysis; the right and left artificial ear signals are voltage-summed for each 1/3-octave band frequency band; these 1/3-octave band data are considered as the input signal to be used for calculations or measurements. Measurements shall be made at 1/3-octave intervals as given by [IEC 61260] for frequency bands from 100 Hz to 8 kHz inclusive. For the calculation, the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dBPa/V.

13.2.3 Idle channel noise

The following limit is recommended at the nominal RLR setting:

– receive noise level maximum –53 dBPa(A).

Spectral peaks shall be less than 10 dB above the average noise spectrum. A narrowband FFT analysis with a frequency resolution of 32 Hz or less is recommended.

The noise produced by the speakerphone terminal providing volume control shall be less than −53 dBPa(A) with the volume control set to its maximum level (minimum RLR setting).

13.2.3.1 Measurement method

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The environment shall comply with the conditions described in clause 7.2.

Under quiet conditions, as defined in clause 7.2, the receiving noise level at DRP with free-field correction is measured with A-weighting.

13.2.4 Distortion

The receiving part between the SS audio input (input of the reference speech encoder of the SS) and the DRP with free-field correction shall meet the requirements in this subclause at the nominal setting of the volume control:

The ratio of signal to total distortion power measured shall be above the limits given in Table 35 when the sound pressure at the DRP with free-field correction is up to 10 dBPa. For a sound pressure ≥ 10 dBPa at the DRP with free-field correction there is no distortion requirement.

Table 35 – Signal to distortion ratio limit, receiving

Frequency (Hz)	Receiving level at the digital interface (dBm0)	Receiving ratio at nominal volume setting (dB)	Receiving ratio at maximum volume setting (dB)
1 020	0	25.5	–
	–3	31.2	–
	–10	33.5	–
	–16	33.5	–
	–20	33	–
	–30	30.5	–
	–40 (Note)	22.5 (Note)	–
	–45 (Note)	17.5 (Note)	–
NOTE –For levels –40 and –45 dBm0 a lower signal-to-total distortion ratio may not be possible, and hence would not be regarded as a failing result. However, the obtained results would be reported.			

13.2.4.1 Measurement method

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The signal used is an activation signal followed by a series of sine wave signals with a frequency of 1 020 Hz, which is the actual test signal. Appropriate signals for activation and signal combinations can be found in [ITU-T P.501]. The signal level shall be calibrated to –16 dBm0.

The signal level shall be applied at the signal input of the SS at the following levels: 0, 3, –10, –16, –20, –30, –40, –45 dBm0. The test signals have to be applied in this sequence, i.e., from high levels down to low levels.

The total distortion power shall be measured at the DRP with free-field correction and calculated with the psophometric noise weighting as specified in [ITU-T O.41].

13.3 TCLw

In order to meet the ITU-T G.131 talker echo objective requirements, the TCLwst should be greater than 46 dB.

For terminals fitted with a volume control, the TCLwst shall be not less than 40 dB for any setting of the volume control.

The limits for TCLwdt are defined in clause 13.5.

13.3.1 Measurement method

The speakerphone is set up in a room with acoustic properties similar to a typical "office-type" room, and the ambient noise level should ≤ 70 dBPa(A).

The set-up for a desktop speakerphone set is described in clause 6.3.1.

The set-up for a handheld speakerphone set is described in clause 6.3.2.

The attenuation from reference point input to reference point output shall be measured using the compressed real speech signal described in clause 7.3.3 of [ITU-T P.501]. The test signal level shall be –10 dBm0.

The TCLw is calculated according to clause B.4 (trapezoidal rule) of [ITU-T G.122], using the frequency range of 300 to 6 700 Hz. For the calculation, the averaged measured echo level at each frequency band is referred to the averaged test signal level measured in each frequency band. The first 17.0 s of the test signal (6 sentences) are discarded from the analysis to allow for convergence of the acoustic echo canceller. The analysis is performed over the remaining length of the test sequence (last 6 sentences).

13.4 Switching characteristics

Many mobile speakerphone terminals use a VAD and a loss controller to control acoustic echo. The VAD distinguishes silent periods (no active speech signal), single-talk periods (near-end speech periods or far-end speech periods), and double-talk periods (near-end and far-end speech signals active at the same time). According to the state determined by the VAD, the loss controller reduces the acoustic echo level by inserting variable losses on the received and/or transmitted audio signals. (The VAD and loss controller may also be used to control electrical echo from the network if the mobile terminal operates on an analogue wireless network without echo control.) An AEC and/or NLP may also be used to control acoustic echo. However, most devices with an AEC and/or NLP also require a loss controller to control residual acoustic echo. Although no VAD is perfect and instantaneous, errors by the VAD may lead to speech clipping and acoustic echo.

The switching parameters and switching functions are described in detail in [ITU-T P.330] and [ITU-T P.340].

The minimum activation level in the sending direction shall be ≤ –20 dBPa. The build-up time for activation in the sending direction shall be:

$$TRst\text{-}s \leq 50 \text{ ms}$$

The minimum activation level in the receiving direction shall be ≤ –35.7 dBm0. The build-up time for activation in the receiving direction shall be:

$$TRst\text{-}r \leq 50 \text{ ms}$$

Requirements for the attenuation and build-up time for receive-to-send states (and send-to-receive states) are for further study.

13.4.1 Measurement method

The set-up for a desktop speakerphone set is described in [ITU-T P.581].

The set-up for a handheld speakerphone set is described in Figure 3.

The TRst-s and TRst-r shall be measured according to [ITU-T P.340].

The echo loss measurement during double talk should be carried out according to [ITU-T P.502].

13.5 Double talk performance

The speech quality during double talk is mainly determined by two parameters: TELR (impairment caused by echo during double talk, and related to TCLwdt), and level variation between single and double talk (attenuation range).

In order to guarantee sufficient quality under double talk conditions, the TELR should be high and the attenuation inserted should be as low as possible. Terminals that do not allow double talk should provide a good echo attenuation realized by high attenuation.

The most important parameters to determine the quality during double talk are as follows:

– Asdt;

– Ardt;

– echo attenuation during double talk.

The requirements for Asdt and Ardt at nominal RLR for each category of speakerphone terminals are provided in Table 4 of [ITU-T P.340]. Higher values for Asdt and Ardt are permitted at lower RLR (due to increasing the received volume level) without affecting the behaviour characteristics. However, the TELRdt must maintain the requirements in [ITU-T P.340] for all received volume settings.

Subjective evaluation methods for double talk quality, as well as single talk quality, are described in [ITU-T P.832].

NOTE – Acoustic echo control is more easily and effectively done in terminals. However, network equipment, as described in [ITU-T G.168], may also include acoustic echo control processing. This leads to tandeming issues of acoustic echo control functionalities. As specified in [ITU-T G.168], the added network component shall prevent any degradation of the overall perceived quality. In practice, however, double-talk capability may be reduced due to network equipment.

13.5.1 Measurement method

The set-up for a desktop speakerphone set is described in [ITU-T P.581].

The set-up for a handheld speakerphone set is described in Figure 3.

The attenuation range in sending and receiving direction during double talk, Asdt and Ardt, shall be measured according to [ITU-T P.340].

13.6 Background noise transmission and comfort noise injection

Mobile speakerphone terminals are typically used in noisy environments.

Most speakerphone terminals have some type of NR capability. The main purpose of a NR system in a device is to reduce the annoying and fatiguing effects of the transmitted background noise. The techniques used to reduce background noise may be classified as analogue only, digital only and combined analogue and digital techniques. These techniques are described in [ITU-T P.330].

Some parameters describing NR and some requirements are also given in [ITU-T P.330]. There is usually a trade-off between the level of digital NR and the voice quality. Too much NR may reduce the transmitted voice level and distort the speech signal.

Without comfort noise, the acoustic echo control system (loss controller, VAD, NLP and AEC) will reduce and/or distort the background noise transmitted during receive and double talk states. This can be annoying to the far-end talker. Some speakerphone terminals inject comfort noise in order to maintain a constant level of background noise transmitted. Some parameters for comfort noise are described in [ITU-T P.330].

The level of comfort noise should be within a range of +2 and –5 dB from the original (transmitted) background noise.

13.6.1 Measurement methods

For further study.

SERIES OF ITU-T RECOMMENDATIONS

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

Printed in Switzerland
Geneva, 2015