



LTE;
Test plan for speech quality and delay through
a headset electrical interface
(3GPP TR 26.954 version 14.0.0 Release 14)



Reference

DTR/TSGS-0426954ve00

Keywords

LTE

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:

<http://www.etsi.org/standards-search>

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (PDF) version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status.

Information on the current status of this and other ETSI documents is available at

<https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx>

If you find errors in the present document, please send your comment to one of the following services:

<https://portal.etsi.org/People/CommiteeSupportStaff.aspx>

Copyright Notification

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.

The copyright and the foregoing restriction extend to reproduction in all media.

© ETSI 2017.

All rights reserved.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are Trade Marks of ETSI registered for the benefit of its Members.

3GPP™ and **LTE™** are Trade Marks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners.

oneM2M logo is protected for the benefit of its Members

GSM® and the GSM logo are Trade Marks registered and owned by the GSM Association.

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: "*Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards*", which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (<https://ipr.etsi.org>).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Report (TR) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities, UMTS identities or GSM identities. These should be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between GSM, UMTS, 3GPP and ETSI identities can be found under <http://webapp.etsi.org/key/queryform.asp>.

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the [ETSI Drafting Rules](#) (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Contents

Intellectual Property Rights	2
Foreword.....	2
Modal verbs terminology.....	2
Foreword.....	4
Introduction	4
1 Scope	5
2 References	5
3 Definitions, symbols and abbreviations	5
3.1 Definitions	5
3.2 Symbols.....	6
3.3 Abbreviations	6
4 Test Configurations	6
4.1 Setup for Terminals	6
4.2 Setup of the electrical interfaces.....	7
4.2.1 Codec approach and specification.....	7
4.2.2 Direct digital processing approach.....	8
4.3 Accuracy of test equipment	9
4.4 Test signals.....	9
5 Test method for delay and speech quality	10
5.1 Delay in sending + receiving direction using "echo" method.....	10
5.2 Speech enhancement processing in the UE and its potential impact on the determination of MOS-LQO scores.....	11
5.3 Speech quality loss in sending direction.....	13
5.4 Speech quality loss in receiving direction	14
6 Test Scenarios	15
6.0 General	15
6.1 Scenario 1: Impairments at the physical layer	15
6.2 Scenario 2: Impairments at the internet layer (IP impairments).....	16
6.3 Scenario 3: Operation with different modes / bit-rates (reference)	16
6.4 Scenario 4: Switching bit-rates / operating modes during a call	17
6.5 Scenario 5: Codec change initiated by remote UE	17
6.6 Scenario 6: Intra-RAT Handover	18
6.7 Scenario 7: Inter-RAT Handover	18
6.8 Scenario 8: Operation with DRX	18
6.9 Scenario 9: Voice call with simultaneous data upload/download	19
Annex A: Test Method for Interruption Time	20
Annex B: Change history	21
History	22

Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

- x the first digit:
 - 1 presented to TSG for information;
 - 2 presented to TSG for approval;
 - 3 or greater indicates TSG approved document under change control.
- y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- z the third digit is incremented when editorial only changes have been incorporated in the document.

Introduction

Testing of speech transmission quality and delay through the UE headset electrical interface is a common practice in the industry for troubleshooting and benchmarking of network and UE performance. The UE headset electrical interface provides a convenient access point for testing that removes acoustic interface related aspects from the measurement. Test equipment (TE) solutions designed to connect to the UE headset electrical interface are commercially available and target both "drive-test" as well as lab environment testing applications. However, the lack of a standardized test methodology causes difficulties in interpreting results between different TEs, leading to possibly misleading conclusions on both network and UE performance. The present document attempts to address this problem by providing a unified test procedure methodology.

1 Scope

The present document provides guidance on testing of speech quality and delay through a UE headset electrical interface, including both analog and digital headset interfaces. The present document includes test set-up configurations, measurement scenarios and an indication of a range of results that can be expected in measurements. It is envisioned that the present document will serve as a reference for UE vendors and Mobile Network Operators wishing to conduct such tests.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

- [1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".
- [2] 3GPP TS 26.445: "Codec for Enhanced Voice Services (EVS); Detailed algorithmic description".
- [3] IEC Publication 61672-1 (2013): "Sound Level Meters - Part 1: Specifications".
- [4] ITU-T Recommendation P.501 (01/2012): "Test signals for use in telephony".
- [5] ITU-T Recommendation P.56 (12/2011): "Objective measurement of active speech level".
- [6] 3GPP TS 26.131: "Terminal acoustic characteristics for telephony; Requirements".
- [7] 3GPP TS 26.132: "Speech and video telephony terminal acoustic test specification".
- [8] ITU-T Recommendation P.381 (07/2016): "Technical requirements and test methods for the universal wired headset or headphone interface of digital mobile terminals".
- [9] ITU-T Recommendation P.863 (09/2014): "Perceptual objective listening quality assessment".
- [10] ITU-T Recommendation P.501 Amendment 3 (06/2015): "New Annex D – Speech files with male/female sentences prepared for use with perceptual based objective speech quality prediction".
- [11] 3GPP TS 36.101: "Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception."

3 Definitions, symbols and abbreviations

3.1 Definitions

For the purposes of the present document, the terms and definitions given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1].

3.2 Symbols

For the purposes of the present document, the following symbols apply:

Ω Ohms

3.3 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

EVS	Enhanced Voice Services
JBM	Jitter Buffer Management
MOS-LQO	Mean Opinion Score – Listening Quality Objective
MTSI	Multimedia Telephony Service for IMS
SPS	Semi-Persistent Scheduling
SRVCC	Single Radio Voice Call Continuity

4 Test Configurations

4.1 Setup for Terminals

The general test setup configuration and access to the UE is described in Figure 1:

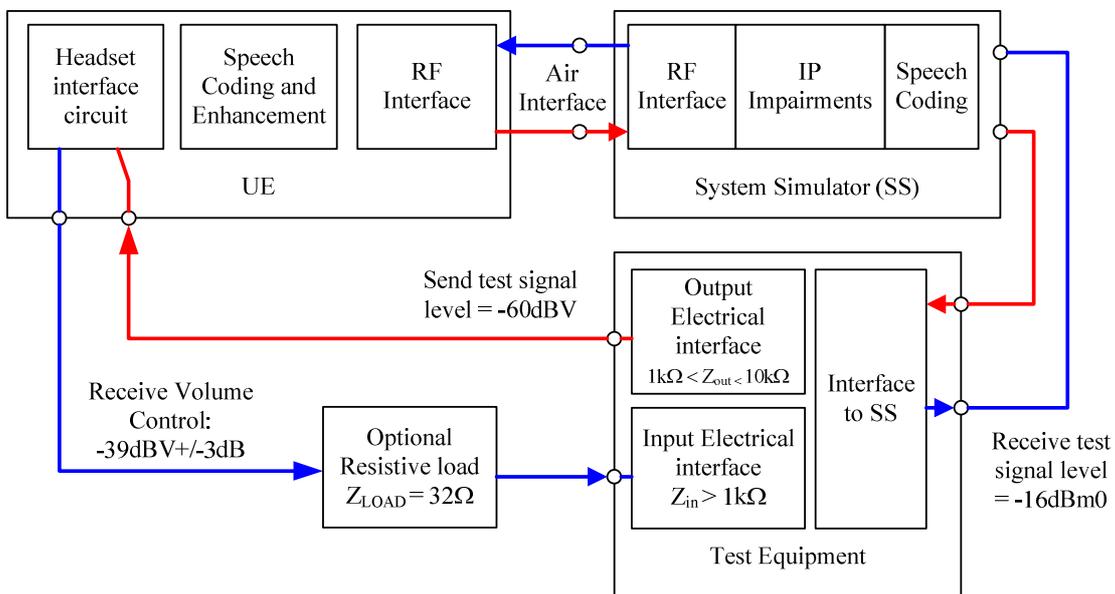


Figure 1: Test Configuration for headset electrical interface

NOTE 1: The nominal Send test signal level is -60dBV. If other send test signal levels are used they will be stated in the test report.

The preferred way of testing is the connection of a UE to the system simulator with exact defined settings and access points. The test sequences are fed in either electrically using a reference codec or using the direct signal processing approach.

The system simulator should simulate the access network and core network including the speech encoding/decoding specified for the test (e.g. AMR, AMR-WB or EVS) but excluding further transcoding beyond linear PCM, see Figure 1.

The output of the test system connected to the interface in Send of the headset connector has to be DC resistant. The output impedance should be between 1 k Ω and 10 k Ω . Care should be taken that the output impedance triggers the UE headset microphone detection. The dynamic range of the test equipment should be consistent with (or exceed) the dynamic range provided by headset microphones.

The input of the test system connected to the receiving interfaces of the headset connectors should have an input impedance greater than 1 k Ω . The dynamic range of the test equipment should be consistent with (or exceed) the dynamic range provided by the UE headset electrical interface output.

An optional 32 Ω load may be used in order to condition the UE with a realistic echo situation and/or to provide a well-defined volume control setting. If the optional 32 Ω load is used, the common ground impedance for the test system (between sending and receiving sides and including the cable used to connect the test equipment to the UE headset electrical interface) should be ≤ 50 m Ω .

NOTE 2: A common ground impedance ≤ 50 m Ω provides a crosstalk attenuation greater than 56 dB between the UE output and input headset electrical interface when using a 32 Ω load. Unless specified otherwise for the respective test, the radio conditions on the air interface will have a block error rate of 0 % and the jitter in the IP transport for MTSI-based speech will be ≤ 1 ms.

In case of MTSI-based speech, the reference client will allow to synchronize to the clock of the device under test and include a de-jitter buffer to equalize possible jitter in the signal received from the UE.

When operating with synchronized clock, the de-jitter buffer will be a static de-jitter buffer and the jitter buffer management will not compensate for clock drift. The reference client will not lose or discard packets, will not trigger retransmission, and will not use error concealment or time-warping. The initial jitter buffer size (filling level) will be higher than the maximum expected network jitter and the maximum jitter buffer size will be at least twice the initial size. During jitter buffer reset, the de-jitter buffer will be emptied/filled to the initial buffer size. In case of buffer over- or underruns, the reference client will give a warning and it will be reported.

NOTE 3: A static de-jitter buffer is a first-in-first-out (FIFO) buffer which at the beginning buffers packets until a given initial buffer size is reached. Due to changing network delays the filling level of the de-jitter buffer can change, but the sum of network delay and jitter buffer delay is constant (as opposed to an adaptive jitter buffer management). The filling level of the de-jitter buffer represents the de-jitter buffer delay.

For measurements with unsynchronized clock e.g. the measurement of clock drift, jitter buffer over- and underruns cannot be avoided due to the unsynchronized clocks. Under the assumption of jitter-free condition the initial jitter buffer size (filling level) will be chosen such that the maximum clock drift can be compensated without any loss of packets for a given time. For the measurement of clock drifts the jitter buffer size should be chosen such that for clock drift of up to 100 ppm no loss of packets due to buffer over- or under-run will occur for a sequence of 160 s.

Unless stated otherwise, for LTE connections, the system simulator will be configured for FDD operation, with a default or dedicated bearer and reference measurement channel scheduling that provides enough resource block allocation for transmitting a full speech packet within a transmission time interval of 1ms. No HARQ re-transmissions will occur. TDD operation, TTI bundling, connected DRX and other forms of scheduling (e.g. SPS) are for further study.

The test setup has to ensure proper clock synchronization of the test equipment to the UE. Clock drift will be negligible and packet loss will not occur during the test.

NOTE 4: Any clock drift may result in improper delay calculation or in wrong positioning of the analysis window.

4.2 Setup of the electrical interfaces

4.2.1 Codec approach and specification

In this approach, a codec is used to convert the digital input/output bit-stream of the system simulator to the equivalent analogue values. With this approach a system simulator simulating the radio link to the terminal under controlled and error-free conditions is required, unless otherwise specified for the respective test. The system simulator has to be equipped with a high-quality codec with characteristics as close as possible to ideal.

Definition of 0 dB_r point:

D/A converter - a Digital Test Sequence (DTS) representing the codec equivalent of an analogue sinusoidal signal with an RMS value of 3,14 dB below the maximum full-load capacity of the codec will generate 0 dBm across a 600 Ω load;

A/D converter - a 0 dBm signal generated from a 600 Ω source will give the digital test sequence (DTS) representing the codec equivalent of an analogue sinusoidal signal with an RMS value of 3,14 dB below the maximum full-load capacity of the codec.

Narrowband telephony testing

For testing of a GSM, 3G or LTE terminal supporting narrowband telephony, the system simulator will use the AMR speech codec as defined in the 3GPP TS 26 series of specifications, at the source coding bit-rate of 12,2 kbit/s, unless otherwise specified.

Wideband telephony testing

For testing of a GSM, 3G or LTE terminal supporting wideband telephony, the system simulator will use the AMR-WB speech codec as defined in 3GPP TS 26 series of specifications, at the source coding bit-rate of 12,65 kbit/s, unless otherwise specified.

Super-wideband telephony testing

For testing of a LTE terminal supporting super-wideband telephony, the system simulator will use the EVS speech codec as defined in 3GPP TS 26 series of specifications at 32 kHz sampling rate, mono, at the source coding bit-rate of 24,4 kbit/s, unless otherwise specified. The system simulator should support all operation modes of EVS speech codec, the complete EVS RTP payload format, and all SDP attributes as specified in TS 26.445 Annex A [2]; it should be possible to control all the SDP attributes declared and negotiated by the system simulator.

Fullband telephony testing

For testing of a LTE terminal supporting fullband telephony, the system simulator will use the EVS speech codec as defined in 3GPP TS 26 series of specifications at 48 kHz sampling rate, mono, at the source coding bit-rate of 24,4 kbit/s, unless otherwise specified.

4.2.2 Direct digital processing approach

In this approach, the digital input/output bit-stream of the terminal connected through the radio link to the system simulator is operated upon directly.

Narrowband telephony testing

For testing of a GSM, 3G or LTE terminal supporting narrowband telephony, the system simulator will use the AMR speech codec as defined in the 3GPP TS 26 series of specifications, at the source coding bit-rate of 12,2 kbit/s, unless otherwise specified.

Wideband telephony testing

For testing of a GSM, 3G or LTE terminal supporting wideband telephony, the system simulator will use the AMR-WB speech codec as defined in the 3GPP TS 26 series of specifications, at the source coding bit rate of 12,65 kbit/s, unless otherwise specified.

Super-wideband telephony testing

For testing of a LTE terminal supporting super-wideband telephony, the system simulator will use the EVS speech codec as defined in 3GPP TS 26 series of specifications in super-wideband mode, at 32 kHz sampling rate, mono, at the source coding bit-rate of 24,4 kbit/s, unless otherwise specified. The system simulator should support all operation modes of EVS speech codec, the complete EVS RTP payload format, and all SDP attributes as specified in TS 26.445 Annex A [2]; it should be possible to control all the SDP attributes declared and negotiated by the system simulator.

Fullband telephony testing

For testing of a LTE terminal supporting fullband telephony, the system simulator will use the EVS speech codec as defined in 3GPP TS 26 series of specifications in fullband mode, at 48 kHz sampling rate, mono, at the source coding bit-rate of 24,4 kbit/s.

4.3 Accuracy of test equipment

Unless specified otherwise, the accuracy of measurements made by test equipment will exceed the requirements defined in table 1a.

Table 1a: Test equipment measurement accuracy

Item	Accuracy
Electrical Signal Power	$\pm 0,2$ dB for levels ≥ -50 dBm
	$\pm 0,4$ dB for levels < -50 dBm
Time	± 5 %
Frequency	$\pm 0,2$ %

Unless specified otherwise, the accuracy of the signals generated by the test equipment will exceed the requirements defined in table 1b.

Table 1b: Test equipment signal generation accuracy

Item	Accuracy
Electrical excitation levels	$\pm 0,4$ dB (see note 1)
Frequency generation	± 2 % (see note 2)
NOTE 1: Across the whole frequency range.	
NOTE 2: When measuring sampled systems, it is advisable to avoid measuring at sub-multiples of the sampling frequency. There is a tolerance of ± 2 % on the generated frequencies, which may be used to avoid this problem, except for 4 kHz where only the -2 % tolerance may be used.	

The measurements' results will be corrected for the measured deviations from the nominal level.

The sound level measurement equipment will conform to IEC 61672-1 Type 1 [3].

4.4 Test signals

Unless stated otherwise, appropriate test signals for GSM/3G/LTE acoustic tests are generally described and defined in ITU-T Recommendation P.501 [4].

More information can be found in the test procedures described below.

For testing the narrowband telephony service provided by the UE, the test signal used will be band limited between 100 Hz and 4 kHz with a bandpass filter providing a minimum of 24 dB/oct. filter roll-off, when feeding into the receiving direction.

For testing the wideband telephony service provided by the UE, the test signal used will be band limited between 100 Hz and 8 kHz with a bandpass filter providing a minimum of 24 dB/oct. filter roll-off, when feeding into the receiving direction.

For testing the super-wideband telephony service provided by the UE, the test signal used will be band limited between 50 Hz and 16 kHz with a bandpass filter providing a minimum of 24 dB/oct. filter roll-off, when feeding into the receiving direction.

For testing the fullband telephony service provided by the UE, the test signal used will be band limited between 20 Hz and 20 kHz with a bandpass filter providing a minimum of 24 dB/oct. filter roll-off, when feeding into the receiving direction.

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 Recommendation P.56 [5] active speech level of the (band limited in receiving direction) test signal, calculated over the complete test sequence.

5 Test method for delay and speech quality

5.1 Delay in sending + receiving direction using "echo" method

The UE delay in loopback mode is obtained by measuring the delay between the output electrical interface of the test equipment (point A in figure 2) and the input electrical interface of the test equipment (point B in figure 2) and subtracting the delays introduced by the test equipment from the measured value.

NOTE: The headset interface circuit depicted assumes the UE includes an analog audio interface. For UEs providing a digital audio interface only, the test should be conducted with a digital to analog and/or analog to digital converters provided by the UE vendor.

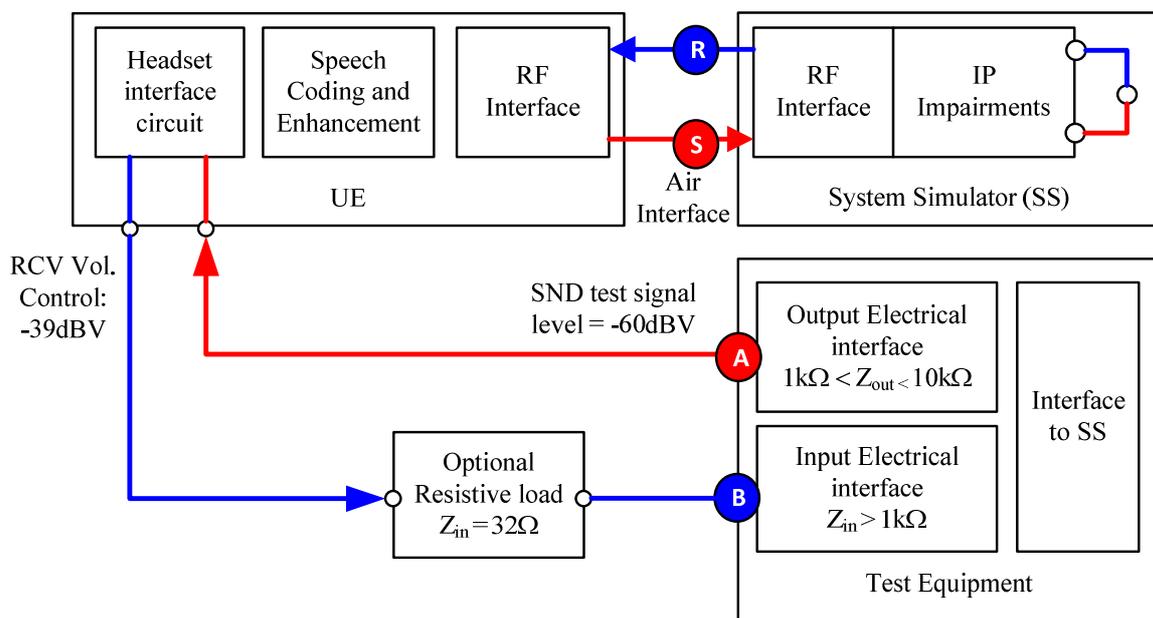


Figure 2: Different entities when measuring the delay in echo method

The delay measured from A to B is ($T_S + T_R + T_{SS}$).

T_{SS} : The delay between the last bit of a speech frame at the system simulator antenna and the first bit of the looped back speech frame at the system simulator antenna (i.e. the delay from point S to point R in the figure 2).

T_S : The UE delay in the sending direction (i.e. the delay from point A to point S in the figure 2).

T_R : The UE delay in the receiving direction (i.e. the delay from point R to point B in the figure 2).

For the measurements a Composite Source Signal (CSS) according to ITU-T Recommendation P.501 [4] is used. It is recommended to use a pseudo-noise sequence of 32k samples (with 48 kHz sampling rate). The test signal level is -60 dBV at the Output Electrical Interface of the test equipment.

The system simulator is configured for "loopback" or "echo" operation with the additional loopback delay as specified below when applicable. In "loopback" or "echo" operation, the packets in the sending direction are routed to the receiving direction by the system simulator.

The reference signal is the original signal (test signal). The setup of the UE is in correspondence to clause 4.1.

The $T_S+T_R+T_{SS}$ delay is determined by cross-correlation analysis between the measured signal at the input electrical interface of the test equipment and the original signal. The analysis window for the cross-correlation will start at an instant $T > 50$ ms in order to discard the cross-correlation peaks corresponding to possible coupling at the electrical headset connector and delayed sidetone signal. The measurement is corrected by subtracting the system simulator delay T_{SS} to obtain the $T_S + T_R$ delay.

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

To account for the possible effect of packet arrival time variations, the measurement is performed [10 to 50] times, depending on the amount of jitter and packet loss in the test case, and the average and 95 % confidence interval are reported. The aim is to achieve a target confidence interval of +/- [0.1] MOS.

For MTSI-based speech with LTE access, a variability of the UE delay with up to 20 ms in the respective sending and receiving direction may be expected due to the synchronization of the speech frame processing in the UE to the bits of the speech frame on the UE antenna. This synchronization is attributed to the UE delay according to the definition of the UE delay reference points. Hence, for test scenarios with no introduced impairments, the UE delay will be reported as the maximum value from at least 5 separate calls each with a different loopback delay T_{SS} in at least 5 steps of 4 ms in the full range from $0+T_{SS}$ to $16+T_{SS}$ ms. All values will be reported in the test report.

5.2 Speech enhancement processing in the UE and its potential impact on the determination of MOS-LQO scores

When tests are conducted over an electrical interface intended for headsets, some headset-related speech enhancement processing is likely active in the UE, unless it is intentionally disabled as described below. Certain speech enhancement is designed to compensate for acoustic aspects related to the headset (e.g. acoustic path losses, strong variations in signal level, etc.). Such speech enhancement is not optimized for conditions where the headset acoustics are absent, such as those found in this test plan (see example in figure 3). Furthermore, this speech enhancement may impact the MOS-LQO scores obtained in this test plan, making it difficult to study non-acoustic related speech quality aspects of interest (e.g. degradations due to JBM, speech coding and radio implementations).

It is therefore recommended to disable the headset-related speech enhancement processing when conducting the objective speech quality assessments described in this test plan (evaluation of non-acoustic related aspects, see table 2, right-most column). Typically only gain settings, the speech codec and the JBM would be enabled. See also table 2.

NOTE: The procedure for disabling the headset-related speech enhancement processing is not standardized and may not be possible for all labs and all UEs. If tests are conducted with speech enhancement enabled, special care should be taken in interpreting the results, for the reasons mentioned above.

Table 2: Examples of test scenarios and recommendations on speech enhancement processing enablement

	Speech quality parameters with regards to various acoustic scenarios	JBM behaviour (changes of delay/ MOS-LQO under varying jitter/loss)	Dedicated studies on aspects that are not acoustics-related (JBM, codec implementation, radio implementation etc.)
Procedure	TS 26.131 [6]/132 [7] (acoustic interfacing)	TS 26.131 [6]/132 [7] is normally used (acoustic interfacing)	Acoustic interfacing is not of interest - the methods in the present document can be more efficient (electrical interfacing)
Speech enhancement processing	The UE is tested "as is" with acoustic input/output, which means headset processing is likely activated	The UE is tested "as is" with acoustic input/output, which means headset processing is likely activated	<i>Where possible it is recommended to disable¹ processing related to acoustic frontends to avoid misleading and blurred data (impact on MOS-LQO)</i>
NOTE: This is expected to be adapted to the test scenario. For instance, in case the test scenario includes double-talk, it may be beneficial to enable echo control processing.			

In figure 3, one aspect of testing via the electrical path is demonstrated. It can be seen that the frequency response from the talker's mouth to a typical headset microphone location might be compensated in the UE. Such compensation is even specified in ITU-T Recommendation P.381 [8]. If such processing is not disabled when measuring with an electrical interface, the MOS-LQO score is likely reduced to a lower value which is not representative and may limit the ability to study the scenarios listed in following clauses, in detail.

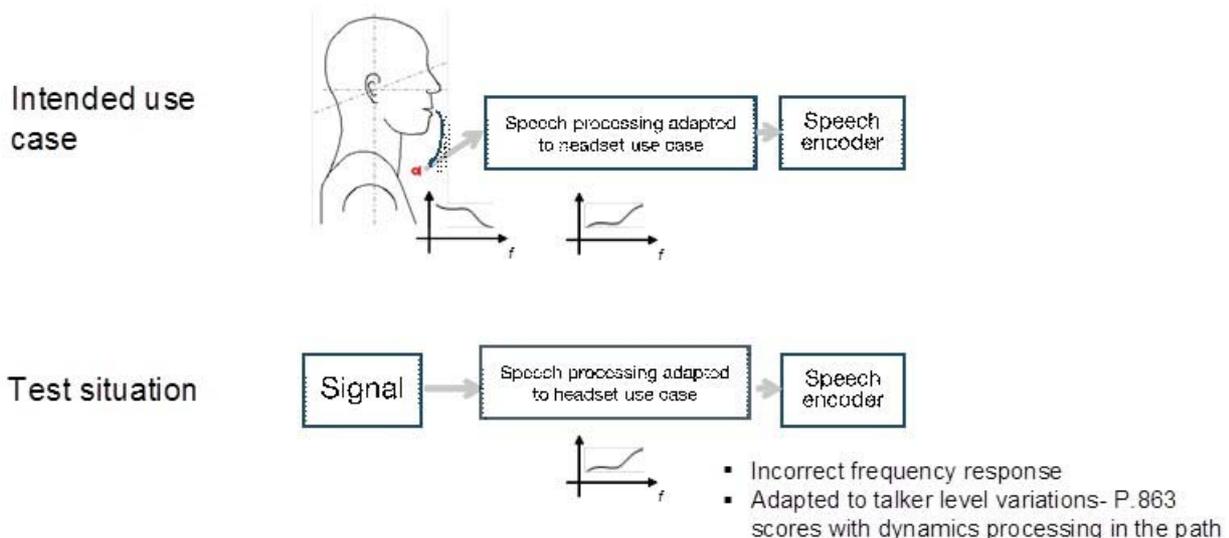
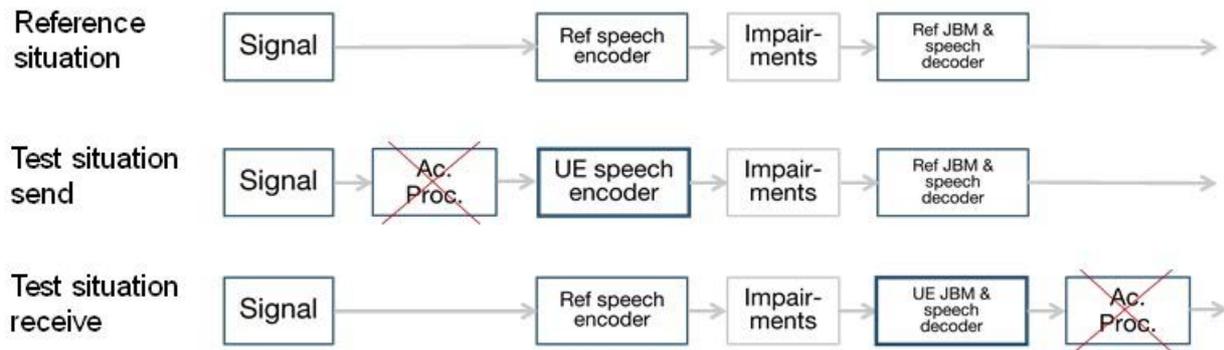


Figure 3: ITU-T Recommendation P.863 [9] applied with signal injection at a headset interface – speech processing intended to be used with an acoustic interface will impact the MOS-LQO values and limit the sensitivity to other degradations of interest

The recommended approach (speech enhancements disabled) is shown in figure 4.



- Can compare test results to reference situation
- Effects of transmission impairments etc can be properly examined

Figure 4: ITU-T recommended ITU-T P.863 [9] usage for assessing UE codec, JBM, effect of transmission radio impairments etc. No speech enhancement processing is enabled. The UE implementation can be directly compared to the reference situation (offline simulation).

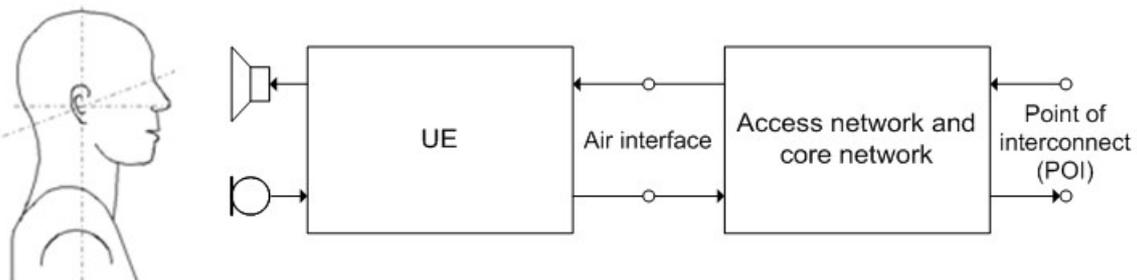


Figure 5: Recommended tests for overall speech quality testing per TS 26.131 [6] and TS 26.132 [7] – include acoustic interfaces and all speech processing – but is less suitable for detailed studies of codec, JBM and effect of transmission impairments

5.3 Speech quality loss in sending direction

For the evaluation of speech quality loss under a given test condition, a fullband speech sentence pair from ITU-T Recommendation P.501 Annex D [10] should be used. P.863 scores are dependent on the speech material used for testing. For consistency in comparing scores across different laboratories, use of a unique sentence pair is recommended. The American English samples from Annex D are an appropriate choice for this purpose. For the P.863 reference input, the fullband test signal will be low-pass filtered at 14 kHz as described in P.863 sub-clause 8.1.1.

NOTE: See also ITU-T Recommendation P.501 Annex D.1 for a rationale on sample selection.

Two recordings are used to produce the speech quality loss metric:

- 1) A recording obtained in jitter and error free conditions (reference condition);
- 2) A recording obtained for the test condition of interest (test condition).

The speech quality of the signal at the output of the speech decoder of the system simulator is estimated using the measurement algorithm described in ITU-T Recommendation P.863 [9]. Certain speech enhancement, as described in clause 5.2, should be disabled.

The measurement will be repeated 20 times for the reference condition and n times for the test condition (as established in clause 6 for each test scenario). For both the reference and the test condition, the first measurement is discarded for convergence of the JBM. A score will be computed for each 6 s speech sentence pair and averaged to produce a mean MOS-LQO value for both the reference and test conditions.

$$\overline{MOS - LQO}_{snd_ref} = \left(\sum_{i=1}^{19} MOS - LQO_{snd_ref}(i) \right) / 19$$

$$\overline{MOS - LQO}_{snd_test} = \left(\sum_{i=1}^{n-1} MOS - LQO_{snd_test}(i) \right) / n$$

where: i is the index of the sentence pair measurement;
 n is the total number of sentence pairs measured;
 $MOS - LQO_{snd_ref}(i)$ is the i_{th} measurement of the reference condition MOS-LQO in sending direction;
 $MOS - LQO_{snd_test}(i)$ is the i_{th} measurement of the test condition MOS-LQO in sending direction.

The mean speech quality loss in sending direction, $\overline{\Delta MOS - LQO}_{SND}$, is determined from:

$$\overline{\Delta MOS - LQO}_{SND} = \overline{MOS - LQO}_{snd_ref} - \overline{MOS - LQO}_{snd_test}$$

5.4 Speech quality loss in receiving direction

For the evaluation of speech quality loss under a given test condition, a fullband speech sentence pair from ITU-T Recommendation P.501 Annex D [10] should be used. Use of [American English samples is recommended for consistency of impairment location across the samples]. For the P.863 reference input, the fullband test signal will be low-pass filtered at 14 kHz as described in P.863 sub-clause 8.1.1.

NOTE: See ITU-T Recommendation P.501 Annex D.1 for a rationale on sample selection.

Two recordings are used to produce the speech quality loss metric:

- 1) A recording obtained in jitter and error free conditions (reference condition)
- 2) A recording obtained for the test condition of interest (test condition)

The speech quality of the signal at the input electrical interface of the test equipment is estimated using the measurement algorithm described in ITU-T Recommendation P.863 [9]. Certain speech enhancement, as described in clause 5.2, should be disabled.

The measurement will be repeated 20 times for the reference condition and n times for the test condition (as established in clause 6 for each test scenario). For both the reference and the test condition, the first measurement is discarded for convergence of the JBM. A score will be computed for each 6 s speech sentence pair and averaged to produce a mean MOS-LQO value for both the reference and test conditions.

$$\overline{MOS - LQO}_{rcv_ref} = \left(\sum_{i=1}^{19} MOS - LQO_{rcv_ref}(i) \right) / 19$$

$$\overline{MOS - LQO}_{rcv_test} = \left(\sum_{i=1}^{n-1} MOS - LQO_{rcv_test}(i) \right) / n$$

where: i is the index of the sentence pair measurement;
 n is the total number of sentence pairs measured;
 $MOS - LQO_{rcv_ref}(i)$ is the i_{th} measurement of the reference condition MOS-LQO in sending direction;
 $MOS - LQO_{rcv_test}(i)$ is the i_{th} measurement of the test condition MOS-LQO in receiving direction.

The mean speech quality loss in receiving direction $\overline{\Delta MOS - LQO_{RCV}}$, is determined from:

$$\overline{\Delta MOS - LQO_{RCV}} = \overline{MOS - LQO_{rcv_ref}} - \overline{MOS - LQO_{rcv_test}}$$

6 Test Scenarios

6.0 General

The test scenarios cover a number of typical use scenarios for IMS based speech services over LTE. The scenarios are described as follows:

- Scenario 1: Impairments at the physical layer (RF Impairments)
- Scenario 2: Impairments at the internet layer (IP Impairments)
- Scenario 3: Operation with different modes/bit-rates
- Scenario 4: Switching bit-rates / codec modes during a call
- Scenario 5: Codec change initiated by remote UE
- Scenario 6: Intra-RAT Handover (e.g. LTE to LTE)
- Scenario 7: Inter-RAT Handover (e.g. LTE to UMTS SRVCC)
- Scenario 8: Operation with DRX
- Scenario 9: Simultaneous data download/upload during a voice call

6.1 Scenario 1: Impairments at the physical layer

This test scenario simulates situations where less than ideal RF operating conditions exist for the UE. The purpose of the test is to verify that the UE operates as expected under those situations.

The UE is setup in a call with the system simulator. The system simulator used for this test scenario should be able to implement the RF fading profiles from TS 36.101 [11]. At a minimum, the EPA5, EPA7 and EVA70 fading profiles should be supported.

Measurements are performed in the receive direction only. For each test, the total number of speech sentence pairs measured is [10].

Speech quality and delay should be measured for the following codec types and rates:

- AMR-WB at 12.65 kbps and 23.85 kbps,
- AMR at 12.2 kbps,
- EVS SWB at 24.4 kbps, 13.2 kbps with and without channel aware mode.

The following propagation conditions should be applied as derived from 3GPP TS 36.101 Annex B:

- "Extended pedestrian A" model (EPA), maximum Doppler frequency 5 Hz
- "Extended vehicular A" model (EVA), maximum Doppler frequency 70 Hz

Low, medium and/or high correlation level may be selected.

6.2 Scenario 2: Impairments at the internet layer (IP impairments)

This test scenario simulates situations where less than ideal EPC operating conditions exist for the UE. The purpose of the test is to verify that the UE (and more specifically, its jitter buffer management) operate as expected under those situations.

The UE is setup in a call with the system simulator. The system simulator used for this test scenario should be able to introduce RTP packet delay and loss impairments in the receiving direction. At a minimum, the system simulator should be able to introduce RTP packet delays according to Gaussian and uniform distributions.

Measurements for this scenario are performed in the receive direction only. For each test, the total number of speech sentence pairs measured is [= 40].

Speech quality and delay should be measured for the following codec types and rates:

- AMR-WB at 12.65 kbps and 23.85 kbps,
- AMR at 12.2 kbps,
- EVS SWB at 24.4 kbps, 13.2kbps with and without channel aware mode.

Tests with file-defined impairment profiles should use the same test method and the same impairment profiles as already specified in 3GPP TS 26.132 clauses 7.10.4 for AMR, 8.10.4 for AMR-WB and 9.10.4 for EVS SWB.

In addition, the following impairments may be used:

- Gaussian distribution: Downlink Jitter: $\sigma = 10$ ms; Packet Loss: 1 %;
- Gaussian distribution: Downlink Jitter: $\sigma = 40$ ms; Packet Loss: 3 %;
- Gaussian distribution: Downlink Jitter: $\sigma = 20$ ms; Packet Loss: 2 %;
- [Uniform/Pareto distribution: Downlink Jitter: $\sigma = 10$ ms; Packet Loss: 1 %];
- [Uniform/Pareto distribution: Downlink Jitter: $\sigma = 40$ ms; Packet Loss: 3 %];
- [Uniform/Pareto distribution: Downlink Jitter: $\sigma = 20$ ms; Packet Loss: 2 %];

While these impairments are often used in commercial testing, they are not representative of typical LTE networks.

6.3 Scenario 3: Operation with different modes / bit-rates (reference)

This test scenario simulates operation of the UE at different bit-rates. The purpose of the test is to verify that the UE has implemented the speech encoder / decoder appropriately.

The UE is setup in a call with the system simulator. The system simulator used for this test scenario should be able to configure the bit-rate / operation mode of the speech codec being used in the call. Different bit rates / operation modes are tested and MOS-LQO scores are compared to expected values.

Measurements for this scenario are performed in the send and receive directions. For each test, the total number of speech sentence pairs measured is [= 20].

- Speech quality and delay should be measured for the following codec types and rates: AMR rates 4.75 kbps to 12.2 kbps
- AMR-WB rates 6.60 kbps to 23.85 kbps
- EVS primary mode NB 5.9 kbps to 9.6 kbps
- EVS primary mode WB 5.9 kbps to 13.2 kbps
- EVS primary mode SWB 9.6 kbps to 24.4 kbps

- EVS AMR-WB IO mode rates 6.60 kbps to 23.85 kbps

Higher bit-rates for EVS primary mode may also be tested.

6.4 Scenario 4: Switching bit-rates / operating modes during a call

This test scenario simulates the situation where bit-rates / operating modes of a codec are switched during a call. The purpose of the test is to verify that the UE has implemented the speech encoder / decoder appropriately and that the speech quality after a bit-rate / mode-switch is as expected.

The UE is setup in a call with the system simulator. The system simulator used for this test scenario should be able to initiate a codec rate change during a call. The codec rate change is initiated and completed during a speech pause. The MOS-LQO scores for the speech material after the codec rate change are compared to expected values.

Measurements for this scenario are performed in the send and receive directions. For each test, the total number of speech sentence pairs measured is [10].

Speech quality should be measured after the following rate changes:

- AMR 12.2 to/from 5.9.
- AMR-WB 23.85 to/from 12.65.
- EVS SWB 24.4 to/from 13.2 with and without channel aware mode.

The results should be compared with the corresponding results of Scenario 3.

6.5 Scenario 5: Codec change initiated by remote UE

This test scenario simulates the situation where a codec change is initiated by a remote UE, e.g. when the remote UE uses SRVCC to handover between an LTE area with wideband speech coverage and an UMTS area with narrowband speech coverage only. The purpose of the test is to verify that the UE can successfully perform the codec change requested and that the speech quality after the codec rate change

The UE is setup in a call with the system simulator. The system simulator used for this test scenario should be able to initiate a codec rate change during a call. The codec rate change is initiated and completed during a speech pause. The MOS-LQO scores for the speech material after the codec rate change are compared to expected values.

Measurements for this scenario are performed in the send and receive directions. For each test, the total number of speech sentence pairs measured is [10].

Speech quality should be measured after the following codec changes:

- AMR-WB 12.65 to AMR 12.2
- EVS Primary SWB 24.4 to AMR 12.2
- EVS Primary SWB 24.4 to AMR-WB 12.65
- EVS Primary SWB 13.2 to AMR 12.2
- EVS Primary SWB 13.2 (with and without channel aware mode) to AMR-WB 12.65
- EVS Primary SWB 24.4 to EVS AMR-WB IO 12.65
- EVS Primary SWB 13.2 (with and without channel aware mode) to EVS AMR-WB IO 12.65

The results should be compared with the corresponding results of Scenario 3.

6.6 Scenario 6: Intra-RAT Handover

This test scenario simulates the situation where a UE performs a handover within a single radio access technology. The purpose of the test is to verify that no speech quality degradation exists after the handover is completed.

The UE is setup in a call with the system simulator. The system simulator used for this test scenario should be able to initiate an intra-RAT handover. The handover is initiated and completed during a speech pause. The MOS-LQO scores for the speech material before and after the handover are compared.

Measurements for this scenario are performed in the send and receive directions. For each test, the total number of speech sentence pairs measured is [= 10].

Speech quality should be measured before and after handover for these codec types and rates:

- AMR-WB at 12.65 kbps, 23.85 kbps
- AMR at 12.2 kbps
- EVS SWB at 24.4 kbps and 13.2 kbps with and without channel aware mode

The results before and after handover should be compared with the corresponding results of Scenario 3.

6.7 Scenario 7: Inter-RAT Handover

This test scenario simulates the situation where a UE performs a handover within two different radio access technologies. The purpose of the test is to verify that no speech quality degradation exists after the handover is completed.

The UE is setup in a call with the system simulator. The system simulator used for this test scenario should be able to initiate an inter-RAT handover from LTE to UMTS (SRVCC). The handover is initiated and completed during a speech pause. The MOS-LQO scores for the speech material before and after the handover are compared.

Measurements for this scenario are performed in the send and receive directions. For each test, the total number of speech sentence pairs measured is [= 10].

Speech quality should be measured after the handover from LTE to UMTS (SRVCC) with these codecs and rates:

- LTE AMR-WB 23.85 to UMTS AMR 12.2
- LTE AMR-WB 12.65 to UMTS AMR 5.9
- LTE AMR-WB 12.65 to UMTS AMR-WB 12.65
- LTE EVS SWB 24.4 to UMTS AMR-WB 12.65
- LTE EVS SWB 24.4 to UMTS AMR 12.2
- LTE EVS SWB 13.2 (with and without channel aware mode) to UMTS AMR-WB 12.65
- LTE EVS SWB 13.2 (with and without channel aware mode) to UMTS AMR 12.2

The results after handover should be compared with the corresponding results from Scenario 3.

6.8 Scenario 8: Operation with DRX

This test scenario simulates the situation where a UE operates with discontinuous reception technology (DRX). The purpose of the test is to verify that no speech quality degradation exists due to the presence of DRX.

The UE is setup in a call with the system simulator. The system simulator used for this test scenario should support DRX operation. The MOS-LQO scores for the speech material with and without DRX are compared.

Measurements for this scenario are performed in the send and receive directions. For each test, the total number of speech sentence pairs measured is [10].

Speech quality should be measured for these codecs and rates:

- AMR-WB at 12.65 kbps, 23.85 kbps
- AMR at 12.2 kbps
- EVS SWB at 24.4 kbps and 13.2 kbps with and without channel aware mode.

The results should be compared with the corresponding results from Scenario 3. The long and short cycle timers describing the cDRX setting should be described in the test report.

6.9 Scenario 9: Voice call with simultaneous data upload/download

This test scenario simulates the situation where a UE is performing an upload/download during a voice call. The purpose of the test is to verify that no speech quality degradation exists due to the upload/download operation

The UE is setup in a call with the system simulator. The system simulator used for this test scenario should be able to upload/download data from/to the network. The MOS-LQO scores for the speech material with and without the data transfer are compared.

Measurements for this scenario are performed in the send and receive directions. For each test, the total number of speech sentence pairs measured is [= 10].

Speech quality and delay should be measured at these rates:

- AMR-WB at 12.65 kbps, 23.85 kbps
- AMR at 12.2 kbps
- EVS SWB at 24.4 kbps and 13.2 kbps with and without channel aware mode

The results should be compared with the corresponding results from Scenario 3. The QCIs, bearers for speech and data connections, data transfer rates and payload should be indicated in the test report.

Annex A: Test Method for Interruption Time

[The test signal is a continuously looped CSS signal, consisting of a voiced signal and a PN sequence, both according to ITU-T P.501 section 7.2.1.1, without pause. This signal provides a periodically recurring voiced signal to keep the voice activity detector activated. The bandwidth of the signal may be adapted to the codec bandwidth, or a narrowband signal may be used for all codec bandwidths. The peak to active speech level of this narrowband signal is about 8 dB (GapTestSignal_NB.wav attached). The test signal level should be -16 dBm0 on the network side.

The measurement is based on peak detection. The signal is analysed in time intervals of at least 1,39583 ms corresponding to 67 samples or a half period of the voiced activation signal part in the CSS (in a 48 kHz sampling system). The analysis result for each time interval is the maximum absolute value of all sample values within this time interval. These maximum absolute values are evaluated on a time axis as shown in the figure below. The time value for each maximum absolute value is the start time of the respective analysis time interval.

The start time of the gap is the lowest time value within the acquisition interval, for which the maximum absolute value falls below a predefined threshold, e.g. -34 dBov. The end time of the gap is the highest time value within the acquisition interval, for which the maximum absolute value falls below this predefined threshold. The interruption time is the difference between the end time of the gap and the start time of the gap.]

Annex B: Change history

Change history							
Date	Meeting	TDoc	CR	Rev	Cat	Subject/Comment	New version
12-2016	TSG SA#74	SP-160920				Presented to TSG SA#74 plenary (for information)	1.0.0
06-2017	TSG SA#76	SP-170326				Presented to TSG SA#76 plenary (for approval)	2.0.0
	TSG SA#76					Approved at TSG SA#76 plenary	14.0.0

History

Document history		
V14.0.0	July 2017	Publication