ETSI TR 102 716-1 V1.1.1 (2009-07)

Technical Report

Speech and multimedia Transmission Quality (STQ); Guidelines, objectives and results of speech quality analysis in the context of interworking Plugtests for multiplay services; Part 1: Guidelines and objectives

Reference

DTR/STQ-00132-1

Keywords

interoperability, quality, voice

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

Contents

Intellectual Property Rights 4
Foreword..4
1 Scope 5
2 References 5
2.1 Normative references 5
2.2 Informative references..5
3 Abbreviations 6
4 Context 7
5 Type of measurement 7
6 Evaluation chain..8
6.1 Take into account electro-acoustic characteristics 8
6.2 Unique evaluation chain implementation..8
6.3 Evaluation chain description 9
7 Metrics..10
7.1 Post Dialling Delay 10
7.2 Listening speech quality..10
7.3 End to End delay 11
7.4 Listening speech quality stability 11
7.5 Active level of speech signal at reception 12
7.6 Reception noise level..12
7.7 Speech signal attenuation (or gain) after transmission 12
7.8 Talker Echo Delay..13
7.9 Talker Echo Attenuation 13
7.10 DTMF integrity 13
8 Results reports..14
8.1 Measurement configuration description 14
8.2 One-view visualization of performances..14
8.3 Visualization of performance variations..15
8.4 Result interpretations..16
9 Results comments..16
History 18

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 (http://webapp.etsi.org/IPR/home.asp).

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 Technical Committee Speech and multimedia Transmission Quality (STQ).

The present document is part 1 of a multi-part deliverable covering Guidelines, objectives and results of speech quality analysis in the context of interworking Plugtests for multiplay services, as identified below:

Part 1: "Guidelines and objectives";

Part 2: "Results".

1 Scope

The present document provides objectives for speech quality evaluation in the specific context of interoperability event. In this context where acoustic measurements are not possible, is specified the type of measurement chain to be implemented and the necessary calibrations. In the present document, are presented a list of metrics that can be assessed and the mode of result presentation. It is also specified comments to be drafted in test reports to clarify the context in which analyses were carried out.

2 References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific.

- For a specific reference, subsequent revisions do not apply.
- Non-specific reference may be made only to a complete document or a part thereof and only in the following cases:
  - if it is accepted that it will be possible to use all future changes of the referenced document for the purposes of the referring document;
  - for informative references.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication ETSI cannot guarantee their long term validity.

2.1 Normative references

The following referenced documents are indispensable for the application of the present document. For dated references, only the edition cited applies. For non-specific references, the latest edition of the referenced document (including any amendments) applies.

Not applicable.

2.2 Informative references

The following referenced documents are not essential to the use of the present document but they assist the user with regard to a particular subject area. For non-specific references, the latest version of the referenced document (including any amendments) applies.

[i.1] ETSI EG 202 765-2: "Speech Processing, Transmission and Quality Aspects (STQ); QoS and network performance metrics and measurement methods; Part 2 : Transmission Quality Indicator combining Voice Quality Metrics".

[i.2] ITU-T Recommendation P.862: "Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs".

[i.3] ITU-T Recommendation P.862.1: "Mapping function for transforming P.862 raw result scores to MOS-LQO".

[i.4] ITU-T Recommendation P.862.2: "Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs".

[i.5] ITU-T Recommendation P.862.3: "Application guide for objective quality measurement based on Recommendations P.862, P.862.1 and P.862.2".

[i.6] ITU-T Recommendation P.800: "Methods for subjective determination of transmission quality".

[i.7] ITU-T Recommendation P.800.1: "Mean Opinion Score (MOS) terminology".

[i.8] ITU-T Recommendation G.114: "One-way transmission time".

[i.9] ITU-T Recommendation P.56: "Objective measurement of active speech level".

[i.10] ITU-T Recommendation O.41: "Psophometer for use on telephone-type circuits".

[i.11] ETSI ES 201 970: "Access and Terminals (AT); Public Switched Telephone Network (PSTN); Harmonized specification of physical and electrical characteristics at a 2-wire analogue presented Network Termination Point (NTP)".

[i.12] ITU-T Recommendation G.131: "Talker echo and its control".

[i.13] ITU-T Recommendation G.168: "Digital network echo cancellers ".

[i.14] ITU-T Recommendation Q.23: "Technical features of push-button telephone sets".

[i.15] ETSI ES 201 235-1: "Specification of Dual Tone Multi-Frequency (DTMF) Transmitters and Receivers; Part 1: General".

[i.16] ETSI ES 201 235-3: " Access and Terminals (AT); Specification of Dual-Tone Multi-Frequency (DTMF) Transmitters and Receivers; Part 3: Receivers".

[i.17] ITU-T Recommendation P.505: "One-view visualization of speech quality measurement results".

3 Abbreviations

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

DSLAM	Digital Subscriber Line Access Multiplexer
DTMF	Dual Tone Multi-Frequency
GPON	Gigabit Passive Optical Network
GW	GateWay
HomeGW	Home GateWay (Referenced also as Residential Gateway)
ID	IDentification
IP	Internet Protocol
ISDN	Integrated Services Digital Network
ITU-T	International Telecommunication Union - Telecommunication standardization sector
MediaGW	Media GateWay
MOS	Mean Opinion Score
MOS-LQOM	Mean Opinion Store - Listening Quality Objective Mixed bandwidths
PDD	Post Dialling Delay
PESQ	Perceptual Evaluation of Speech Quality
PSTN	Public Switched Telephone Network
TV	TeleVision
VoD	Video on Demand
VoIP	Voice over Internet Protocol
xDSL	x Data Subscriber Line (where x represent the different associated technologies)

4 Context

Frequently during Plugtests, speech quality analysis come to complete inter-working tests. It was the case for the different xDSL and GPON Plugtests organized by ETSI at Lannion in 2006, 2007 and 2008. The tests conditions are perfectly adapted to protocol and inter-working analysis: the different participants are located and installed in the same area (very large room). A precise test planning allows several tests of manufacturer equipments in various conditions and configurations.

Installing different equipment and expects (manufacturer and organizer teams) in the same area, makes easier interworking test organization but does not permit at all acoustic measurements. These environments without control of acoustic condition require adaptation of speech quality analysis and evaluation restriction to specific indicators.

Figure 1: Photo presenting a part of the area used for xDSL and GPON interoperability event organized by ETSI at Lannion in June, 2007

Moreover, timeslots fixed to every manufacturer for every test condition are relatively short (1 or 2 hours) to allow performing a maximum of different configurations analysis. These relatively short timeslots are also required to restrict speech quality analysis to main parameters such as speech signal distortion after transmission and one way transmission delay.

5 Type of measurement

Considering conditions in which these tests are carried out, the speech quality analysis should be performed only in electric to electric configuration (analysers connected on electrical interfaces of network). Acoustic conditions are absolutely not controlled in the area where interoperability event take place, the speech quality measurements should be carried out from electric interfaces like ISDN or analogue accesses of the PSTN, electric accesses of equipment or handset interface of terminal.

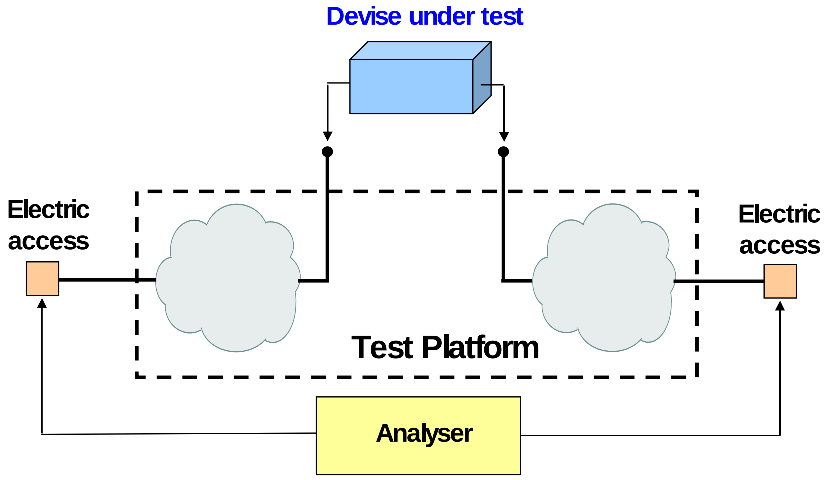

Figure 2: Overview diagram schematizing the measurement bench to be carried out

The device to be tested is characterized on an electric-electric configuration.

6 Evaluation chain

6.1 Take into account electro-acoustic characteristics

Although analyses are carried out from electric interfaces, it is necessary to take into account -during analysis- the transducers (microphones and loudspeakers) characteristics. This restriction (implementation of electric to electric configuration) requires a pre-calibration of the evaluation chain in acoustic situation. This calibration allows determination of electro-acoustic characterization of transducers, characterization which is used to balance measurements performed via electric interfaces.

Information concerning the pre-calibration using the electro-acoustic parts of the evaluation chain is available in EG 202 765-2 [i.1], annex B.

6.2 Unique evaluation chain implementation

The interoperability event context consists in assessing different equipment industrialized by different manufacturers. Therefore, test conditions consist in carrying out inter-work between 2 (or several) types of equipment, like HomeGW with DSLAM, IPphone with gatekeepers, HomeGW with MediaGW, etc.

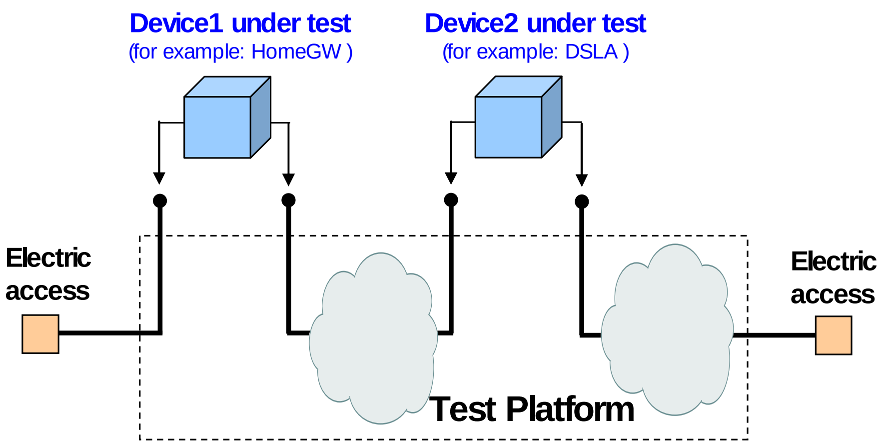

Figure 3: Overview diagram schematizing a configuration of interfunctioning test

In the case presented above, it can be an interfunctioning between a HomeGW and a DSLAM.

In any case, for a given type of validated equipment (HomeGW, IPphone, MediaGW) it is necessary that speech quality evaluation should be performed on the same evaluation chain. This condition allows results and equipment performances comparison. The analyses of speech quality require carrying out a sort of "reference platform" where all equipment has well known characteristics.

The reference platform should utilize equipment whose characteristics are well known either because equipment is already deployed in the network or because it is already used in laboratory. The reference platform should not utilize equipment whose characteristics are not well known.

In these conditions, the speech quality analysis will be performed on a platform where all the equipment of the chain has perfectly known characteristics except the device concerned by the tests of speech quality.

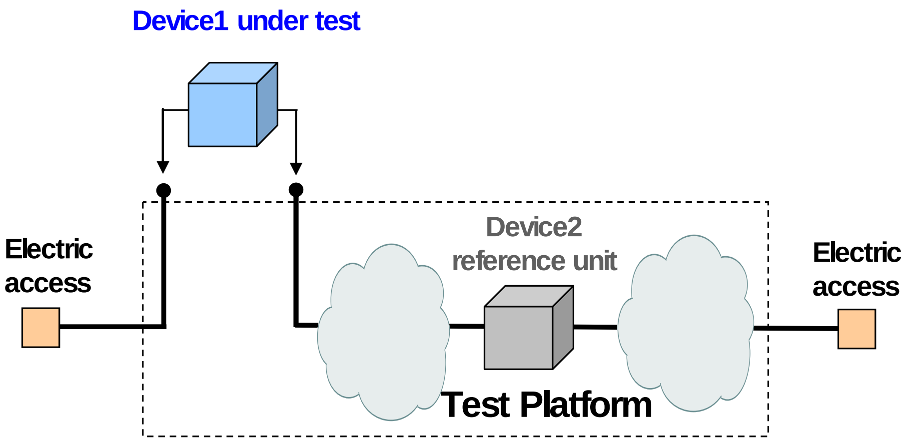

Figure 4: Overview diagram schematizing the measurement bench to carry out speech quality characterizations on Device 1 (like HomeGW for example)

To carry out analysis on Device 1, Device 2 under test is replaced by a reference equipment. With this configuration, any Device 1 going through speech quality analysis will be tested in the same conditions with the same platform.

6.3 Evaluation chain description

Besides the fact that the measurement chain should be perfectly controlled with perfectly known characteristics, it is necessary that this chain of measure will be described exactly. This requirement allows to know exactly the conditions of analysis and to save the context in which these analyses were conducted. By providing and recording all this information it will be possible (if necessary) to repeat analyses in the same conditions.

Concerning the test platform it is necessary to produce:

- a description of the deployed architecture (equipment and connection between them);
- a description of every equipment mentioning manufacturer, model, firmware version and so on;
- a description of the implemented configuration in the case of equipment having several functioning modes;
- a description of characteristics of the test platform relevant to the tests conducted.

Concerning the device to be tested it is necessary to produce:

- description mentioning manufacturer, model and firmware version;
- description of the device configuration during the tests (for example: the negotiated codec, the payload size, the signalization protocol).

Concerning the chain of analysis, it is necessary to give a description of the analyzers and their connections on test platform.

Concerning the indicators it is necessary to specify:

- the list of metrics and their meaning;

- the assessment method associated with these indicators;
- and the target values for reporting.

7 Metrics

The measurable metrics during tests are the ones described in the following clauses.

7.1 Post Dialling Delay

Definition	Time interval between the end of dialling by the caller and the reception back by him of the appropriate ringing tone or recorded announcement. Metric determines on one of the two access of the communication.
Assessment method	Indicator determines sequentially from the two access of call configuration. This indicator characterizes only the caller part of the configuration.
Unit	Millisecond.
Standardization reference	
Comment	This indicator has to be separated between call types (IP to IP, IP to PSTN, etc.) for a detailed analysis.

7.2 Listening speech quality

Definition	Intrinsic quality of speech signal after transmission. This indicator takes into account degradations generate on the signal by the transmission links.
Assessment method	Voice quality is evaluated by using the ITU-T Recommendation P.862 [i.2] with the mapping functions according to ITU-T Recommendations P.862.1 [i.3] and P.862.2 [i.4]. MOS or Mean Opinion Score (calculated using the Perceptual Evaluation of Speech Quality, or PESQ method) provides an objective view on the quality of the voice signal as it may be perceived by the customer. The MOS score is obtained by comparing speech samples: - the original signal sent by the far end of the connection; - the degraded signal received at the local end, where the measurement is applied. For this analysis, we can use a speech sample corresponding to several speakers pronouncing different sentences. The use of a speech sample (test sequence) corresponding to 4 speakers pronouncing each two sentence is perfectly adapted to this measure. Such a sample can have a duration of about twenty seconds. The voice quality indicator is determined in the two directions of transmission. Several MOS scores are determined in series during the same call. So for a given transmission way, Listening Speech quality performance during the call is defined by the mean value of MOS-LQOM measurements (in the same direction). The indicator evaluation requires repeating measurement successively during a period of 10 minutes at least. During this period, a minimum of 10 measures should be performed. This indicator is characterized by the mean value of the test sequence and also by the maximal and the minimal values. The analysis duration and the number of measures should also be indicated in the results presentation.
Unit	Note between 1 (=very bad) and 5 (=excellent) determines on MOS-LQOM scale.
Standardization reference	ITU-T Recommendations P.800 [i.6], P.800.1 [i.7], P.862 [i.2], P.862.1 [i.3], P.862.2 [i.4], P.862.3 [i.5].
Comment	The value of this indicator depends on the used codec, but also on impairments like IP packet loss or low signal to noise ratio. It is a one-way indicator, therefore it should be measured separately in both transmission directions, with no average between them afterwards. This indicator may be separated between call types (IP to IP, IP to PSTN, etc.) for a detailed analysis.

7.3 End to End delay

Definition	Global delay from one access to the other one. This indicator takes into account transmission delay on networks but also processing delay in sending and receiving terminals. In these specific conditions, the network delay is often very low because it is delay introduced by a test platform and not a real network. So it is delay measurement in optimal conditions.
Assessment method	The end to end delay is the delay from mouth to ear, which means the transmission delay over the whole transmission path. For the purpose of the present document, end to end delay does not take into account the transducers delay (loudspeaker and microphone) while measurements are done at the electrical interfaces of the end terminals. To measure end to end delay it is needed to ensure a synchronization of both transmission ends of the measurement device. In the context of the present document both end transmission access are located in the same area, so synchronization may be done directly by the analyser. End to end delay is determined in the two directions of transmission. The indicator evaluation requires repeating measurement successively during a period of 10 minutes at least. During this period, a minimum of 10 measures should be performed. This indicator is characterized by the mean value of realized measurements and also by the maximal and the minimal values. The analysis duration and the number of measures should also be indicated in the results presentation.
Unit	Millisecond.
Standardization reference	ITU-T Recommendation G.114 (session 1) [i.8].
Comment	The standards (ITU-T Recommendation G.114 [i.8] in particular) recommend not going beyond 150 ms in one-way.

7.4 Listening speech quality stability

Definition	Speech quality stability estimator during a call. This metric represent the stability of the voice quality during a communication of several minutes. This indicator takes into account degradations generate on the signal by the transmission links. Concerning voice over IP, a single measurement of speech quality once at the very beginning of a call is not enough. They should be analysed all along the duration of the call, typically several minutes.
Assessment method	Several measurements of MOS score performed with ITU-T Recommendation P.862 [i.2] in conjunction with ITU-T Recommendation P.862.1 [i.3] are performed in series in the same call. For this analysis, the data acquired according to clause 7.2 are used. The assessment of Listening Speech Quality Stability is preformed as follow: For each MOS value the absolute difference with the previous MOS value, is calculated. In order to take into account the subjective perception and measurement accuracy, the absolute difference values are set to 0 when the values are equal to or lower than 0.1. The MOS instability associated to speech quality during the call is defined by mean value of these absolute difference values. A linear weighting function is applied in order to express MOS Stability on a 0 to 100 scale: MOS Stability = 100 - (250 × MOS instability); and MOS Stability = 0 if [100-(250 × MOS instability)] < 0. This indicator is determined in the two directions of transmission. All information concerning listening speech quality stability can be found in EG 202 765-2 [i.1], clause 7.12 and annex A.
Unit	Statistics on MOS score variation are plotted on a 0 to 100 scale.
Standardization reference	EG 202 765-2 [i.1].
Comment	For this analysis the measure should be repeated at least during 10 minutes to have at least 10 measurements of speech vocal during the same call. Only successfully completed calls should be taken into account for this indicator.

7.5 Active level of speech signal at reception

Definition	Speech signal level at the reception after transmission. The level of the signal heard by the user has an impact on the quality he will perceive. A too low signal will be hardly audible and masked by the noise, while a too high level will be painful. Therefore, a measurement of the speech signal level is necessary to ensure a good listening comfort.
Assessment method	The received decoded signal used for instance of ITU-T Recommendation P.862 [i.2] can be used also to assess this parameter. A typical method for the measurement of this parameter, based on a sample by sample approach and a moving threshold between noise and speech, is given in ITU-T Recommendation P.56 [i.9].
Unit	dBm.
Standardization reference	ITU-T Recommendation P.56 [i.9].
Comment	Each sample of signal has a level, generally express in mV. The mean speech level is the transformation on as appropriate logarithmic scale of the mean signal voltage. The samples taken into account for this measurement are the ones seen as speech (the other ones are taken into account for noise measurements). It is recommended to fall within classical speech levels values, i.e. between -25 dBm and -10 dBm.

7.6 Reception noise level

Definition	Noise level determined at reception in non-speech segment of speech sample. The noise present besides the speech signal can have characteristics that can become a disagreement, for instance if they have a varying spectrum (crowd, noise, for instance). But the more important source of annoying due to noise is simply its level.
Assessment method	The received decoded signal used for instance for ITU-T Recommendation P.862 [i.2] can be used also to assess this parameter. The measurement of these parameters is normally performed as for speech signal level (see clause 7.6), but on the samples identified as non-speech.
Unit	dBmOp.
Standardization reference	ITU-T Recommendation O.41 [i.10].
Comment	Each sample of signal has a level, generally express in mV. The mean noise level is the transformation on as appropriate logarithmic scale of the mean signal voltage of the noise samples. To get a more accurate noise level measure, a frequency transform needs to be done in order to apply a *psophometric* weighting (see ITU-T Recommendation O.41 [i.10]). It is recommended not to have noises louder than -50 dBmOp.

7.7 Speech signal attenuation (or gain) after transmission

Definition	Difference between the active speech level at receiving and sending access.
Assessment method	The received decoded signal used for instance of ITU-T Recommendation P.862 [i.2] can be used also to assess this parameter. Once the speech signal level has been computed (see clause 7.2), it is compared with the level of the sent signal. The attenuation is the difference between these two levels. There are other methods to compute this parameter, based for instance on intrusive measurement made with sine waves and a specific weighting function. Indicator determines in the two directions of transmission.
Unit	dB.
Standardization reference	ES 201 970 [i.11].
Comment	It is recommended to comply with PSTN attenuation rules, i.e. an attenuation between 6 dB and 10 dB. When ISDN access is used in the test platform, the attenuation should be as close as possible to 0 dB. This indicator characterizes the compliance with ES 201 970 [i.11].

7.8 Talker Echo Delay

Definition	Time it takes for the speech signal to go from the mouth of a subscriber back to the ear of the same subscriber, with one or more reflections occurring along the transmission path. In electric to electric condition, talker echo delay is the time it takes for the speech signal to go from the measurement access back to this same access. In Telecommunications, the term Echo describes delayed and unwanted feedback of the send signal into the receive path. The so-called echo source is the reflection point between send and receive directions, which could be caused by one of the following: - 4-wire/2-wire Hybrid Circuits (multiple reflections possible); - Coupling in Handset Cords; - Structure Borne Coupling in Handsets; - Acoustical Coupling between Earpiece and Microphone. This phenomenon is characterized by two parameters: its attenuation and its delay. With the increased delays present in today's IP networks, echo has the potential to be much more perceivable and annoying than in classical PSTN. In order to achieve a similar user perception with higher delays the attenuation of the talker echo should be increased, i.e. active echo cancellation is necessary.
Assessment method	Indicator should be assess in both directions.
Unit	Milliseconds.
Standardization reference	ITU-T Recommendation G.131 [i.12].
Comment	For fully digital networks the talker echo delay can be assumed to be equivalent to twice the mean one-way delay.

7.9 Talker Echo Attenuation

Definition	Difference of level between the sending signal and the receiving signal both measured at the access.
Assessment method	Indicator should be assess in both directions.
Unit	dB.
Standardization reference	ITU-T Recommendations G.131 [i.12] and G.168 [i.13].
Comment	If active echo cancellation is used, ITU-T Recommendation G.168 [i.13] should be applied; VoIP should provide echo attenuation of 55 dB. If the delay under all circumstances is known to not exceed 50 ms lower values for the talker echo attenuation (down to 35 dB) may be acceptable. Echo annoyance depends on two metrics: the attenuation and the delay. For a similar attenuation level greater the delay is more important the annoyance will be for the user(s). Echo-Annoyance factor is defined as K. $K = EA - 40 \times \mathrm{Log}[\,(1+delay/10) / (1+delay/150)\,] + 6 \times \exp(-0{,}3 \times delay^2)$

7.10 DTMF integrity

Definition	Indicator characterizing the capacity of the transmission chain to transmit correctly the DTMF codes.
Assessment method	A sequence containing all DTMF codes is sent on one transmission access of evaluation chain. The signal received on the other transmission access is analyzed to verify if DTMF code characteristics are compliant with recommendations of the subject. Analysis performs in the two directions of transmission.
Unit	Indicator express by a Boolean metric: Past or Failed.
Standardization reference	ITU-T Recommendation Q.23 [i.14], ES 201 235-1 [i.15], ES 201 235-3 [i.16].
Comment	DTMF code sequence like "0123456789*#ABCD" can be used for analysis. It is recommended to perform several time (10 for example) this analysis to sure that DTMF codes are correctly transmitted and recognized.

We will consult EG 202 765-2 [i.1] for more information on these indicators and on the measurement methods.

The metrics Listening speech quality, End to End delay, Listening speech quality stability, Talker Echo Delay and Talker Echo Attenuation are the most important metrics. They should be evaluated as much as possible during this type of event.

8 Results reports

Results are produced in 2 types of report:

- individual test report for each manufacturer summarizing this equipment performances; and
- an anonymous report reproducing all obtained results during the measurement campaign or during the Plugtests.

8.1 Measurement configuration description

In the 2 types of report, it is necessary to describe the analysis chain including network architecture and speech quality analyser implemented to carry out measurements.

All information concerning the measurement chain (type of equipment, manufacturer and model of these equipments, firmware version and configuration) should be present in these reports. In the same manner, information concerning device tested (model and firmware ID, configuration during analysis) will be presented.

To respect the anonymity of manufacturers, in the general report presenting all the results of the analysis, the information identifying the tested devices will be removed.

It is also necessary to introduce the methodology: evaluated indicators, what they represent and the method to determine them. The means implemented to realize analyses (analyzers) as well as the measurement methods should be described.

8.2 One-view visualization of performances

In order to give a quick overview of all quality parameters, a specific representation (overview visualization defined in ITU-T Recommendation P.505 [i.17]) of the metric value will be used. This representation reveals at one glance the strengths and weaknesses of devices under test with respect to target values.

This specific presentation (Pie Diagram) provides an "aggregate" results view. This presentation is particularly adapted to show all results (all tested equipment performances), allowing an easy performance comparison of the tested equipment. The Pie Diagram makes also easier the comparison of the results when several configurations of the device under test are analyzed.

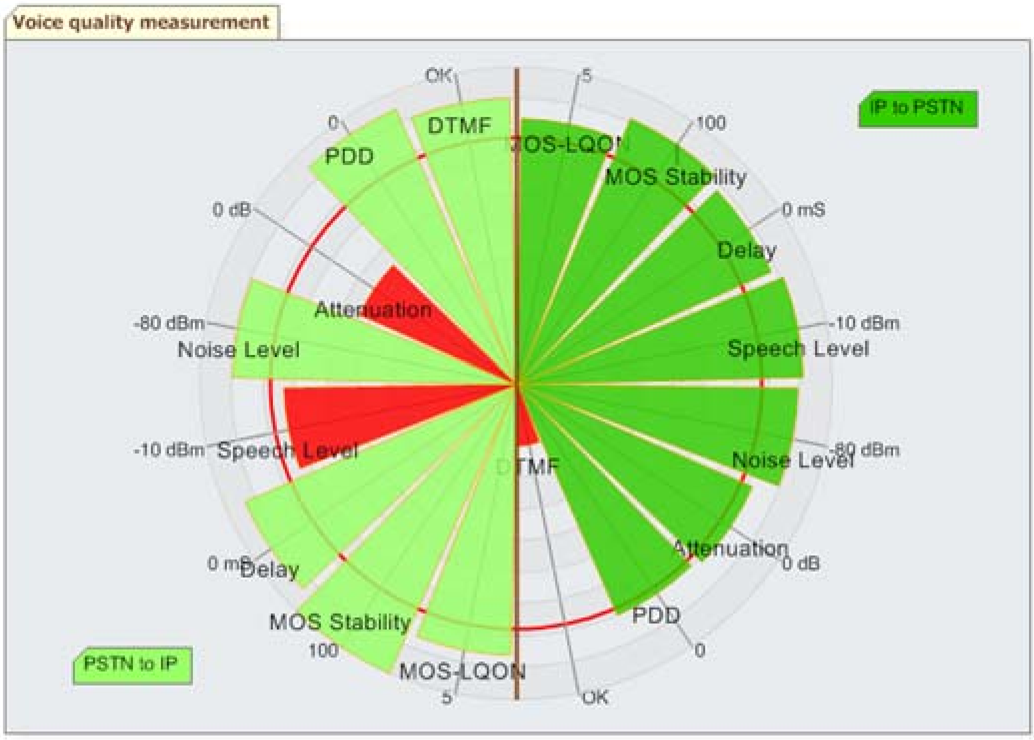

Figure 5: Example of results presented with the Pie Diagram presentation

On this figure, results are presented for the two transmission paths (IP to PSTN and PSTN to IP) and for every indicator the measured value is compared to an acceptance threshold.

8.3 Visualization of performance variations

For the report to the manufacturer (individual test report), it is also important to present the variation of indicators. For the speech quality and End to End delay, the indicators show the time variation within the same communication. For PDD and DTMF integrity, the indicators show the variation between successive analyses. The presentation of these indicator variations can be shown through graphs as those below.

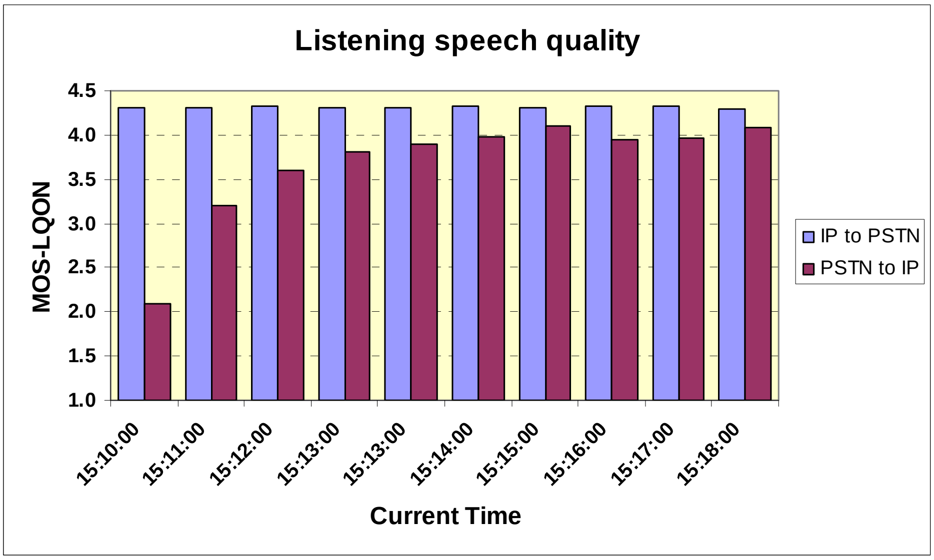

Figure 6: Example of graph for speech quality indicator

On this figure, MOS scores (determined successively in the same call) are presented for the two transmission paths (IP to PSTN and PSTN to IP). For speech quality, the graphic presentation of MOS score values versus time completes information given by the **speech quality stability indicator**.

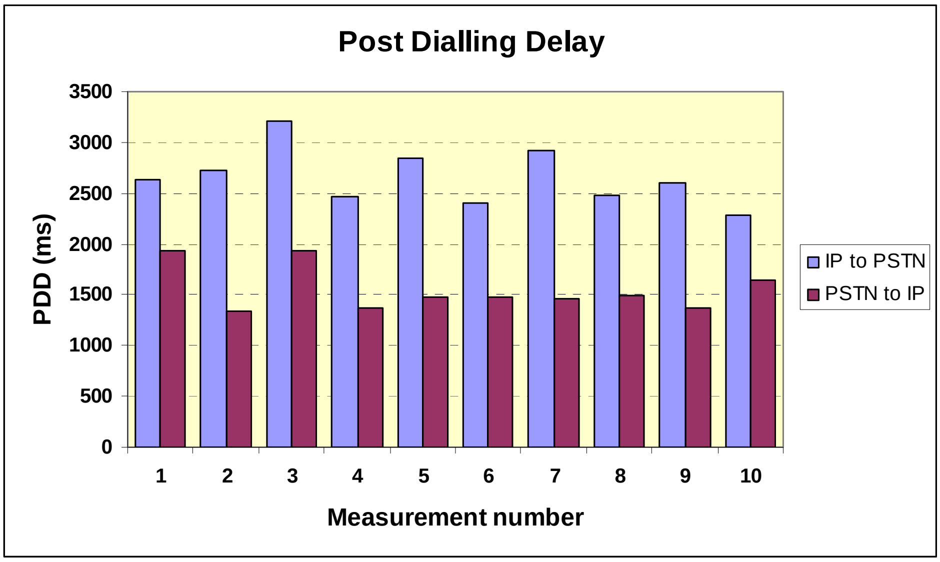

Figure 7: Example of graph for PDD

On this figure, indicator values (determined by successive call attempts) are presented for the two call paths (IP to PSTN and PSTN to IP).

8.4 Result interpretations

An interpretation of the results would be presented in the reports of the analysis. For each metric, it is necessary to explain the results obtained during the tests. In the individual reports to the manufacturer, the result interpretation is given according to the target values by trying, if necessary, to identify the cause of degradation or low performance. In the general report, the analysis of the results has to show the comparison of the performance levels (maximum, medium and minimum) of the tested devices.

9 Results comments

Comments are attached to the results, particularly on reference to the target values.

Independently from the results comments, it is also necessary to point out clearly the limited impact of characterization type. The reasons of limited characterization are:

- measurements performed from electric interfaces (no acoustic evaluation);
- a limited number of tested configurations;
- a limited number of measured indicators;
- measurements performed in situation of voice traffic only (no data, TV and VoD flow simultaneously).

In the report, it should be clearly mentioned that evaluation is partial, that even if the analyses results are correct (or good) it does not mean that the equipment (or tested service) respects all requirements associated with voice quality.

Here is an example indicating the limits of these tests:

Specific note to manufacturer

The analyses of speech quality performed during these Plugtests are limited:

- by the conditions of test (electric to electric conditions);
- by the number of implementation conditions (PSTN to IP and IP to PSTN conditions only);
- by the number of metrics (MOS score, MOS stability, delay, delay stability, signal attenuation after transmission, noise level at the reception, PDD and DTMF integrity).

Obtaining good results for the measurement of the main indicators characterizing voice quality does not mean that the device is completely validated. To validate the device, it is absolutely necessary to complete the tests performed during this Plugtests by more complete analyses requiring particular conditions like acoustic to electric test configuration and also use cases such as double talk.

History

Document history		
V1.1.1	July 2009	Publication