ETSI TR 102 648-1 V1.1.1 (2006-12)

**Speech Processing, Transmission and Quality Aspects (STQ);
Test Methodologies for ETSI Test Events and Results;
Part 1: VoIP Speech Quality Testing**

Reference

DTR/STQ-00079-1

Keywords

interoperability, quality, speech, VoIP

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

Individual copies of the present document can be downloaded from:
http://www.etsi.org

The present document may be made available in more than one electronic version or in print. In any case of existing or
perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF).
In case of dispute, the reference shall be the printing on ETSI printers of the 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
http://portal.etsi.org/tb/status/status.asp

If you find errors in the present document, please send your comment to one of the following services:
http://portal.etsi.org/chaircor/ETSI_support.asp

Copyright Notification

Contents

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 Processing, Transmission and Quality Aspects (STQ).

The present document is part 1 of a multi-part deliverable covering the procedures and results of the ETSI VoIP Speech Quality Test Events, as identified below:

> **Part 1:** **"VoIP Speech Quality Testing";**
>
> Part 2: "Results of the 1st ETSI VoIP Speech Quality Test Event";
>
> Part 3: "Results of the 2nd ETSI VoIP Speech Quality Test Event";
>
> Part 4: "Results of the 3rd ETSI VoIP Speech Quality Test Event";
>
> Part 5: "Results of the 4th ETSI VoIP Speech Quality Test Event".

Introduction

ETSI VoIP speech quality test events have been organized by ETSI Plugtest (formerly ETSI Bake-Off Service) since the year 2000. The main goals of the events always have been:

- Test of different VoIP implementations, terminals and gateways under the identical conditions for all participating manufacturers.

- Use and improve existing ESTI standards and give appropriate feedback to the ETSI standards bodies.

- Give a feedback to the manufacturers of VoIP equipment with respect to their performance in comparison to other vendors.

The idea behind the ETSI VoIP test events is to measure, analyze and compare speech quality parameters for VoIP equipment. All conversational aspects like speech sound quality, echo measurements, double talk performance and the transmission quality in the presence of background noise are considered. These test events can be regarded as very useful for all sides:

The manufacturer:

- participate in a **tutorial** about speech quality measures;

- have **one exclusive testing day;**

- learn how **the equipment performs** in various test conditions;

- can derive useful information for **system optimization** in a special "**consulting part**" during the testing day;

- get all his **individual results** including detailed information about potential improvements;

- can compare his individual results to the results of all other participants being published in an **anonymous test report**. Comparison to the published results can be made to the results of former SQTEs.

Standardization bodies:

- achieve important data about current speech quality testing methods;

- can demonstrate the trend and development of speech quality testing methods;

- obtain an overview about conversational speech quality of current VoIP implementations available;

- can demonstrate the trend and development of VoIP speech quality by comparing the results to those from the previous events.

The present document describes the latest procedures used in the ETSI VoIP test events. It is used as a guideline for the implementation of such tests.

1 Scope

The present document is a guideline for tests to be conducted in the ETSI VoIP speech quality test events. The present document describes:

- the general test conditions;

- the test setup;

- the test methodologies;

- and the result representations.

Testing as described in the present document ensures comparability of the results between tests performed in different test events.

The test principles described in the present document are applicable for tests:

- between the two acoustic interfaces of a connection;

- between the acoustic interface and the electrical access point of a connection;

- between two electric interfaces of a VoIP connection.

The present document covers narrowband connections and to some extent wideband connections. Besides gateways and handset terminals hands-free terminals are addressed. Conference configurations are out of scope of the present document.

2 References

For the purposes of this Technical Report (TR) the following references apply:

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

[1] 2nd ETSI VoIP Speech Quality Test Event, Test Specification, Version 2.01, ETSI Plugtests, T-Systems Nova GmbH, Berkom, HEAD acoustics, April 2002.

[2] 1st ETSI VoIP Speech Quality Test Event, Test Specification, ETSI Bake-off Service, Deutsche Telekom Berkom, HEAD acoustics, October 2000.

[3] ETSI TS 101 329-5: "Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON) Release 3; End-to-end Quality of Service in TYPHON systems; Part 5: Quality of Service (QoS) measurement methodologies".

[4] "Proposal for enhancing the test program of the 3rd speech quality test event, HEAD acoustics", ETSI, STQ#15 meeting, 8th to 12th December 2003, Düsseldorf, Germany.

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

[6] ITU-T Recommendation P.501: "Test Signals for Use in Telephonometry".

[7] ETSI EG 201 377-1: "Speech Processing, Transmission and Quality Aspects (STQ); specification and measurement of speech transmission quality; Part 1: Introduction to objective comparison measurement methods for one-way speech quality across networks".

[8] ITU-T Recommendation P.862: "Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs".

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

[10] ITU-T Recommendation P.502: "Objective test methods for speech communication systems, using complex test signals".

[11] ITU-T Recommendation P.58: "Head and torso simulator for telephonometry".

[12] ITU-T Recommendation P.57: "Artificial ears".

[13] ITU-T Recommendation P.64: "Determination of sensitivity/frequency characteristics of local telephone systems".

[14] ETSI TIPHON temporary document 17TD135: "Subjective and objective speech quality evaluation on speech data recorded at the SuperOp 99 event in Hawaii". Sophia Antipolis, March 2000.

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

[16] ITU-T Recommendation P.79: "Calculation of loudness ratings for telephone sets".

[17] ITU-T Recommendation G.122: "Influence of national systems on stability and talker echo in international connections".

[18] F. Kettler, H.W. Gierlich, F. Rosenberger:"Application of the Relative Approach to Optimize Packet Loss Concealment Implementations" DAGA, March 2003, Aachen, Germany.

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

[20] ITU-T Recommendation P.340: "Transmission characteristics and speech quality parameters of hands-free terminals".

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

[22] F. Kettler; F. Rosenberger; H.W. Gierlich: "Speech Quality "Quick Check" for VoIP Terminals", DAGA, March 22.-28., 2004, Strasbourg, France.

[23] ETSI EG 201 377-2: "Speech Processing, Transmission and Quality Aspects (STQ); Specification and measurement of speech transmission quality; Part 2: Mouth-to-ear speech transmission quality including terminals".

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

[25] 3rd ETSI VoIP Speech Quality Test Event, Test Specification, ETSI Plugtests Service, HEAD acoustics, April 2004.

[26] 3rd ETSI VoIP Speech Quality Test Event, Test Specification, ETSI Plugtests Service, Anonymous Test Report Gateways.

[27] 3rd ETSI VoIP Speech Quality Test Event, Test Specification, ETSI Plugtests Service, Anonymous Test Report Terminals.

[28] ETSI TR 102 526: "Speech Processing, Transmission and Quality Aspects: Wideband telephony considerations".

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

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

3 Abbreviations

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

AEC	Acoustic Echo Cancellers
AGC	Automatic Gain Control
BGNT	BackGround Noise Transmission
CN	Comfort Noise
DSS1	Digital Subscriber System No.1
DT	Double Talk
ETSI	European Telecommunications Standards Institute
HATS	Head And Torso Simulators
HFT	Hand-Free Telephone
IP	Internet Protocol
ISDN	Itegrated Services Digital Network
ITU-T	International Telecommunication Union - Telecommunication Standardization Sector
JLR	Junction Loudness Rating
MOS	Mean Opinion Score
MOS-LQO	Mean Opinion Score - Listening speech Quality Objective
MOS-LQON	Mean Opinion Score - Listening speech Quality Objective, Narrowband
MOS-LQOM	Mean Opinion Score - Listening speech Quality Objective, Mixed
MOS-LQOW	Mean Opinion Score - Listening speech Quality Objective, Wideband
NISTnet	National Institute of Standards and Technology network emulation tool
NLP	Non-Linear Processor
OLR	Overall Loudness Rating
PESQ	Perceptual Evaluation of Speech Quality
PLC	Packet Loss Concealment
RLR	Receiving Loudness Rating
SLR	Sending Loudness Rating
SQTE	Speech Quality Test Event
TOSQA2001	Telecommunications Objective Speech Quality Assessment, version 2001
TMOS	TOSQA Mean Opinion Score

NOTE: Output of TOSQA2001.

VAD	Voice Activity Detection

4 The general structure of ETSI speech quality test events

The goal of the ETSI SQTE's is to evaluate the performance of different VoIP equipment with respect to speech quality and conversational quality. ETSI and ITU-T standardized methods are used in order to determine the different parameters influencing the speech and conversational quality. It is expected that manufacturers bring either IP gateways or IP terminals for the tests. The IP terminals may be operated with handset, headset, in loud-hearing or hands-free mode.

One test lab typically performs speech quality measurements on the VoIP equipment of different manufacturers. These tests are instrumental (objective) and described in the present document. The test program of the 3rd ETSI VoIP SQTE [26], which was an extension to the 1st and 2nd program of the events [1], [2] and the TS 101 329-5 [3] is kept as a basis since it was well accepted by all participants and exact comparisons can be made to the previous events. The program is described in this technical report. A typical test event covers:

- Two days preparation for each manufacturer.

- A **½ day tutorial** with presentations is given by the test lab about speech covering speech quality aspects, testing methods, test signals including appropriate analysis methods and algorithms. The whole program for the testing day and the settings that need to be configured in the equipment are also discussed in the tutorial.

- An exclusive testing day being subdivided in predefined **test sessions** and so-called **"freestyle" sessions**.

- The **test sessions** are carried out with fix test conditions (packet loss rates, jitter conditions etc.) and system settings on those the participating manufacturers have agreed on (e.g. speech coder). Note that this agreement is necessary in order to compare the results.

- During the **freestyle session** each manufacturer can choose the tests or test conditions which he decides to be most interesting for him.

A **consulting part** is offered during the testing day. This shall give guidelines for **system optimization**. It is therefore highly recommended for participants to:

- bring implementations that allow the online configuration of parameters;

- join the test event together with the developers of the signal processing algorithms.

4.1 Tests and test sessions

The Tests

Manufacturers can bring VoIP gateways and VoIP terminals to the test event. The equipment is connected to an IP simulation tool and - in case of VoIP gateways - to an ISDN/E1 simulator providing the appropriate interfaces (e.g. E1, etc.). The interfaces are provided by the test lab and are used at all locations where the event takes place (Asia, EU, etc.). Details are described in the appropriate clauses below. VoIP terminals can be operated and measured in handset, headset, loud-hearing or hands-free mode.

The tests carried out cover **all conversational speech quality aspects** taking into account:

- the one-way speech transmission quality;

- echo performance tests including a detailed evaluation of implemented echo cancellers (gateway echo cancellers, acoustic echo cancellers, echo suppression);

- interactive double talk tests; and

- the transmission quality in the presence of background noise.

The tests check common requirements and pass/fail criteria in telephonometry. Moreover the tests are designed to identify parameters which may lead to auditory perceived conversational quality degradation. On the one hand the tests therefore provide an important, **detailed quality description** of the current implementation for the manufacturers. On the other hand the test event may even provide more: The specific tests analyze the relevant parameters determining this current quality. The results and discussions can therefore also be used to **optimize the performance** (see also the "Consulting part" of the Test Event).

The Test Sessions

During some **test sessions** the types of test, the test conditions and the necessary VoIP equipment configurations are pre-defined. During the preparation of the test event manufacturers have to agree on these conditions for the tests sessions in order to guarantee the comparability of results. A maximum possible number of manufacturers have to agree on identical test conditions and settings for the comparison process.

The Freestyle Sessions

The **freestyle sessions** give the manufacturer the opportunity to choose the tests and test conditions, which he decides to be of most interest to him. They can be used e.g. to measure the performance of an additional speech coder, evaluate the influence of PLC, VAD or CN injection, a more detailed evaluation of the implemented echo cancellers or any other aspect covered by the prepared tests and test systems. The freestyle session is scheduled for one of the afternoon sessions, given the manufacturers the opportunity to get an overview about the testing capabilities provided during the event before deciding.

The Consulting Part

Due to the high benefit for the manufacturers during the previous SQTEs the tests emphasis the **consulting part** during the one day test session in order to give guidelines for system optimization. Implementations allowing the free configuration of parameters are recommended. Moreover manufacturers should consider joining the test event together with the developers of the signal processing algorithms in order to discuss details, benefits and drawbacks of the individual implementation and consider further improvements.

4.2   Test Reports

The results for each manufacturer are summarized in a detailed **individual test report** together with listening examples on CD. In addition an **anonymous test report** is published after the event containing the results of all participants including average scores, minimum and maximum scores. Anonymity needs to be guaranteed, the anonymous report can therefore only contain the results if a minimum number of participants agreed on identical test conditions and settings during the event (see also "Test Session").

This anonymous test report and the experiences gained are also fed back into the **standardization bodies** within ETSI looking after the improvement of quality aspect in the voice transmission area. The anonymous test report is made available to the public by issuing it as a Technical Report by ETSI TC STQ.

5   Test description

5.1   General Test Description

The tests are based on the test program successfully used already in previous SQTEs [1] and [26]. The test plan is subdivided into different parts, measurements are done with three basic configurations:

- Electrical - Electrical Connection (IP gateway to IP gateway);

- Electrical - Acoustical Connection (IP gateway to IP terminal);

- Acoustical - Acoustical Connection (IP terminal to IP terminal).

For each configuration the measurements are conducted using two kinds of input signals:

- speech samples, used to calculate MOS-LQO values (MOS-LQON or MOS-LQOM for narrowband systems, MOS-LQOM for wideband systems) (see note) according to ITU-T Recommendation P.800.1 [5];

- speech like test signals according to ITU-T Recommendation P.501 [6].

NOTE: In the 2nd to 4th ETSI VoIP test events MOS-LOQ was used when applying P.862 [8]. This is replaced by MOS-LQON according to the latest revision of P.862.1 [24]. This value can be transformed in MOS-LQOM based on the information given in figure 3 of TR 102 526 [28].

The testlab provides the test systems, the electrical and acoustical interfaces and the analysis tools. ***Electrical, acoustical or combined electrical/acoustical end to end measurements*** are performed. In order to reproduce realistic conditions for acoustical end to end quality measurements subscribers are substituted by dummy heads (Head And Torso Simulators, HATS [11]) during the tests, each equipped with an artificial mouth and artificial ears (type 3.4 according to [12]). The positioning of handsets is made according to ITU-T Recommendation P.64 [13].

The speech samples are to be provided in German and English and transmitted over the connections. For all conditions, identical speech material (samples of about 30 seconds) is used to achieve a good comparability between different test conditions. These recorded samples are analysed using the Telecommunications Objective Speech Quality Assessment method TOSQA2001 [1], [2],[7]. These measures lead to objective TMOS values. This analysis method has already been used in the 1st, 2nd and 3rd ETSI VoIP Speech Quality Test Event [1], [2] and [26]. In addition to the TOSQA2001 analysis, speech quality measures according to ITU-T Recommendation P.862 [8] and [24] are performed for those scenarios where ITU-T Recommendation P.862 [8] is suitable.

Fundamental to all objective tests are subjective evaluations. A good correlation between the subjective methods and the objective method TOSQA 2001 was confirmed during the first event. ITU-T recommends P.862 (PESQ) which was proven to give a good correlation between subjective results in case of electrical to electrical connections. P.862 [8] was not available during the first ETSI VoIP test event, these tests were added later.

Auditory (subjective) testing is required if the objective tests are not validated for the equipment to be tested. If auditory tests are requested, additional speech material according to ITU-T Recommendation P.800 [9] is to be processed for the requesting participants. A randomly selected subset of these recordings has to be assessed in listening only tests according to ITU-T Recommendation P.800 [9]. Note that auditory tests are not part of the present document.

Instrumental measurements for the chosen scenarios are carried out using speech like test signals and analysis methods as published and described in [6] and [10]. These signals and methods are specially developed to determine *instrumental quality parameters influencing the conversational quality* like double talk performance, switching characteristics, echo performance, implementation of VAD, quality of background noise transmission, comfort noise characteristics and others.

In addition to these tests with their specific parameters and results as described in detail below, time frames of a daily session are reserved for *manufacturers to chose any condition or system setting to be tested*, measured and analysed ("freestyle session"). This test can be made very flexible and allow to specifically test conditions as defined by the manufacturer within this time frame.

5.2 Tests Based on Instrumental Assessment of Speech Samples

Speech samples that are acquired during the tests are evaluated using instrumental speech quality measures. In principle two test methods, TOSQA2001 and PESQ according to ITU-T Recommendations P.862 [8], P.862.1 [24], P.862.2 [29] and P.862.3 [30] respectively are applied depending on the measurement scenario used.

Both analysis methods lead to a one dimensional test results with a high correlation to auditory perceived speech sound quality for one-way transmission (e.g. MOS-LQO values according to ITU-T Recommendation P.800.1 [5]). These methods have been validated for VoIP transmission scenarios [14] and are therefore applicable for those scenarios using electrical interfaces. TOSQA2001 has been used and validated in the 1st SQTE [2], PESQ has not been available at that time. Both methods, TOSQA2001 and PESQ have been successfully used for quality assessment of recordings carried out at electrical interfaces during the previous SQTE's [1] and [26]. Both methods lead to highly correlated results for these test scenarios [1].

For recordings at the acoustical interface, i.e. for electrical - acoustical scenarios as well as for acoustical - acoustical scenarios, only TOSQA2001 is used.

The result is a one-dimensional score (TMOS or MOS-LQO according to ITU-T Recommendation P.800.1 [5]). It is influenced by parameters like:

- the type of speech coder;

- the type of AGC, VAD and silence suppression at the sending side;

- comfort noise generation at the receiving side;

- the system reaction on packet loss and jitter in the network (e.g. the quality of PLC - packet loss concealment and jitter buffer design);

- in case of terminals being connected (electrical-acoustical setup, acoustical-acoustical setup) the results are further influenced by frequency responses, distortion and other terminal related parameters.

The TMOS or MOS-LQO scores provide a useful, comprehensive quality score for one-way speech transmission but provide little information about the parameter "being responsible" for the quality observed. In order to provide additional information - with the main focus on how to optimize the current quality- detailed tests are carried out during the event using sophisticated test signals and analysis methods according to ITU-T Recommendation P.501 [6] and P.502 [10].

Further information about the detailed test setups can be found in EG 201 377-1 [7] and EG 201 377-2 [23].

5.3 Tests Based on Speech like Test Signals according to ITU-T Recommendation P.501

The overall quality for speech controlled, non-linear or time-variant systems like VoIP scenarios can be separated into:

- one-way transmission quality in sending and receiving direction (listening speech quality as described above);

- echo performance (talking related);

- double talk performance (interactive conversational aspects); and

- quality of background noise transmission.

The combination of these parameters determines the overall quality of the complete system. Tests based on sophisticated test signals and analysis methods were developed to determine the corresponding instrumental parameters. Depending on the interfaces used during the tests (electrical-electrical, electrical-acoustical or acoustical-acoustical, see clause 4 for details) parameters according to the following list are measured [10], [15], [16] and [17], examples given for electrical-electrical setup, i.e. a gateway to gateway connection:

- one-way delay in send and receive direction, echo delay;

- jitter buffer characteristics and control mechanism;

- quality of PLC implementation (packet loss concealment) using the Relative Approach and cross correlation analysis method [18] and [23];

- junction loudness rating JLR and frequency responses, behaviour of implemented AGC (Automatic Gain Control);

- behaviour of VAD (voice activity detection), silence suppression and CN generation;

- switching characteristics, minimum activation level, sensitivity of double talk detection;

- tests of implemented echo cancellers beyond ITU-T Recommendation G.168 [19] variation of ERL between 6 dB and 40 dB and infinite, single talk echo characteristics, double talk echo using specific AM/FM modulated test signals according to [6] and [10];

- switching characteristics of non-linear processor or centre clipper (NLP);

- detailed double talk performance tests;

- background noise transmission tests at idle mode, with near end signal, with far end signal, variation of background noise signal characteristics.

Due to the acoustic characteristics of the human ear - reproduced by the artificial head measurement system with its flexible pinna according to ITU-T Recommendation P.57 [12] - some specific tests are carried out when **IP terminals with handset or headsets** are tested using the acoustical interfaces (electrical-acoustical test setup):

- Sending Loudness Ratings (SLR) and Receiving Loudness Ratings (RLR);

- Terminal Coupling Loss (TCL_w);

- pressure force depended Receiving Loudness Ratings (RLR), Receiving Frequency Responses including the leakage sensitivity for 2N, 8N and 13N application force between handset and artificial ear.

VoIP hands-free terminals require specific tests which include the following tests (electrical-acoustical test setup):

- adjustment of gains in sending and receiving direction (SLR, RLR);

- performance of acoustic echo cancellers (AEC) and echo suppression;

- configuration of implemented level switching;

- characterization (type 1, type 2a, 2b, 2c or type 3) according to ITU-T Recommendation P.340 [20].

The basis for the implementation of the tests is EG 201 377-2 [23]. In the present document information about the detailed test setups and the test procedures can be found.

NOTE: It is recommended to provide the appropriate implementation that allows:

- the free configuration and parameter setting of signal processing components;

- the enabling and disabling capability of signal processing blocks like AEC, echo suppressions, etc.

6 Detailed Test Plan

6.1 Electrical - Electrical Measurements
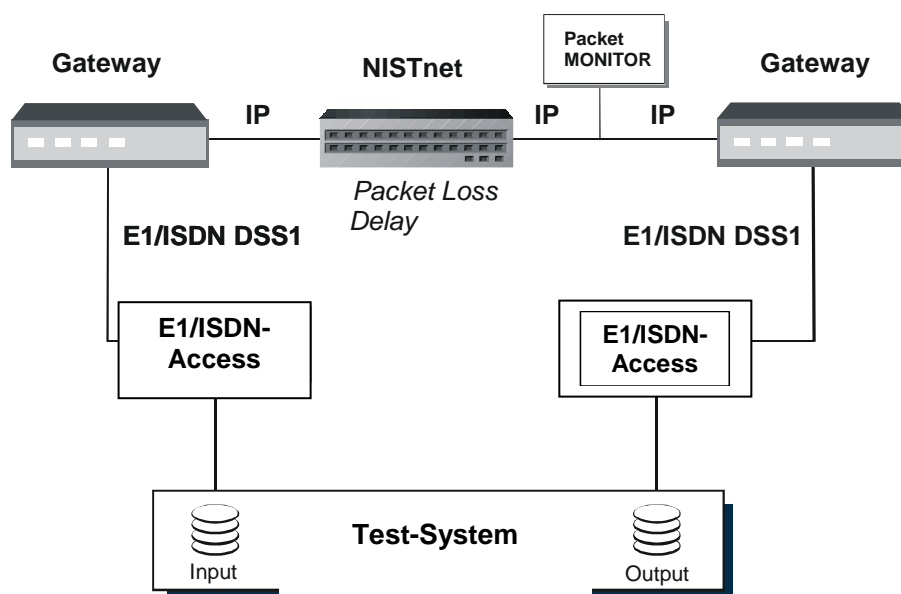
Figure 1: Electrical - Electrical Measurement Setup

Packet loss emulation can be made using either NISTnet with the parameters described or a packet loss simulator performing exactly the same way.

The input signals (speech samples designed according to ITU-T Recommendation P.800 [9] and test signals according to ITU-T Recommendation P.501 [6]) are transmitted and recorded simultaneously, that means that the record process starts at the same time as the transmit process begins. Therefore exact delay assessment is possible.

For all kind of measurements the packet loss generator NIST-Net V.2.0.10 and a packet loss monitor is included in the setup.

In order to ensure comparability to the ETSI VoIP Speech Quality Test, the following IP network conditions are used for "electrical - electrical" measurements.

Table 1: Network Conditions for Electrical - Electrical Measurements (Speech Samples)

Condition	Packet Loss (Equal)	Additional Delay[1]	Delay Variation
0a [3] (VAD)	0	0	No
1a	0	0	No
2a	1 %	0	No
3a	2 %	0	No
4a	3 %	0	No
5a	5 %	0	No
6a	1 %	50 ms	20 ms [2]
NOTE 1: Additional IP network delay is introduced by NISTnet.			
NOTE 2: Delay Variation produced with a Pareto-Distribution and r = 0,5 as provided by NISTnet V.2.0.10.			
NOTE 3: VAD on, all other conditions (1a-6a) tested with VAD off.			

The additional delay in condition 6a is intended to ensure proper jitter (delay variation) generation by NISTnet. In such jitter condition the test network can cause situations where packets are reordered, if the packet size is very small.

Table 2: Network Conditions for - Electrical Measurements (Test Signals)

Condition	Packet Loss (Equal)	Additional Delay[1]	Delay Variation
0b [3] (VAD)	0	0	No
1b	0	0	No
2b	5 %	0	No
3b	0	50 ms	20 ms [2]
4b	5 %	50 ms	20 ms [2]
NOTE 1: Additional IP network delay is introduced by NISTnet.			
NOTE 2: Delay Variation produced with a Pareto-Distribution and r = 0.5 as provided by NISTnet V. 2.0.10.			
NOTE 3: VAD on, all other conditions (1b-4b) tested with VAD off.			

Under these conditions transmission quality parameters:

- can be measured with and without VAD (0b and 1b);

- can be determined without the influence of packet loss and delay variation (condition 1b);

- can be determined and compared to condition 1b separately for packet loss (condition 2b) or delay variation (condition 3b); and

- can be assessed for the combination of both packet loss and delay variation (condition 4b). Again these results can be compared to the other network conditions (condition 1b, 2b and 3b).

These test conditions are in accordance with the 3rd ETSI VoIP Speech Quality Test. The analysed parameters under these conditions can also be directly compared to the corresponding results from the previous events for the two extreme network conditions (1b and 2b). Moreover, these test conditions provide the evaluation of the influences of delay variation and packet loss separately.

It is proposed to carry out 4 different test-settings for each participant during one day (see table 3). Test setting 1 and 2 are fixed settings for each participant with fix gateway conditions. That means all participants have to agree on this condition, e.g. voice codec G.711 with PLC on (Packet Loss Concealment) or (example) G.729 without VAD. This implies that for setting 1 and 2 two appropriate IP gateways have to be provided by the manufacturer. In these two settings 1 and 2 all parameters are tested.

In test setting 3 the participant can decide which further gateway condition (codec, VAD, etc.) he wants to test and in which kind (electrical - electrical, acoustical - electrical) the condition should be tested. Alternatively session 3 can be set up as a "freestyle" test session like session 4. In test setting 4, "freestyle" testing, the input signals, the kind of measurements and the gateway condition can be chosen.

**Table 3: Time Allocation for Electrical - Electrical Measurements
(Gateway to Gateway Configuration)**

No.	Kind of Measurement	Measurement Signal	Set 1 [time]	Set 2 [time]	Set 3 [time]	Set 4 [time]
1	Electric - Electric (figure 1)	Voice	40 min			
		P.501	1,5 h			
2	Electric - Electric (figure 1)	Voice		40 min		
		P.501		1,5 h		
3	Electric - Electric (like set1, diff. settings) or freestyle	Voice			40 min	
		P.501			1,5 h	
4	freestyle test (free choice of kind of measurement and measurement signals)	Voice				x min
		P.501				x min
	Gateway-condition		FIXED G.711	FIXED G.7xx - to be agreed	FREE	FREE
	Day Schedule		AM		PM	

6.2   Electrical - Acoustical Measurements
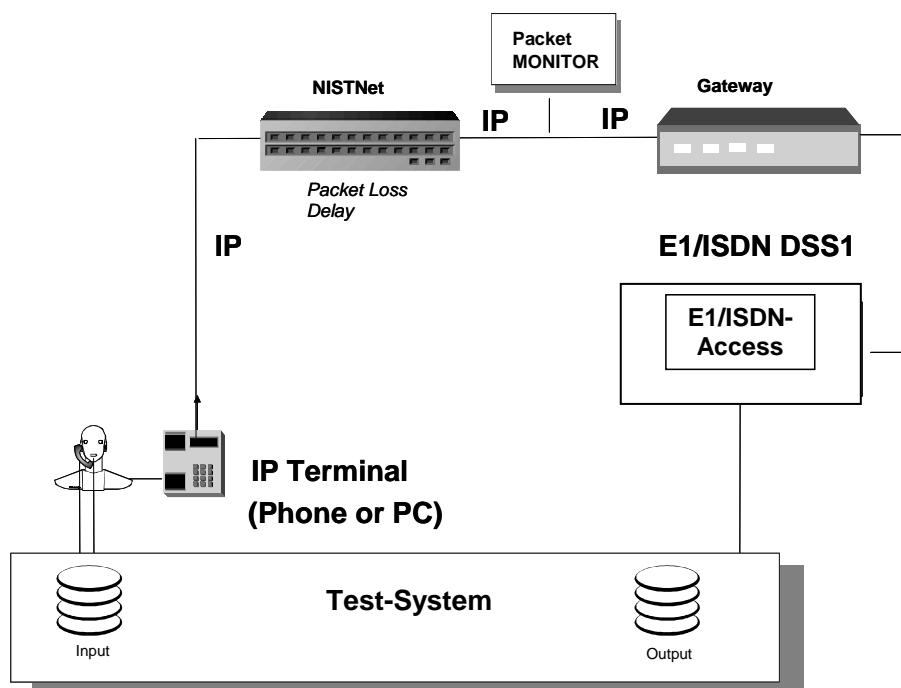
**Figure 2: Measurement Setup Acoustical - Electrical for IP-Terminal
to Gateway Configuration**

For the tests the handsets of the terminals are applied to the HATS using the positioning as described in ITU-T Recommendation P.64 [13] with defined application force. The ITU-T Recommendation P.57 [12] type 3.4 artificial ear is used. The tests can also be carried out in other operation modes like e.g. hands-free. In this case the HATS and the terminal are positioned according to ITU-T Recommendation P.340 [20]. Note that the test room characteristics have a higher influence if the hands-free option is used instead of the handset or headset. In order to evaluate the different kinds of implemented signal processing influencing one-way transmission quality on the one hand and interactive conversation on the other hand, it is recommended:

- to use the handset or headset mode for one way transmission tests using the speech samples and the TOSQA2001 analysis method; and

- to use both modes (handset and hands-free) for evaluating conversational aspects with the P.501 test signals.

The input signals (speech samples designed according to ITU-T Recommendation P.800 [9] and test signals according to ITU-T Recommendation P.501 [6]) are transmitted and recorded simultaneously, that means that the record process starts at the same time as the transmit process begins. Therefore exact delay assessment is possible.

For all kind of measurements the packet loss generator NIST-Net V.2.0.10 and a packet loss monitor is included in the setup.

In order to ensure comparability previous ETSI VoIP Speech Quality Test, the following IP network conditions are used for "electrical - acoustical" measurements.

Table 4: Network Conditions for Electrical - Acoustical Measurements (Speech Samples)

Condition	Packet Loss (Equal)	Additional Delay[1]	Delay Variation
0c [3] (VAD)	0	100 ms	No
1c	0	100 ms	No
2c	0	100 ms	20 ms [2]
3c	1 %	100 ms	No
4c	1 %	100 ms	20 ms [2]
5c	3 %	100 ms	No
NOTE 1: Additional IP network delay is introduced by NISTnet.			
NOTE 2: Delay Variation produced with a Pareto-Distribution and r = 0.5 as provided by NISTnet V. 2.0.10.			
NOTE 3: VAD on, all other conditions (1c-5c) tested with VAD off.			

Note that the additional delay does not influence test results derived by the TOSQA2001 analysis. The additional delay is also intended to ensure proper jitter (delay variation) generation by NISTnet. In such jitter condition the test network can cause situations where packets are reordered, if the packet size is very small.

Table 5: Network Conditions for Electrical - Acoustical Measurements (Test Signals)

Condition	Packet Loss (Equal)	Additional Delay[1]	Delay Variation
0d [3] (VAD)	0	100 ms	No
1d	0	100 ms	No
2d	3 %	100 ms	No
3d	0	100 ms	20 ms [2]
4d	3 %	100 ms	20 ms [2]
NOTE 1: Additional IP network delay is introduced by NIST Net.			
NOTE 2: Delay Variation produced with a Pareto-Distribution and r = 0.5 as provided by NISTNet V. 2.0.10.			
NOTE 3: VAD on, all other conditions (1d - 4d) tested with VAD off.			

These conditions provide the possibility to measure transmission quality parameters:

- with and without VAD (0d and 1d);

- without the influence of packet loss and delay variation (condition 1d);

- separately if influenced by packet loss (condition 2d) or by delay variation(condition 3d); and

for the combination of both packet loss and delay variation (condition 4d). These results can be compared to the other network conditions (condition 1d, 2d and 3d).

These test conditions are in accordance to the 3rd ETSI VoIP Speech Quality Test Event. The parameters analysed under these conditions can also be compared directly to the corresponding results from the previous events for the two extreme network conditions (1b and 2b). Moreover, these test conditions provide the evaluation of the influences of delay variation and packet loss separately.

It is recommended to carry out 4 different test-settings for each participant during one day (see the following table). Test setting 1 and 2 are fixed settings for each participant with fix conditions. That means all participants have to agree on this condition, e.g. (example) voice codec G.711 with PLC on (Packet Loss Concealment) or (example) G.729 without VAD. This implies that for setting 1 and 2 one appropriate IP gateway and the appropriate IP terminal have to be provided by the manufacturer. In these two settings 1 and 2 all parameters are tested.

In test setting 3 the participant can decide which further condition (codec, VAD, hands-free, etc.) he wants to test and in which kind the condition should be tested. Alternatively session 3 can be set up as a "freestyle" test session like session 4. In test setting 4, "freestyle" testing, the input signals, the kind of measurements and the gateway condition can be chosen.

**Table 6: Time Allocation for Electrical - Acoustical Measurements
(Gateway to IP-Terminal Configuration)**

No.	Kind of Measurement	Measurement Signal	Set 1 [time]	Set 2 [time]	Set 3 [time]	Set 4 [time]
1	Electric - Acoustic (figure 2)	Voice	40 min			
		P.501	1,5 h			
2	Electric - Acoustic (figure 2)	Voice		40 min		
		P.501		1,5 h		
3	Electric - Acoustic (different speech coder, handset, loud-hearing or hands-free mode) or freestyle test	Voice			40 min	
		P.501			1,5 h	
4	freestyle test (free choice of kind of measurement and measurement signals)	Voice				x min
		P.501				x min
	Gateway Condition		FIXED G.711	FIXED G.7xx - to be agreed	FREE	FREE
	IP terminal condition		FIXED G.711	FIXED G.7xx - to be agreed	FREE	FREE
	Day Schedule		AM		PM	

6.3   Acoustical - Acoustical Measurements
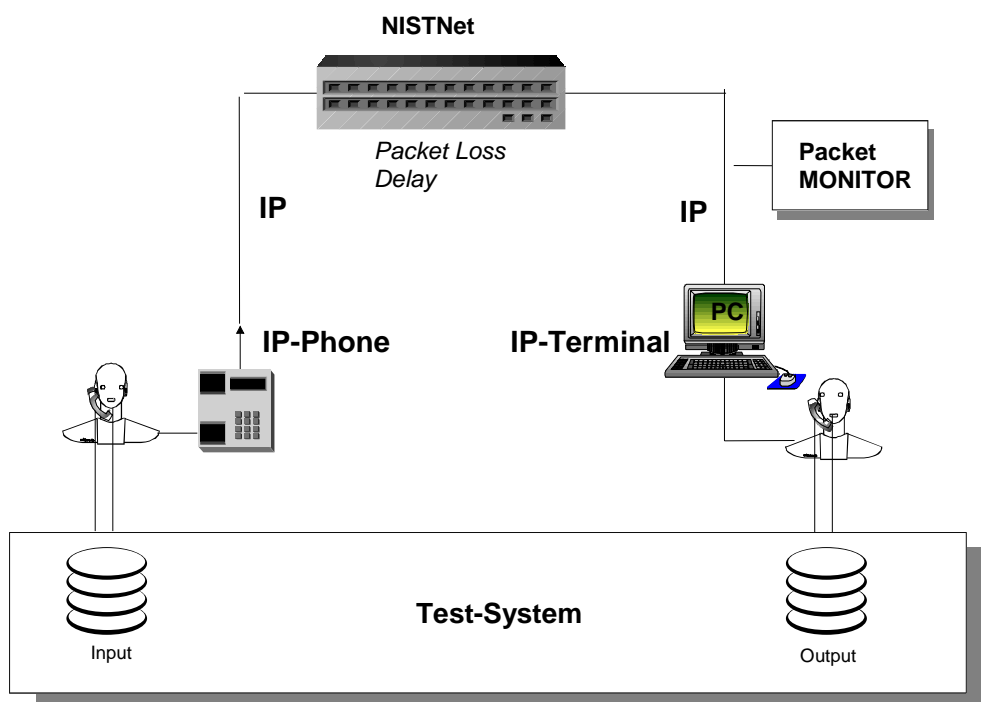
**Figure 3: Acoustical - Acoustical Measurement Setup with IP Terminals
(handsets, headsets or hands-free terminals can be used during the tests)**

For the tests the handsets of the terminals are applied to the HATS using the positioning as described in
ITU-T Recommendation P.64 [13] with defined application force. The ITU-T Recommendation P.57 [12] type 3.4
artificial ear is used. The tests can also be carried out in other operation modes like e.g. hands-free. In this case the
HATS and the terminal are positioned according to ITU-T Recommendation P.340 [20]. Note that the test room
characteristics have a higher influence if the hands-free option is used instead of the handset or headset. In order to
evaluate the different kinds of implemented signal processing influencing one-way transmission quality on the one hand
and interactive conversation on the other hand, it is recommended:

- to use the handset or headset mode for one way transmission tests using the speech samples and the
  TOSQA2001 analysis method;

- to use both modes (handset and hands-free) for evaluating conversational aspects with the P.501 test signals;
  and

- if possible, to evaluate the implemented signal processing in detail using an electrical-acoustical setup
  (see figure 2).

The input signals (speech samples designed according to ITU-T Recommendation P.800 [9] and test signals according
to ITU-T Recommendation P.501 [6]) are transmitted and recorded simultaneously, that means that the record process
starts at the same time as the transmit process begins. Therefore exact delay assessment is possible.

For all kinds of measurements the packet loss generator NIST-Net V.2.0.10 and a packet loss monitor is included in the
setup.

In order to ensure comparability to the previous ETSI VoIP Speech Quality Tests, the following IP network conditions
are used for "acoustical - acoustical" measurements.

Table 7: Network Conditions for Acoustical - Acoustical Measurements (Speech Samples)

Condition	Packet Loss (Equal)	Additional Delay[1]	Delay Variation
0c [3] (VAD)	0	100 ms	No
1c	0	100 ms	No
2c	0	100 ms	20 ms [2]
3c	1 %	100 ms	No
4c	1 %	100 ms	20 ms [2]
5c	3 %	100 ms	No
NOTE 1: Additional IP network delay is introduced by NIST Net.			
NOTE 2: Delay Variation produced with a Pareto-Distribution and r = 0,5 as provided by NISTnet V. 2.0.10.			
NOTE 3: VAD on, all other conditions (1d-4d) tested with VAD off.			

Note that the additional delay does not influence test results derived by the TOSQA2001 analysis. The additional delay is also intended to ensure proper jitter (delay variation) generation by NISTnet. In such jitter condition the test network can cause situations where packets are reordered, if the packet size is very small.

Table 8: Network Conditions for Acoustical - Acoustical Measurements (Test Signals)

Condition	Packet Loss (Equal)	Additional Delay[1]	Delay Variation
0d [3] (VAD)	0	100 ms	No
1d	0	100 ms	No
2d	3 %	100 ms	No
3d	0	100 ms	20 ms [2]
4d	3 %	100 ms	20 ms [2]
NOTE 1: Additional IP network delay is introduced by NIST Net.			
NOTE 2: Delay Variation produced with a Pareto-Distribution and r = 0,5 as provided by NISTnet V. 2.0.10.			
NOTE 3: VAD on, all other conditions (1d-4d) tested with VAD off.			

These conditions provide the possibility to measure transmission quality parameters:

- with and without VAD;

- without the influence of packet loss and delay variation (condition 1d);

- separately if influenced by packet loss (condition 2d) or by delay variation (condition 3d); and

- for the combination of both packet loss and delay variation (condition 4d). These results can be compared to the other network conditions (condition 1d, 2d and 3d).

These test conditions are in accordance to previous ETSI VoIP Speech Quality Test. The parameters analysed under these conditions can also be compared directly to the corresponding results from the 1st Test Event for the two extreme network conditions (1b and 2b). Moreover, these test conditions provide the evaluation of the influences of delay variation and packet loss separately.

It is recommended to carry out 4 different test-settings for each participant during one day (see table 9). Test setting 1 and 2 are fixed settings for each participant with fix conditions. That means all participants have to agree on this condition, e.g. (example) voice codec G.711 with PLC on (Packet Loss Concealment) or (example) G.729 without VAD. This implies that for setting 1 and 2 two appropriate IP terminals have to be provided by the manufacturer. In these two settings 1 and 2 all parameters are tested.

In test setting 3 the participant can decide which further condition (codec, VAD, hands-free, etc.) he wants to test and in which kind the condition should be tested. Alternatively session 3 can be set up as a "freestyle" test session like session 4. In test setting 4, "freestyle" testing, the input signals, the kind of measurements and the gateway condition can be chosen.

**Table 9: Time Allocation for IP Terminal to IP Terminal Configuration**

No.	Kind of Measurement	Measurement Signal	Set 1 [time]	Set 2 [time]	Set 3 [time]	Set 4 [time]
1	Acoustic - Acoustic (figure 3)	Voice	40 min			
		P.501	1,5 h			
2	Acoustic - Acoustic (figure 3) (like 1 but with different settings)	Voice		40 min		
		P.501		1,5 h		
3	Acoustic - Acoustic (figure 3, but with different settings) or Electric - Acoustic (figure 2 (see note)) or freestyle test	Voice			40 min	
		P.501			1,5 h	
4	freestyle test (free choice of kind of measurement and measurement signals)	Voice				x min
		P.501				x min
	IP terminal condition		FIXED G.711	FIXED G.7xx - to be agreed	FREE	FREE
	Day Schedule		AM		PM	
NOTE:	Due to the acoustical characteristics in an acoustical - acoustical measurement setup (like side-tone in the handset, the acoustical coupling between artificial mouth and artificial ear, etc.) and the fact of having always two transmission characteristics involved - the sending direction of one terminal and the receiving direction of the other terminal - it is recommended to choose an electrical - acoustical setup for some measurements. The electrical - acoustical setup provides the advantage of a digital 4-wire access on the electrical side. For manufacturers planning to bring and test IP terminals it is therefore recommended to additionally provide an IP gateway in order to test the IP terminal in an electrical - acoustical measurement setup.					

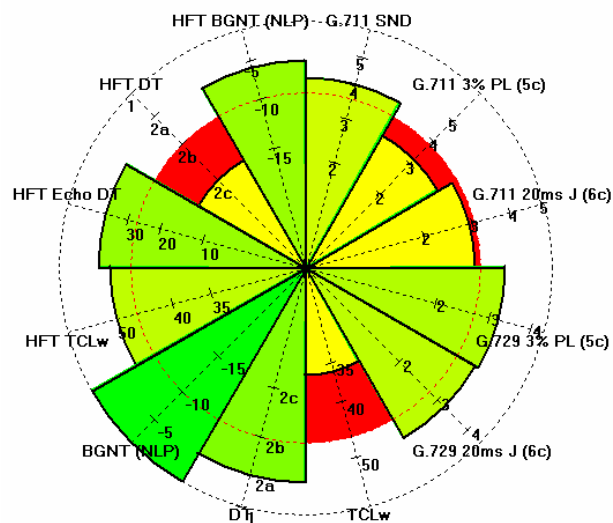# 7   Representation and Documentation of Test Results

In order to provide a condensed overview about the results for each gateway under tests for all speech quality aspects a graphical result representation is derived in accordance to ITU-T Recommendation P.505 (see [21] and [26]). The focus of this conversational speech quality representation is:

- To provide a condensed, "quick and easy to read" overview about the current implementation.

- To provide a variety of measurement results and compare them to the recommended values and numbers in current ITU-T or ETSI standards.

- To provide a comparison to average results from all manufacturers participating in this event.

- To give an indication about strength and weakness of the different implementations.

- To provide detailed enough information for engineering and development in order to improve the performance.

The results are summarized in one diagram, best described as a "Quality Pie". Due to the different parameters for gateways and IP phones the corresponding quality pies differ. But the principle structure described here applies for both realizations.
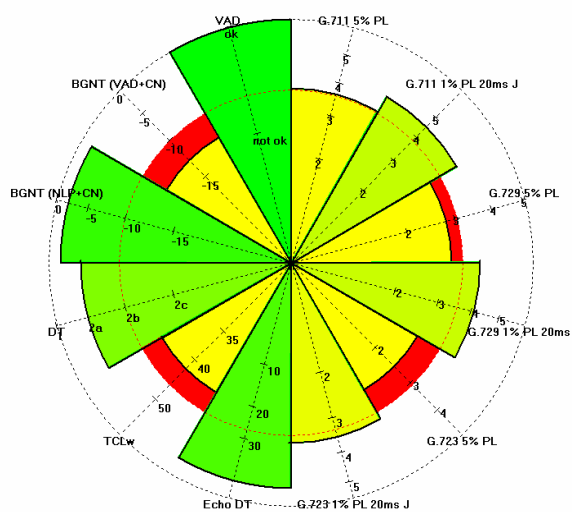
EXAMPLE "Gateway Pie"   EXAMPLE "Terminal Pie"

Each pie slice represents a transmission performance parameter like the codec performance under 5 % packet loss, the echo attenuation under single talk conditions, the quality of background noise transmission or others. The size of each slice represents a measure for the quality of this parameter. Bigger slices indicate a better performance.

7.1   Gateway Pies

The results are summarized in one diagram, best described as a "Quality Pie". Each pie slice represents a transmission performance parameter like the codec performance under 5 % packet loss, the echo attenuation under single talk conditions, the quality of background noise transmission or others. The size of each slice represents a measure for the quality of this parameter. Bigger slices indicate a better performance.

All relevant analyses for this representation are derived from the instrumental measurements. The following example of this "Gateway Pie" does not represent an existing gateway implementation, it is used as an example in order to introduce this result overview.

The following assumptions are made:

- Each parameter is represented by a pie slice.

- The pie slices are independent from each other. Interaction between different parameters like the echo perception due to the combination of echo attenuation and speech distortions (introduced by speech coders) are not considered.

- The size of each slice directly correlates to quality. The pie slice size is area equivalent.

- The minimum requirement for a parameter or the average results from all manufacturers participating during the event is indicated by an inner red circle. If the measured parameter exceeds the recommended requirement or indicates a quality better than the average performance during the test event, the red circle is not visible and overlapped by the pie slice.

- In addition the size of a pie slice is colour coded from yellow (low quality scores or low values) to green (high quality scores or high values).

- The axis scale of each pie slice is parameter dependent.

The following example introduces this conversational speech quality representation and explains, "how to read it". The example does not represent an existing gateway.

EXAMPLE:

conversational aspects like echo behaviour, double talk performance and background noise transmission
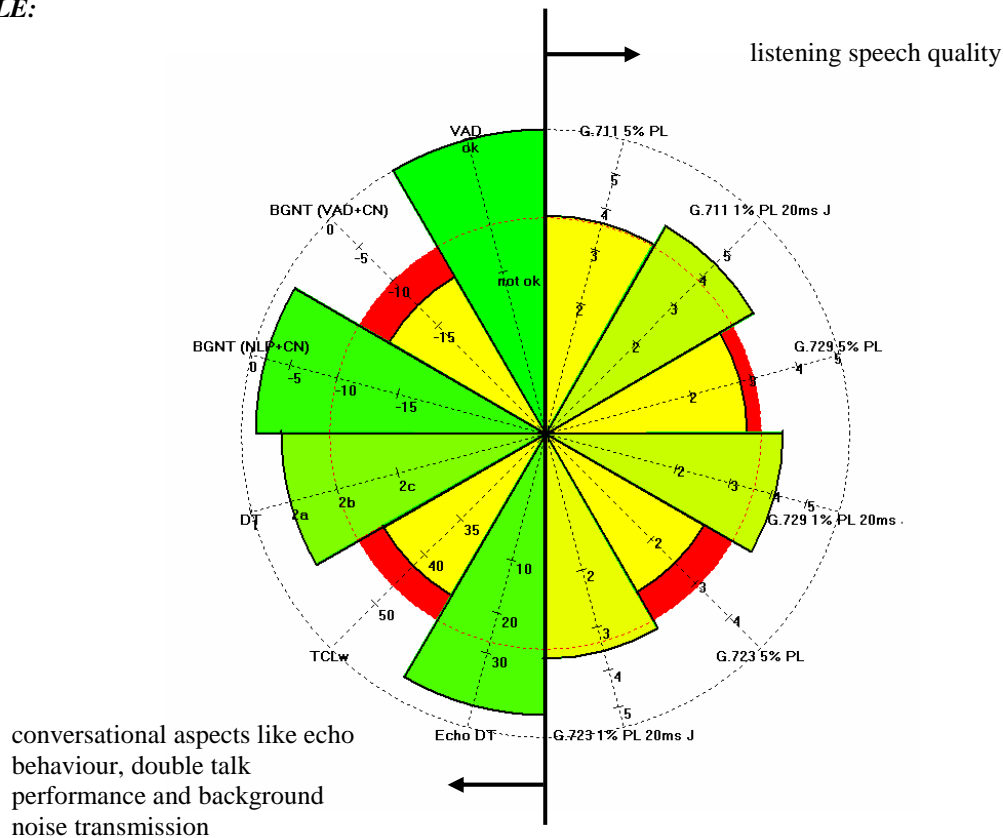
The right hand half of the pie represents the listening speech quality for the different speech codecs. These results consider the influence of packet loss and jitter (G.711, G.729 and G.723 under test condition 5a and 7a).

The left hand side represents the conversational aspects:

- Echo performance during double talk and single talk ("Echo DT", "TCL$_w$").

- The double talk performance ("DT"), characterization in accordance to ITU-T Recommendation P.340 [20].

- The quality of background noise transmission ("BGNT(NLP+CN)"), modulation introduced by the echo suppression unit and its associated comfort noise injection during the application of far end signal.

- The quality of background noise transmission in one way scenarios ("BGNT(VAD+CN)"), modulation caused by voice activity detection of comfort noise injection under single talk conditions.

- The performance of the implemented VAD respectively automatic gain control ("VAD").

The following examples explain each transmission quality parameter ("pie slice") with its scaling and requirement in detail. Again these examples are not derived from real existing gateways.

EXAMPLES (listening speech quality):
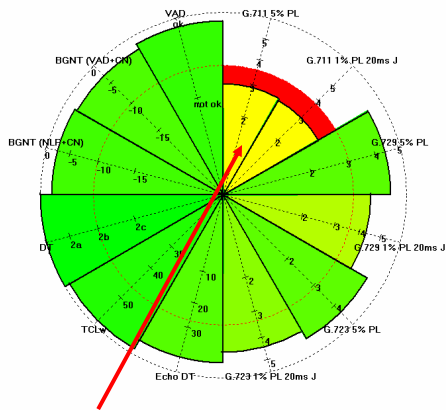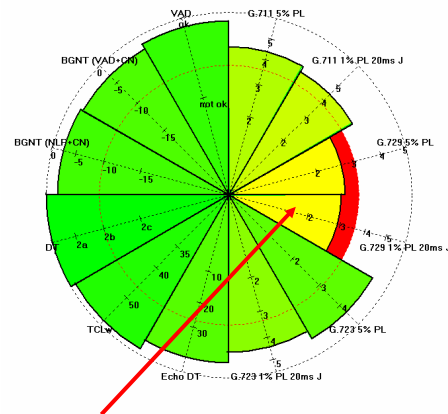
"G.711 listening speech quality below average results under both condition 5a and 7a"
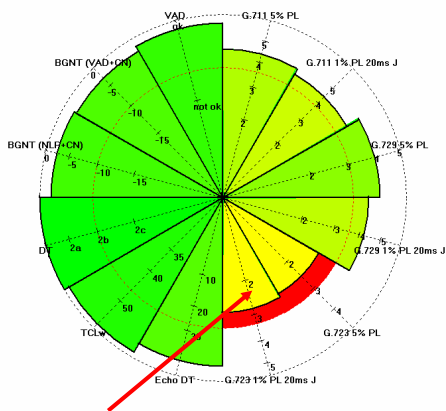
*"G.729 listening speech quality below average results under both condition 5a and 7a"*

The listening speech quality result for each speech coder is represented by two slices, one for the packet loss condition 5a (5 %), one for the jitter condition 7a (20 ms jitter, 1 % packet loss). The values are taken from the MOS-LQO result tables in the individual reports for the G.711, G.729 and G.723 speech coder. Each axis is scaled between 1 and 5 representing the MOS scale.

The limit (radius of the red circle) is given by the average MOS-LQO result over all participants under this test condition (average result from a test, e.g. **tables 5.2**, **5.4** and **5.6 of** [26]). It should be considered that this limit is codec dependent, thus the limits are different for the three speech coders.

"G.723 listening speech quality below average results under both condition 5a and 7a"

EXAMPLES (echo during double talk and single talk):

The echo attenuation during double talk was measured using the AM/FM modulated test signal. These tests were carried out with a 40 dB ERL and 6 dB ERL echo path.

The minimum attenuation (indicated by the inner red circle) is 27 dB. This value, derived from subjective tests can be found in ITU-T Recommendation P.340 [20]. 27 dB echo attenuation during double talk would lead to a full duplex characterization assuming a 100 ms one-way delay in the network. This value can be regarded as a minimum requirement.
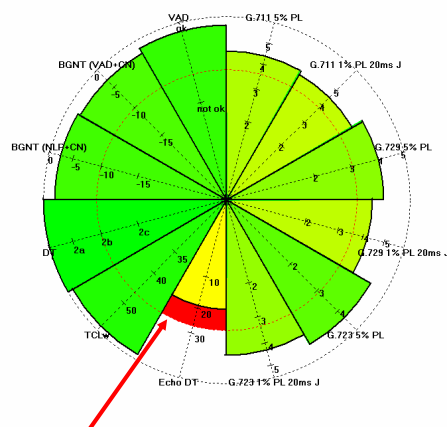
"Echo attenuation under double talk conditions lower than recommended"

The echo loss results are taken form the individual tests. The relevant results for this representation are taken from the 6 dB and the 40 dB ERL measurement.

The lower value from both measurements is used for the pie. The requirement represented by the inner red circle is 46 dB.
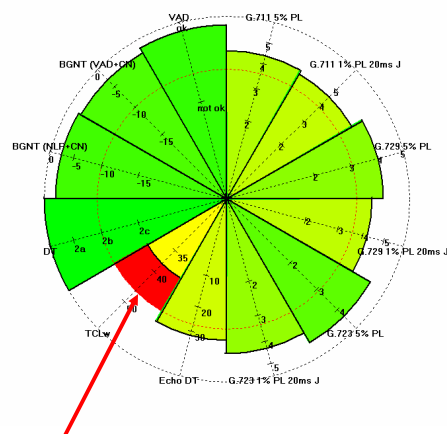
"Echo attenuation according to G.122 under single talk condition below 46 dB"

EXAMPLE (attenuation during double talk, characterization):

The double talk performance is influenced by the attenuation inserted during a double talk period. Most double talk tests are carried out during the event with a 40 dB ERL and 6 dB ERL echo path.

The test results are taken from each manufacturer individually for the pie slice. Examples can be found in **figures 5.27** and **5.32 of [**26**]**. The analyses in these figures represent extreme conditions (high level differences) of the whole sequence.

In accordance to listening examples recorded during the event using real speech, the analysed sequence from the test signal (CS signals) is chosen during the time, where the receive signal and the near end signal are applied with the same level.
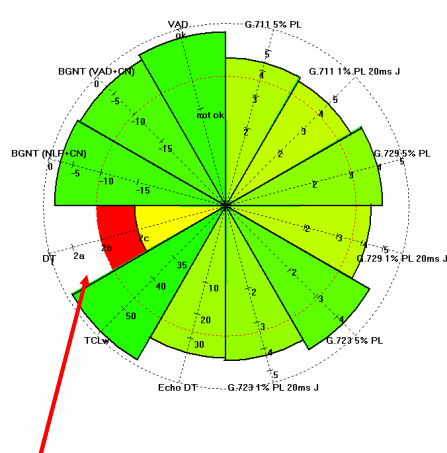
"Double talk performance influenced by level variation"

The level of the transmitted signal is referred to the near end signal level (double talk signal) and analysed vs. time. The average level difference is used to classify the double talk performance.

These results correlate to the listening examples recorded during the event. Level differences between 0 dB and 3 dB lead to a type 1 characterization (full duplex capability). This was achieved by all implementation under test.

Listening examples are recorded under the same conditions (same level in both directions).

More detailed double talk performance tests are carried out during the event with higher test signal level variations (see e.g. **figures 5.27**, **5.28**, **5.32** and **5.33 of [**26**]**). These analysis results may provide useful information for manufacturers in order to optimize double talk performance.

EXAMPLE (quality of background noise transmission with far end signal):

During the application of far end signals the echo suppression unit may introduce audible and disturbing noise modulation (level variation).

The relevant tests can be found in **figures 5.34** and **5.24 of [**26**]**. The level difference between the transmitted signal with and without the application of far end signals is measured.

This difference should not exceed 10 dB, neither for the pub noise nor for the café noise.
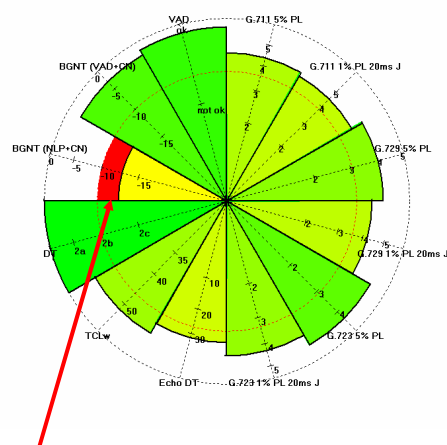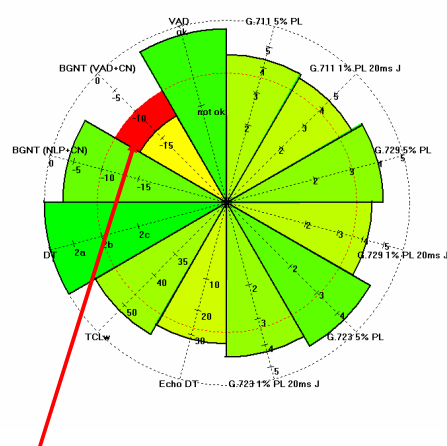

"Background noise modulation introduced by echo suppression and/or comfort noise generation too high"

EXAMPLE (quality of realistic background noise transmission):

Realistic background noise scenarios like the pub noise or the café noise used during this test event should be transmitted without significant level variation. The relevant tests can be found in **figures 5.21** and **5.22 of [**26**]**. The level difference between the transmitted signal with and without VAD is measured.

This difference should not exceed 10 dB, neither for the pub noise nor for the café noise.

"Background noise modulation introduced by VAD or comfort noise generation too high"

EXAMPLE (VAD and AGC test):

The level of a transmitted test signal should follow the original test signal level, if VAD is enabled. Comfort noise -if implemented- should be level adaptive. The relevant analyses can be found in **figure 5.14 of** [26].

The level of the transmitted signal should meet the tolerance scheme in **figure 5.14 of** [26]. This tolerance scheme was derived from test results during the 2nd ETSI SQTE [1].
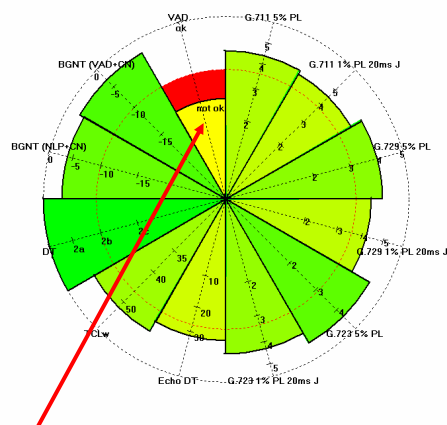
"Level of transmitted signal violates the tolerance scheme"

7.2   Terminal Pies

As for the gateways the results are summarized in one diagram, best described as a "Quality Pie". Each pie slice represents a transmission performance parameter like the codec performance under 3 % packet loss, the echo attenuation under single talk conditions, the quality of background noise transmission or others. The size of each slice represents a measure for the quality of this parameter. Bigger slices indicate a better performance.

All relevant analyses for this representation are given in this report above. The following example of this "IP Terminal Pie" does not represent an existing IP phone implementation, it is used as an example in order to introduce this result overview.

The following assumptions are made:

- Each parameter is represented by a pie slice.

- The pie slices are independent from each other. Interaction between different parameters like the echo perception due to the combination of echo attenuation and speech distortions (introduced by speech coders) are not considered.

- The size of each slice directly correlates to quality. The pie slice size is area equivalent.

- The minimum requirement for a parameter or the average results from all manufacturers participating during the event is indicated by an inner red circle. If the measured parameter exceeds the recommended requirement or indicates a quality better than the average performance during the test event, the red circle is not visible and overlapped by the pie slice.

- In addition the size of a pie slice is colour coded from yellow (low quality scores or low values) to green (high quality scores or high values).

- The axis scale of each pie slice is parameter dependent.

The following example introduces this conversational speech quality representation and explains, "how to read it". The example does not represent an existing IP phone.

EXAMPLE:

conversational aspects,
hands-free - echo and double talk
performance background noise
transmission

handset
listening speech quality

conversational aspects, *handset* - echo
and double talk performance,
background noise transmission
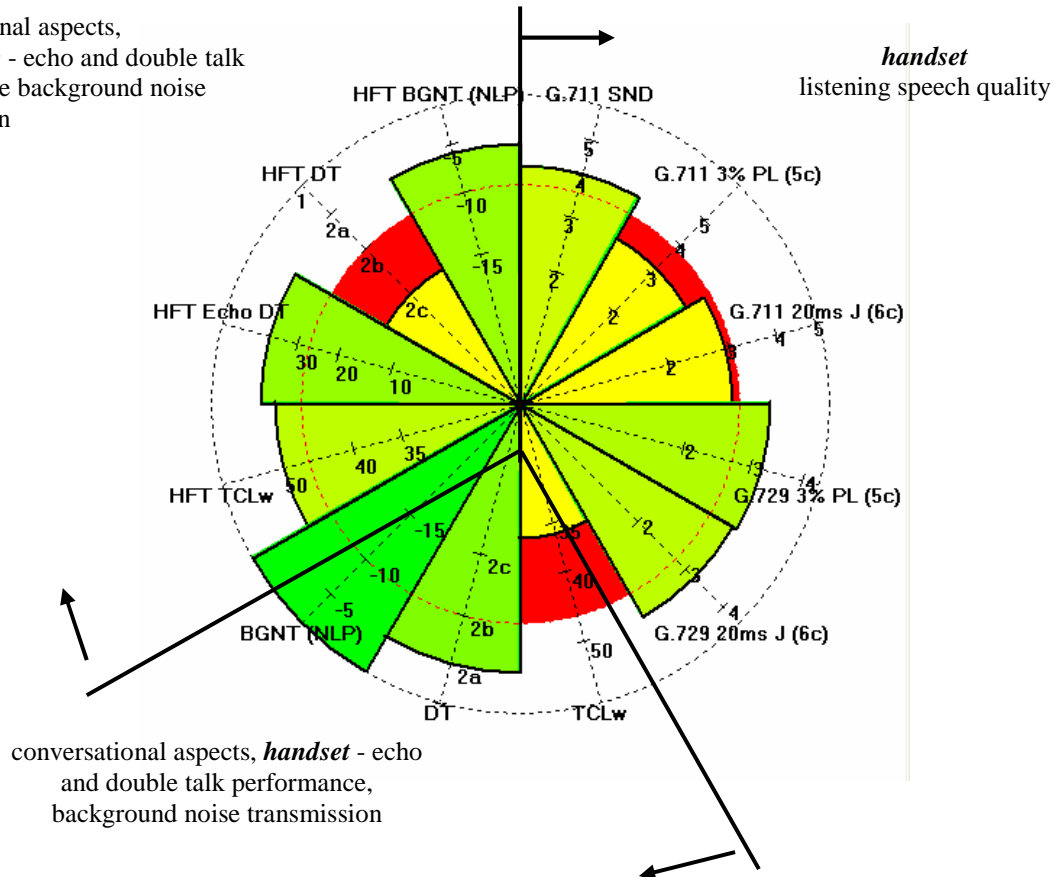
The right hand half of the pie represents the listening speech quality in handset mode for the different speech codecs. The sending direction (with G.711 speech coder) is considered in the first pie slice. The other results were measured in receiving direction with 8 N pressure force between the handset and the artificial ear and cover the influence of packet loss and jitter (G.711 and G.729 under test condition 5c and 6c).

The left hand side represents the conversational aspects in handset mode and in hands-free mode like:

- The echo performance during single talk ("TCL$_w$" respectively "HFT TCL$_w$").

- The echo performance under double conditions (only in hands-free mode, "HFT Echo DT").

- The double talk performance, characterization in accordance to ITU-T Recommendation P.340 [20] for handset and hands-free mode ("DT", respectively "HFT DT").

- The quality of background noise transmission in sending direction during the application of a far end signal, modulation caused by echo suppression or comfort noise injection ("BGNT(NLP)" respectively "HFT BGNT(NLP)".

The following examples explain each transmission quality parameter ("pie slice") with its scaling and requirement in detail. Again these examples are not derived from real existing IP phones.
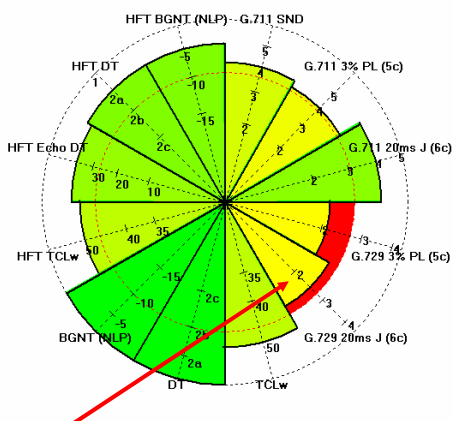
EXAMPLES (listening speech quality):
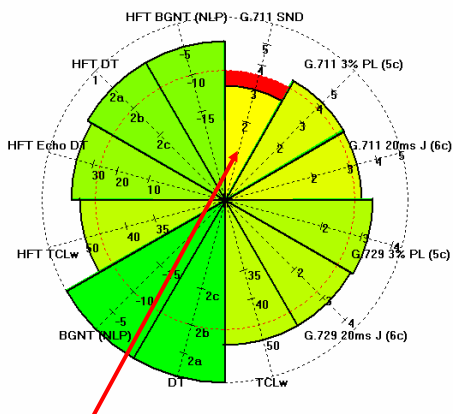
"Listening speech quality in sending direction using G.711 (handset) below average"
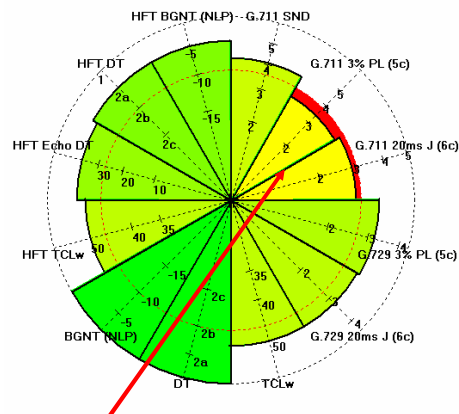
"G.711 listening speech quality in receiving direction (handset) below average results under both conditions 5c and 6c"

"G.729 listening speech quality in receiving direction (handset) below average results under both conditions 5c and 6c"

The listening speech quality result measured in sending direction is represented by the first slice. In receiving direction each speech coder is represented by two slices, one for the packet loss condition 5c (3 %), one for the jitter condition 6c (20 ms jitter, 1 % packet loss). The values are taken from the TMOS result tables in the individual reports for the G.711 and G.729 speech coder. Each axis is scaled between 1 and 5 representing the MOS scale.

The limit (radius of the red circle) is given by the average TMOS result under this test condition (average result see e.g. **table 5.1 of** [27]). It should be considered that these limits are different for each test condition and each speech coder.

EXAMPLES (echo during single talk and double talk):

The echo loss results (TCLw) are derived from each manufacturer. Examples can be found in **table 5.4** for handset mode and in **table 5.9** of [27] for hands-free mode in this report.

The requirement represented by the inner red circle is 46 dB.
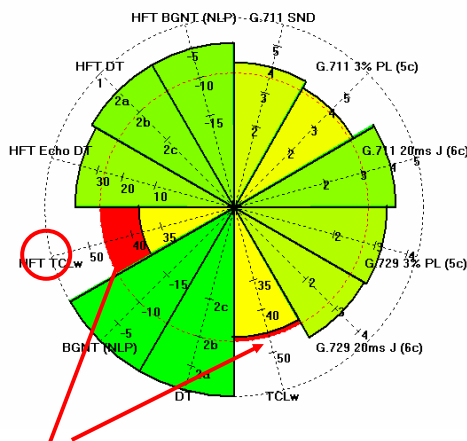
"Echo attenuation according to G.122 under single talk condition below 46 dB"

The echo attenuation of the hands-free implementation during double talk is measured using the AM/FM modulated test signal (see [6]). Examples for such results can be found in [27]. Here the results for the "Quality Pie" are taken from **figures 5.40** to **5.43**.

The minimum attenuation (indicated by the inner red circle) is 27 dB. This value, derived from subjective tests can be found in ITU-T Recommendation P.340 [20]. 27 dB echo attenuation during double talk would lead to a full duplex characterization assuming a 100 ms one-way delay in the network. This value can be regarded as a minimum requirement.
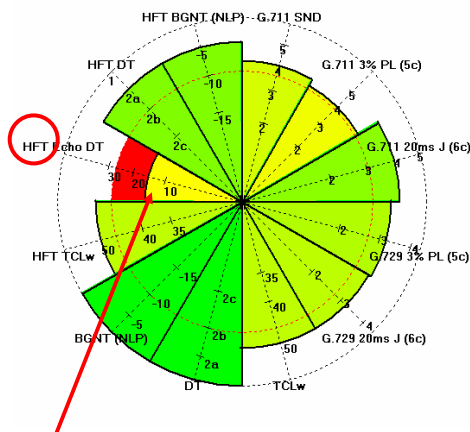

"Echo attenuation under double talk conditions lower than recommended"

EXAMPLE (attenuation in sending direction during double talk, characterization):

The double talk performance is influenced by the attenuation inserted during a double talk period. Different double talk tests are carried out during the tests using different test signals.

The test results for the pie slices for the handset respectively hands-free mode are taken from each manufacturer. Examples can be found in **figures 5.15** and **5.44 of** [27].

The level of the transmitted signal is referred to the near end signal level (double talk signal) and analysed vs. time. In this example the attenuation in sending direction leads to a type 2c characterization in handset and in HFT mode.
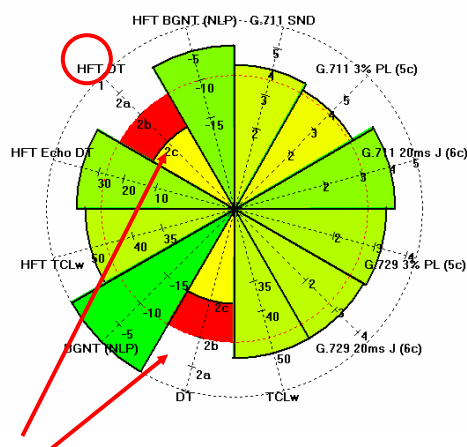

"Double talk performance influenced by level variation leads to "type 2c" characterization"

EXAMPLE (quality of background noise transmission with far end signal):

During the application of far end signals the echo suppression unit may introduce audible and disturbing noise modulation (level variation).

The level difference between the transmitted signal with and without the application of far end signals is measured. Examples for the relevant tests can be found in **figures 5.23** and **5.24 of [**27**]** for the handset mode respectively **figures 5.45** and **5.46 of [**27**]** for the hands-free mode.

This difference should not exceed 10 dB, neither for the pub noise nor for the café noise.
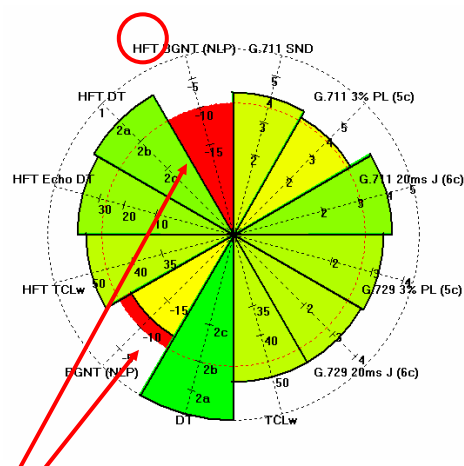
"Background noise modulation introduced by echo suppression and/or comfort noise generation too high"

History

Document history		
V1.1.1	December 2006	Publication