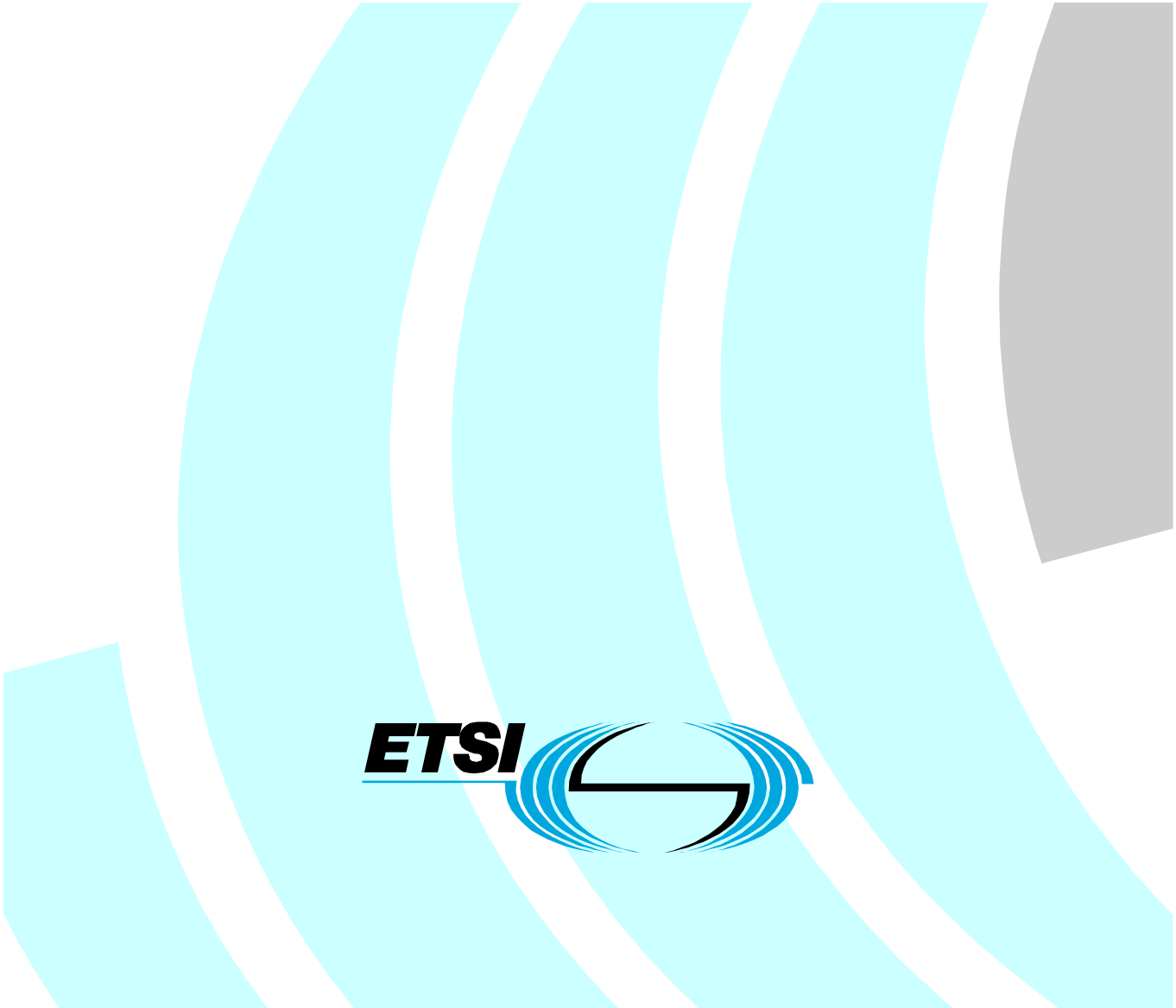
ETSI TR 102 648-3 V1.1.1 (2007-02)

Technical Report

**Speech Processing, Transmission and Quality Aspects (STQ);
Test Methodologies for ETSI Test Events and Results;
Part 3: 2nd ETSI Plugtests Speech Quality Test Event Report**



Reference
DTR/STQ-00079-3

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

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 3 of a multi-part deliverable. Full details of the entire series can be found in part 1 [18].

1 Scope

The present document contains the anonymous Test Report from the 2nd Speech Quality Test Event 2002.

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

[2] 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".

[3] ITU-T Recommendation P.501: "Test signals for use in telephonometry".

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

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

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

[7] 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.

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

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

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

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

[12] ITU-T Recommendation P.830: "Subjective performance assessment of telephone-band and wideband digital codecs".

[13] ITU-T Recommendation P.810: "Modulated noise reference unit (MNRU)".

[14] ITU-T COM12-D41 Deutsche Telekom AG (Q9/12): "Enhancement of P.862 results by post-processing using "TOCQ"".

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

[16] Genuit, K.: "Objective Evaluation of Acoustic Quality based on a Relative Approach", InterNoise Proceedings, Liverpool, UK, 1996.

[17] Sottek, R.: "Modelle zur Signalverarbeitung im menschlichen Gehör", PHD thesis RWTH Aachen, 1993.

[18] ETSI TR 102 648-1: "Speech Processing, Transmission and Quality Aspects (STQ); Test Methodologies for ETSI Test Events and Results; Part 1: VoIP Speech Quality Testing".

[19] 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".

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

[21] ETSI TBR 008: "Integrated Services Digital Network (ISDN); Telephony 3,1 kHz teleservice; Attachment requirements for handset terminals".

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

3 Abbreviations

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

ACQUA	Advanced Communication Quality Analysis
ASL	Active Speech Level
BRI	Basic Rate Interface
CSS	Composite Source Signal
DSS1	European ISDN Protocol
ERL	Echo Return Loss
HATS	Head And Torso Simulator
IP	Internet Protocol
IRS	Intermediate Reference System
ISDN	Integrated Services Digital Network
MNRU	Modulated Noise Reference Unit
MOS	Mean Opinion Score
NIST-Net	Network Simulation Tool from National Institute of Standards and Technology
PABX	Private Automatic Branch Exchange
PLC	Packet Loss Concealment
RTP	Real Time Transport Protocol
TMOS	TOSQA Mean Opinion Score

NOTE: Output of TOSQA.

TOSQA	Telecommunications Objective Speech Quality Assessment
VAD	Voice Activity Detection
VAD	Voice Activity Detection

4 Summary

The present document describes the test methodologies, the assessment methods and the results of the measurements which were carried out during the 2nd ETSI Speech Quality Test Event. The tests were conducted by T-Systems Nova GmbH Berkom (Berkom) and HEAD acoustics GmbH (HEAD acoustics).

The aim of the test event was to determine the speech quality of various Voice over IP equipment under certain IP network conditions. During the test event, speech material as well as measurement data were collected by transferring voice samples and artificial signals across the Voice over IP setup. This material was analysed and the results are reported in the present documentation.

The analysis of the collected data can be split in two parts. In the first part the assessment of the one-way speech quality (listening quality) was performed by instrumental assessments. The one-way speech transmission quality was evaluated by processing real speech samples and analyzing it using the TOSQA2001 algorithm and in addition using the current ITU-T Recommendation P.862 [19] "PESQ". TOSQA2001 and PESQ lead to MOS-comparable results on a scale from 1 to 5 or -1 to 5, respectively. For all measurements at the acoustical interface using artificial ears (e.g. measurements of IP-Phones) only TOSQA2001 was used, because PESQ is not applicable for evaluation of acoustical recorded signals according to P.862 documentation. PESQ and TOSQA were used in all measurement scenarios that were based on measurements at the electrical interfaces at sending and receiving side (e.g. Gateway-Gateway).

In the second part more detailed analyses were carried out. The various strategies of PLC implementations were investigated in detail as well as VAD- and comfort noise implementations. Furthermore additional situations were evaluated. The system performance with background noise during double talk and specifically the echo canceller characteristics. These tests were based on TS 101 329-5 [1] and ITU-T Recommendations P.501 [3], P.502 [4] and G.168 [20].These test show the performance of the different implementations with respect to speech quality in detail and can be used to optimize the system design.

5 Introduction

The tests conducted during the 2nd ETSI TIPHON VoIP Speech Quality Test Event have been optimized and adapted considering the feedback from all participating manufactures after the 1st ETSI VoIP Speech Quality Test Event, concentrating on a more detailed evaluation of specific parameters, optimizing the number of network conditions and discussing the manufacturers implementation based on the test results already during the test event. In order to ensure consistency of the measurements of the 1st ETSI VoIP Speech Quality Test Event and to allow comparisons of results as well, instrumental measures for electrical and acoustical VoIP scenarios were conducted with TOSQA2001. In addition, instrumental measurements were conducted according to ITU-T Recommendation P.862 for those scenarios where P.862 is suitable.

Instrumental measurements were performed for one-way speech transmission as well as double-talk situations by evaluating background noise transmission performance, PLC implementations and echo cancelling characteristics in detail. The measurements conducted by HEAD acoustics were based on test signals and test procedures as described in TS 101 329-5 [1] using test signals and analysis procedures described in ITU-T Recommendations P.501 [3], P.502 [4] and G.168. Furthermore, each company conducted a "free-style" testing block to evaluate specific test conditions. In addition to the tests, an extra ½ day tutorial session was held for each manufacturer. Within ETSI, two bodies are actively involved in this event. ETSI Project TIPHON (Telecommunications and Internet Protocol Harmonization Over Networks) and ETSI Technical Committee STQ (Speech Processing, Transmission and Quality Aspects) look for the improvement of quality aspects in the voice transmission area.

5.1 Time and location

The measurements were conducted during the 2nd ETSI Speech Quality Test Event at ETSI Headquarters in Sophia Antipolis, France from 16th to 24th of April 2002.

6 Responsibilities

The present document contains results of both types of measurements, for measurement at the electrical network-interface of the used PBX as well as at the acoustical interface using a telephone-handset which was connected to the PBX. IP-phones were connected to the IP-network. Technical details for the various types of measurements can be found later in the present document. This clause is intended to provide information about the responsibility of the test labs (Berkom, HEAD acoustics) for the respective measurements, quality assessments and parts of the reports. So, if you have any questions concerning the present document you may contact the responsible test lab. The contact information can be found in the former clause.

The data acquisition for the evaluation of the one-way speech transmission quality at the acoustical interfaces as well as at the electrical interface was performed by HEAD acoustics. The assessment of all TMOS values was carried out by Berkom. The data acquisition and the evaluation of the various transmission parameters, double-talk and background noise performance using artificial test signals were performed by HEAD acoustics, at the electrical interface as well as at the acoustical interface.

The IP-network simulator NIST-Net were provided and controlled by Berkom. For this test event the latest NIST-Net version 2.0.10 was used.

7 Test description

7.1 General test description

The tests are based on the ETSI TIPHON specification as described in the latest version of TS 101 329-5 [1].

The test plan is subdivided in different parts.

Measurements were done with three basic configurations:

- Electrical - Electrical Connection.
- Acoustical - Electrical Connection.
- Acoustical - Acoustical Connection.

Measurements were conducted using two kinds of input signals:

- Real speech samples (English for demonstration, German for Listening Quality measurements).
- Artificial test signals (according to ITU-T Recommendation P.501 [3]).
- Electrical - Electrical Connection: Voice samples and artificial test signals.
- Acoustical - Electrical Connection: Voice samples and artificial test signals.
- Acoustical - Acoustical Connection: Voice samples and artificial test signals.

After recording the samples were analysed by T-Systems Nova, Berkom using the ***Telecommunications Objective Speech Quality Assessment*** method TOSQA2001 [2]. This measure leads to objective TMOS values which allow direct comparisons to the TOSQA values of corresponding test conditions of the 1st ETSI VoIP Quality Test Event. The TOSQA2001 analysis was carried out for all kind of measurement scenarios.

In addition to the TOSQA2001 analysis, speech quality measures according to ITU-T Recommendation P.862 were performed for those scenarios where electrical - electrical scenarios were built up.

Instrumental measurements for the chosen scenarios were performed by HEAD acoustics using sophisticated speech like test signals and analysis methods as published and described in [3], [4] and [12]. The description of the signals is included in clause 7.4. ***Acoustical, electrical or combined acoustical/electrical end to end measurements***. In order to reproduce realistic conditions for acoustical end to end quality measurements, both subscribers were substituted by dummy heads (Head And Torso Simulators, HATS [5]). During the tests, each HATS was equipped with an artificial mouth and artificial ears (type 3.4 according to [6]). The positioning of handsets was made according to ITU-T Recommendation P.64 [8]. Measurements using the electrical interfaces were carried out in the same way by HEAD acoustics based on these test signals and analysis methods. These signals and methods [3] and [4] were specially developed to determine *instrumental quality parameters influencing the conversational quality* like double talk performance, switching characteristics, echo performance and others. These methods are described in clause 7.5.

In addition to these prepared tests with their specific parameters and results as described in detail below, two time frames of a daily session were reserved for *manufactures to choose any condition or system setting to be tested*, measured and analysed ("freestyle test"). This test allowed to test specific conditions as defined by the manufacturer.

7.2 Measurement scenarios

7.2.1 Measurements using electrical interfaces

7.2.1.1 Measurement setup

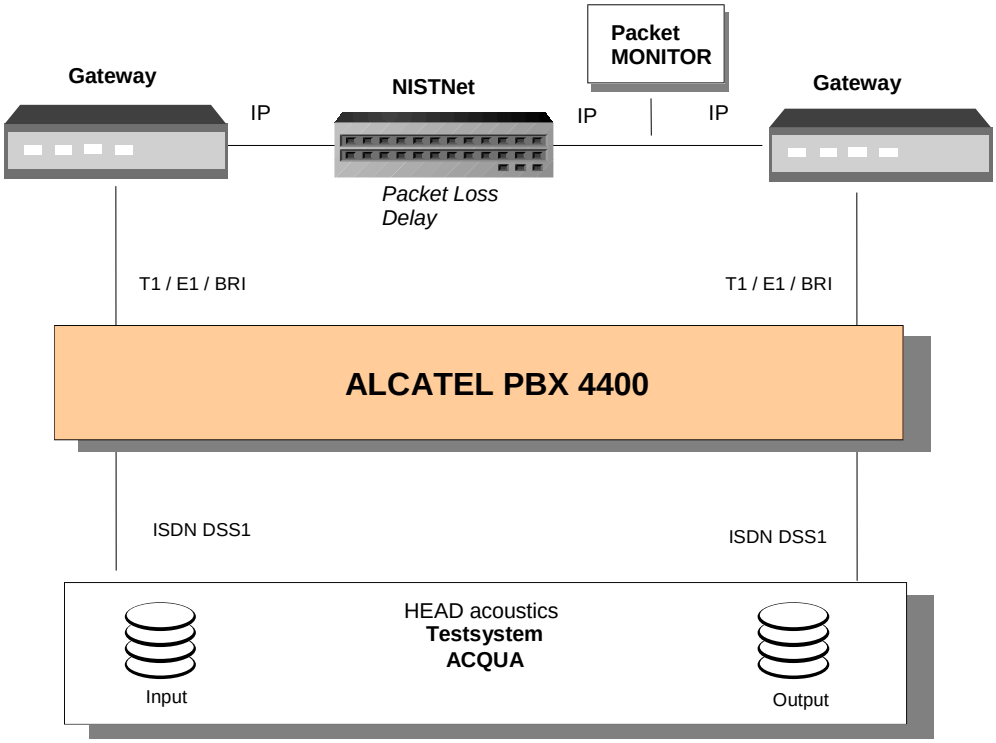


Figure 1: Electrical - Electrical measurement setup

For the electrical-electrical measurements two kinds of input signals were used, speech samples designed according to ITU-T Recommendation P.800 [15] and test signals according to ITU-T Recommendation P.501 [3].

The input signals 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.

7.2.1.2 Measurement conditions

In order to ensure comparability to the 1st ETSI VoIP Speech Quality Test, the following IP network conditions were used for "electrical-electrical" measurements using speech samples according to ITU-T Recommendation P.800 [15].

Table 1: Network conditions for electrical-electrical measurements (speech samples)

Condition	Packet loss (equal)	Additional delay (note 1)	Delay variation
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 (note 2)
NOTE 1: Additional IP network delay is introduced by NIST Net. NOTE 2: Delay Variation produced with a Pareto-Distribution and r = 0,9 as provided by NISTNet V. 2.0.10.			

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.

The measurements using the test signals according to ITU-T Recommendation P.501 [3] were adapted for the 2nd Test Event based on the experience gained during the 1st Test Event and were carried out under the following conditions.

Table 2: Network conditions for electrical-electrical measurements (test signals)

Condition	Packet loss (equal)	Additional delay (note 1)	Delay variation
1b	0	0	No
2b	5 %	0	No
3b	0	50 ms	20 ms (note 2)
4b	5 %	50 ms	20 ms (note 2)
NOTE 1: Additional IP network delay is introduced by NIST-Net. NOTE 2: Delay Variation produced with a Pareto-Distribution and r = 0,9 as provided by NIST-Net V. 2.0.10.			

Under these conditions transmission quality parameters:

- can be determined without the influence of packet loss and delay jitter (condition 1b);
- can be determined and compared to condition 1b separately for packet loss (condition 2b) or delay jitter (condition 3b); and
- can be assessed for the combination of both packet loss and delay jitter (condition 4b). Again these results can be compared to the other network conditions (condition 1b, 2b and 3b).

7.2.2 Acoustical measurements

7.2.2.1 Measurement setup

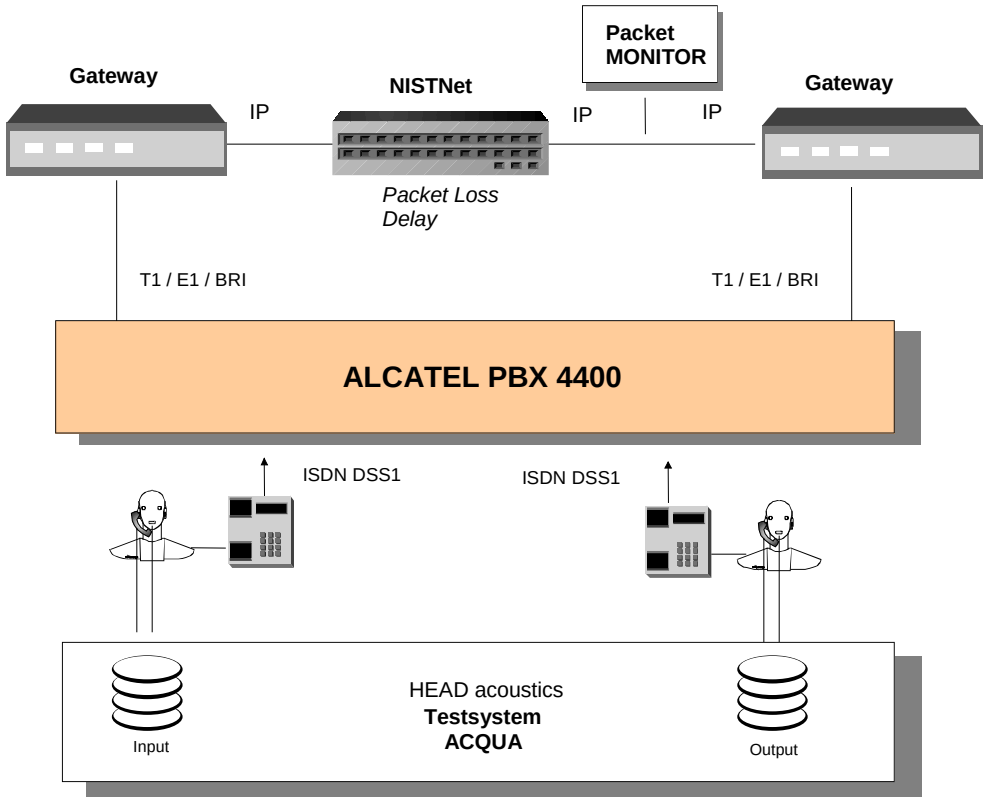
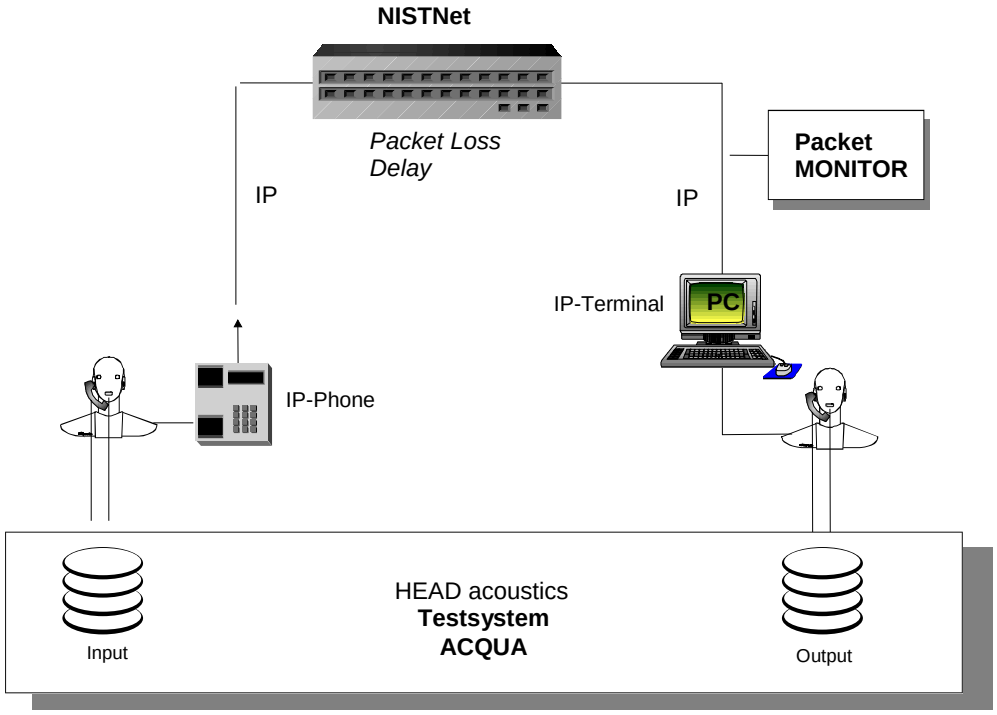


Figure 2: Acoustical - Acoustical measurement setup with reference ISDN terminals



NOTE: Handsets, headsets or hands-free terminals could be used during the tests.

Figure 3: Acoustical - Acoustical measurement setup with IP terminals

The reference terminals provided were standard digital handset terminal ("Europa 10") according to TBR 8 [21]. For all kind of measurements a packet loss generator and a packet loss monitor was included in the setup.

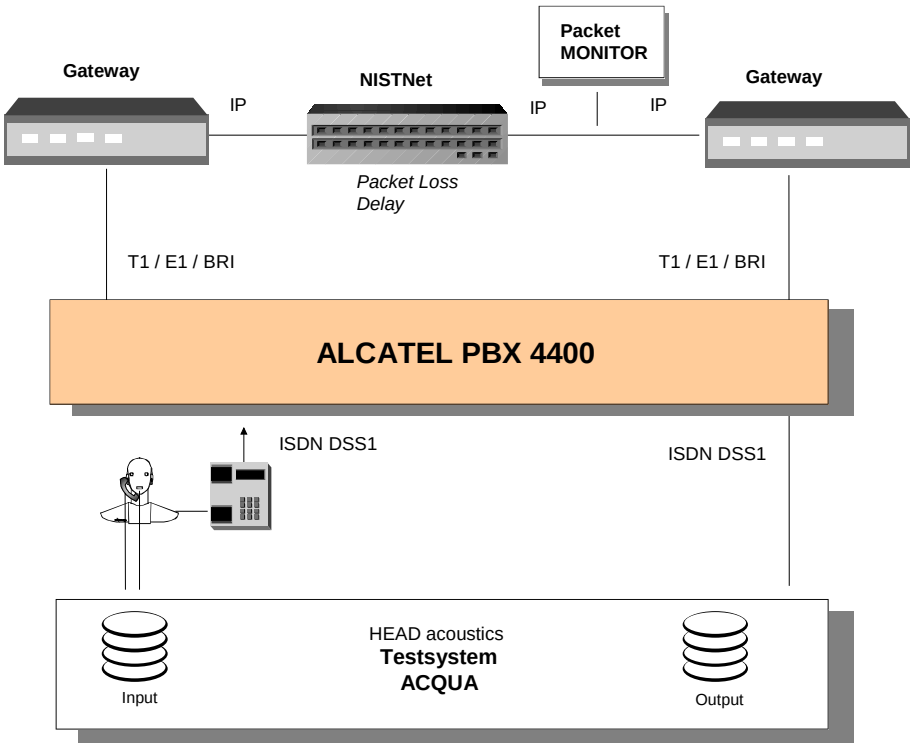


Figure 4: Measurement setup acoustical - Electrical for gateway to gateway configuration

For the tests the handsets of the terminals are applied to the HATS using the positioning as described in ITU-T Recommendation P.64 [8] with defined application force. The test sequences, speech as well as the artificial sequences then were automatically generated by the test system ACQUA and recorded on hard disc.

Instead of the PABX telephone an IP telephone could be used if provided in combination with a gateway which interfaces to the PABX. This is shown in figure 5.

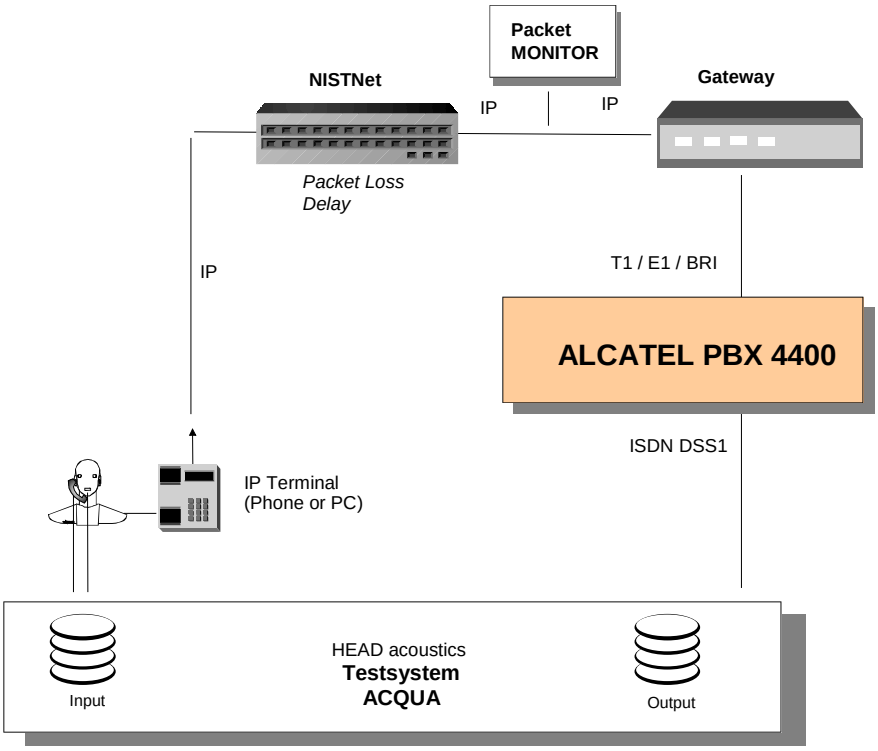


Figure 5: Measurement setup acoustical - Electrical for IP-terminal to gateway configuration

7.2.3 Measurement conditions

In order to ensure comparability to the 1st Test Event, the IP network conditions for all kinds of acoustical measurements ("electrical - acoustical" and "acoustical - acoustical") using speech samples are as follows:

Table 3: Network conditions for all kinds of acoustical measurements (speech samples)

Condition	Packet loss (equal)	Additional delay (note 1)	Delay variation
1c	0	100 ms	No
2c	0	100 ms	20 ms (note 2)
3c	1 %	100 ms	No
4c	1 %	100 ms	20 ms (note 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 = 0,9 as provided by NIST-Net V. 2.0.10.			

The measurements using the test signals as specified in ITU-T Recommendation P.501 [3] were carried out under the following network conditions.

Table 4: Network conditions for all kinds of acoustical measurements (test signals)

Condition	Packet loss (equal)	Additional delay (note 1)	Delay variation
1d	0	100 ms	No
2d	3 %	100 ms	No
3d	0	100 ms	20 ms (note 2)
4d	3 %	100 ms	20 ms (note 2)
NOTE 1: Additional IP network delay is introduced by NIST-Net. NOTE 2: Delay Variation produced with a Pareto-Distribution and r = 0,9 as provided by NIST-Net V. 2.0.10.			

Again these conditions provide the possibility to measure transmission quality parameters:

- without the influence of packet loss and delay jitter (condition 1d);
- separately if influenced by packet loss (condition 2d) or by delay jitter (condition 3d); and
- for the combination of both packet loss and delay jitter (condition 4d). These results can be compared to the other network conditions (condition 1d, 2d and 3d).

7.3 Measurement methodology

This clause describes the measurement procedure in detail, especially the verification procedure that specific adjustments on the behaviour of IP network were achieved.

As displayed in several figures of clause 7.2, a real-time network simulator (NIST-Net) was used to generate the IP network conditions. NIST-Net is a network emulation based on Linux. It is developed by the National Institute of Standards and Technology (NIST). NIST-Net allows a normal PC to act as a router emulating a wide variety of network conditions. NIST-Net is implemented as a module extension to the Linux kernel and is distributed as OpenSource Software.

In terms of packet loss generation this network simulator uses a shaped random number generator to drop packets. To achieve the given percentage of packet loss, a fairly large number of packets need to pass the NIST-Net device. If for example the packet loss rate is configured to 1 % the NIST-Net device would need to drop one packet out of 100. Because of the underlying random number generator it may occur that out of 100 packets 0 or 2 packets will be dropped. The configured packet loss rate will actually be achieved just after a large number of packets (> 1 000) with a certain maturity.

During the 2nd ETSI-VoIP-Test-Event the following parameters have been varied:

- packet loss;
- (fixed) delay;
- jitter (delay variation).

The tests were divided into two separate parts, each with different setups for NIST-Net:

1) acoustical tests (IP-Phones);
2) electrical tests (Gateways).

The actual values for the parameters had the following ranges:

packet loss: 0 %, 1 %, 2 %, 3 %, 5 %;

fixed delay: 0 ms, 50 ms (electrical) / 100 ms (acoustical);

delay jitter: 20 ms.

A matter of special interest, the statistical distribution of the delay variation (jitter) was identified. NIST-Net by default implements the Pareto distribution. This distribution offers a good reproduction of the situation in real wide area networks. The probability for delayed packets is higher than the probability for packets coming earlier than the fixed delay time. A typical distribution of the delay jitter is shown in figure 6. This example is based on the evaluation of about 2 000 RTP packets under condition of 100 ms fixed delay and 20 ms delay jitter (condition 4 c in table 3). The behaviour described above can be observed easily, including the steep slope to the left and the shallower slope to the right side (longer delays).

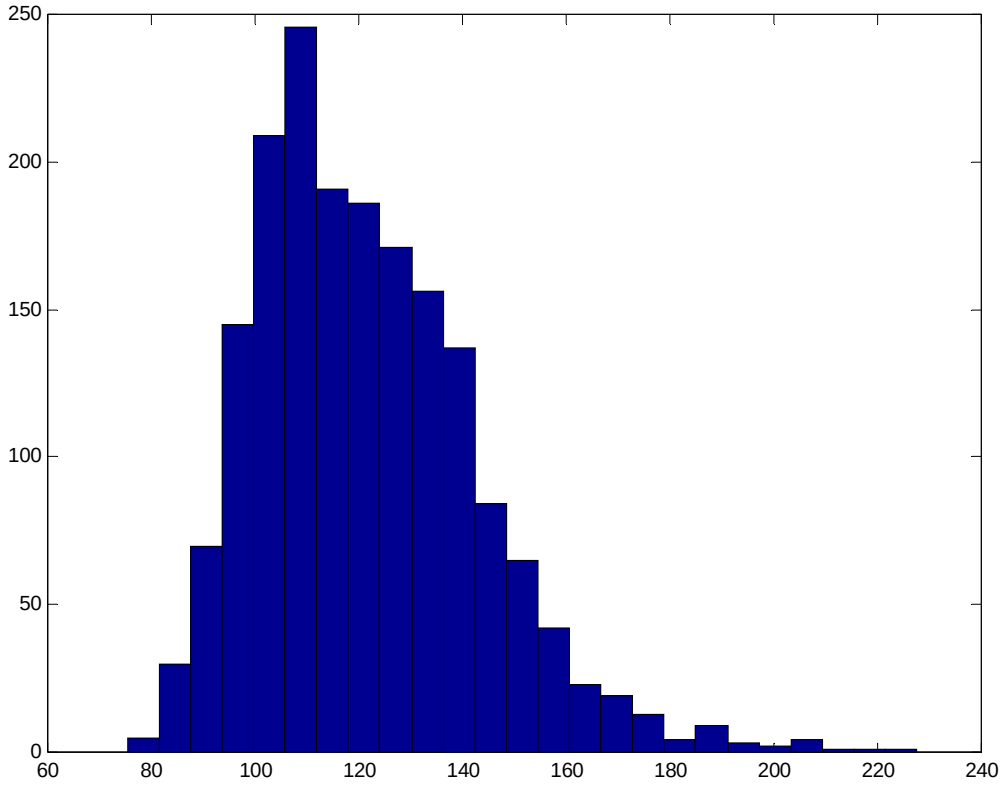


Figure 6: Typical example for the distribution of delay jitter generated by NIST-Net

The speech quality tests have been carried out using NIST-Net's version 2.0.10 running on a Linux-PC (1,6 GHz Pentium4, 256 MB RAM, 100 Mbit/s Ethernet Cards) with SuSE 7.3. They were controlled by NIST-Net's command-line interface in combination with PERL-Scripts developed by Berkom.

Using 4 different voice samples (according to ITU-T Recommendation P.800 [15], 8 seconds each) for one particular condition leads to reasonably accurate results for time invariant voice transmission systems (e.g. PBX test, codec test). Because of the nature of VoIP technology, the VoIP transmission system is time variant, especially in cases of packet loss. Even if it could be ensured that the number of packet losses matches the packet loss rate (by checking the packet losses using a packet monitor), the system is nevertheless time variant in terms of the position of packet losses in the speech signal. To avoid those influences of the location of dropped packets in the speech signal, it was necessary to transmit more speech samples in order to compensate this effect in average. A number of 16 different speech samples, still 8 seconds each, leads to reasonable results and increases the maturity of the results. For transmission 4 speech samples (8 seconds each) were concatenated to one file of 32 s. Four of such 32 s speech files were transmitted and on this basis the packet loss was monitored and controlled.

In order to observe the behaviour of the IP, a packet monitor observes the network and calculates the actual packet loss rate based on RTP sequence numbers. After transferring a 32 s voice sample, the real packet loss rate was checked and if necessary the same speech sample was repeated several times until the required packet loss rate was encountered within 95 % accuracy for all transmitted files in the chosen condition.

7.4 Test signals

7.4.1 Voice signals

The transmitted voice signals are the basis for the instrumental evaluation of one-way speech transmission quality using PESQ and TOSQA. Therefore speech samples in German language were used. In addition to these English speech samples were transmitted and used for presentation of typical effects.

Four concatenated speech files (32 s each) were generated and transmitted. Each of these files contain four different sentence pairs (8 seconds long each) uttered by different male and female speakers. Each of these sentence pairs fulfils the requirements of ITU-T Recommendation P.800 [15] and shows a speech activity factor of about 50 %. The figure 7 shows a typical structure of the used 32 second speech files.

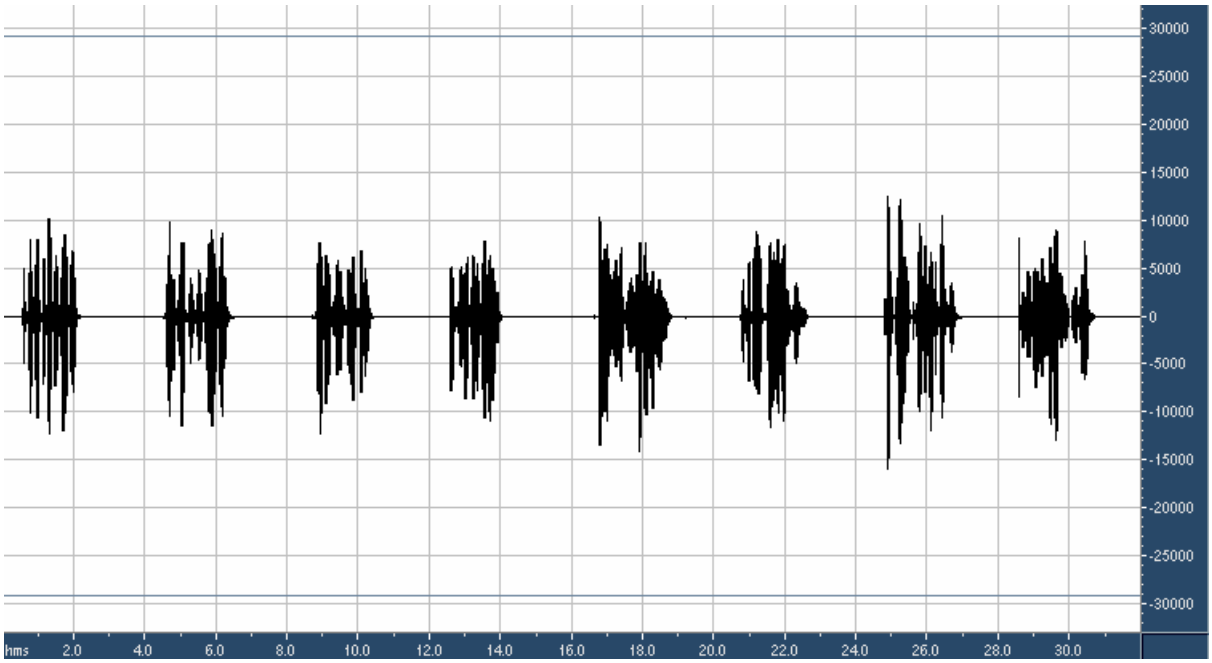


Figure 7: Typical structure of a 32 s speech file for evaluation of one-way transmission quality

For each tested condition, all four of these 32 s files were transmitted. The transmission was repeated until the defined packet loss could be achieved within an accuracy of 95 % for this test condition. For the final evaluation of the speech quality values the 32 s files were divided in the original 8 s sentence pairs. The resulting 16 speech samples were used for instrumental evaluation. The instrumental evaluation was performed for each of the 8 s speech samples separately by the instrumental speech quality assessment methods TOSQA2001 and PESQ. The final result for each test condition was achieved by averaging the 16 individual quality values.

All speech samples which were used as electrical input signals were pre-filtered with a modified IRS(send) filter [12]. The active speech level (ASL, [11]) at the sending side was adjusted in these conditions to -16 dBm0 at 600 Ohm. For all acoustical input signals, unfiltered source material without any band limitation was used. Here the active speech level was 89 dB (SPL) at the mouth-reference-point.

7.4.2 Test signals according to ITU-T Recommendation P.501

The test signals which were used for the objective tests conducted by HEAD acoustics during the event are published and defined in ITU-T Recommendation P.501 [3]. These speech-like test signals represent important characteristics of real speech, e.g. voiced parts such as the vowels in real speech, unvoiced parts such as most of the consonants, power density spectrum or signal modulation vs. time. These signals have the advantage of not being limited to one language or one speaker, but being adapted to measure the relevant implemented technical parameters for speech quality perception. The corresponding analysis methods are described in ITU-T Recommendation P.502 [4].

These test signals were used during the event in order to measure the conversational parameters like delay and delay variation, echo attenuation, switching characteristics, activation sensitivity, double talk performance or the quality of background noise transmission. Some of these measurements compare the results to given requirements, others analyse the current implementation and provide hints for system optimization. Examples are the test for the implemented signal processing like Packet Loss Concealment (PLC), Voice Activity Detection (VAD) and comfort noise generation.

Following is a brief description of these test signals together with some analysis demonstrating the specific characteristics of each signal.

Figure 8 shows the "Composite Source Signal" in the upper window. It consists of 2 signal bursts with a duration of 250 ms each and a pause of approximately 100 ms. Each burst consists of a voiced part, a shaped pseudo random noise sequence. The power density spectrum (lower window in figure 8), derived from the Fourier-Transform of this test signal, reproduces the spectral characteristics of real speech. The pitch frequency and the harmonics of the voiced part of the signal can clearly be seen in the power density spectrum and the spectral representation.

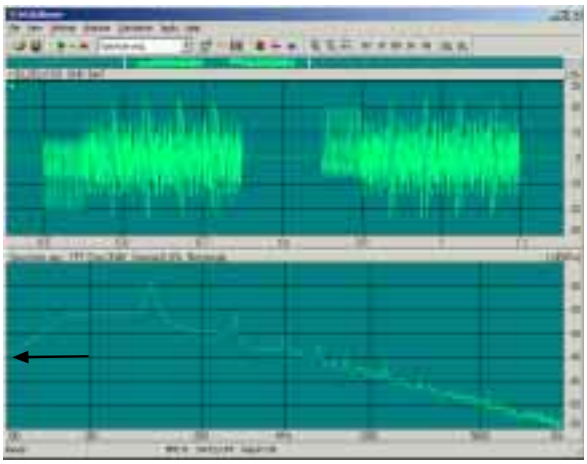


Figure 8: Composite Source Signal (upper window: time signal, lower window: power density spectrum derived) by Fourier transform

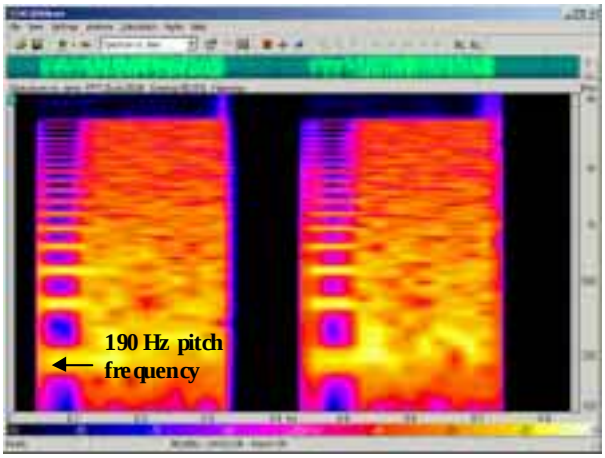


Figure 9: Spectral representation of the Composite Source Signal (time scale on the x-axis, frequency scale between 100 Hz and 5 kHz on the y-axis)

For tests under single talk conditions, this signal is fed into the measurement setup, and the transmitted signal is recorded and analysed. Parameters such as one way transmission delay, switching characteristics and others can be determined using this test signal.

A periodical repetition of this Composite Source Signal with variable signal level for each signal burst is shown in figure 10. This test signal is used to determine the sensitivity threshold and the switching characteristic of Voice Activity Detection (VAD).

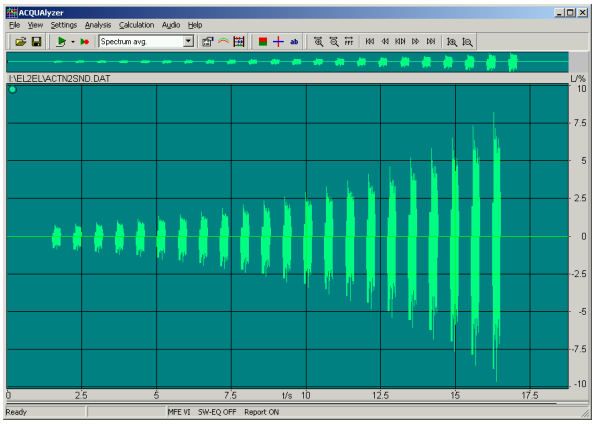


Figure 10: Periodical repetition of the Composite Source Signal to determine the activation sensitivity under single talk conditions

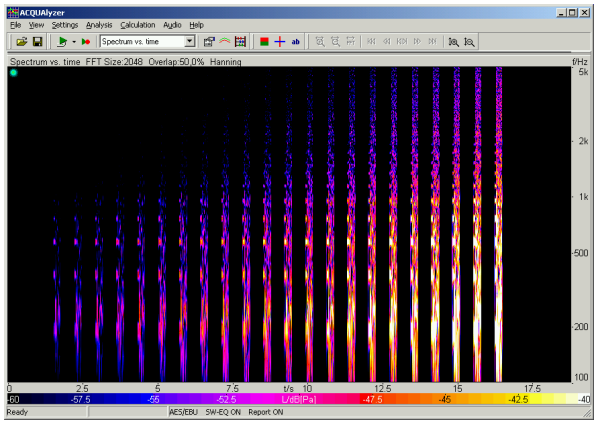


Figure 11: Spectral representation of the test signal (time scale on the x-axis, frequency scale between 100 Hz and 5 kHz on the y-axis)

A test signal developed to determine the behaviour of implemented Automatic Gain Control (AGC) is shown in figure 12. It consists of a periodical repetition of a voiced sound (the same which is used in the voiced part of the CSS in figure 8) with decreasing and increasing level variation vs. time. The duration of the voiced sound is approximately 3 ms, the complete signal length amounts to 10 seconds.

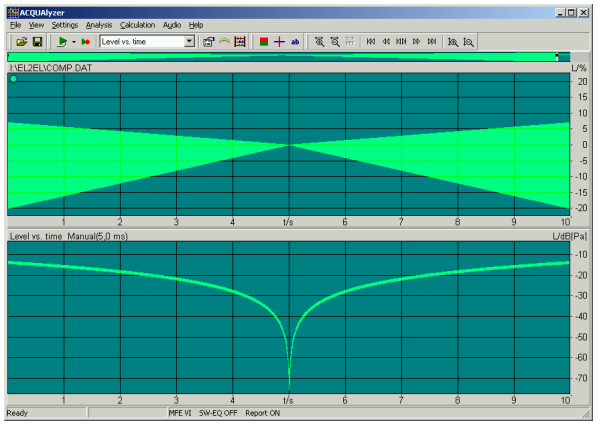


Figure 12: Test signal to measure AGC behaviour (upper window: time signal, 10 s, lower window: signal level vs. time, determined with a time constant of 5 ms)

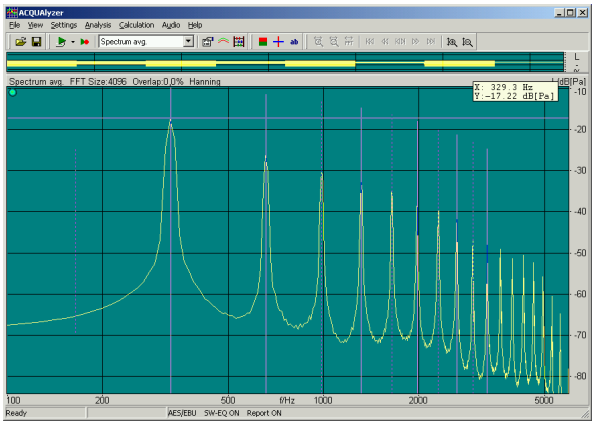


Figure 13: Power density spectrum of the voiced sound with its typical harmonics (pitch frequency approximately 330 Hz)

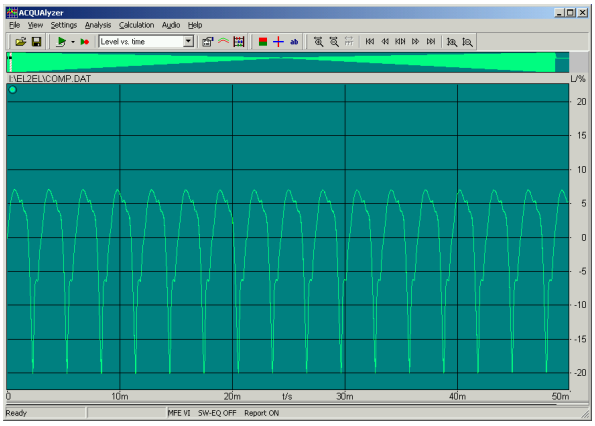


Figure 14: Enlarged time sequence from figure 12 (50 ms) showing in detail the periodical repetition of the voiced sound

Due to its deterministic and periodic characteristic (see an enlarged time sequence in figure 14) the signal is also suited to analyse packet loss, the performance of implemented packet loss concealment and the jitter buffer behaviour.

The evaluation of double talk performance - both subscriber talk simultaneously - requires a second test signal to be applied simultaneously at the opposite transmission path. The two signals that simulate double talk need to be uncorrelated. Figures 15 and 16 show a second uncorrelated Composite Source Signal.

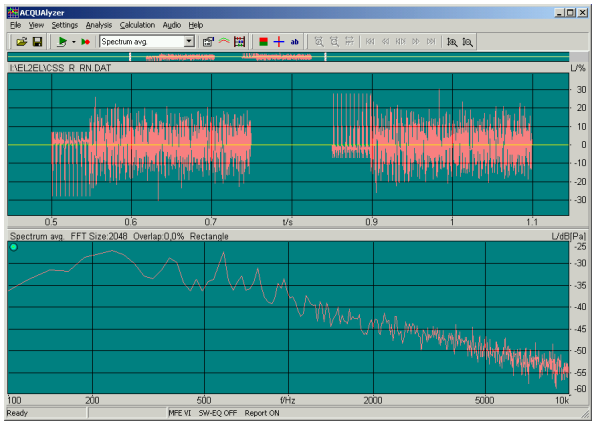


Figure 15: Second -uncorrelated- Composite Source Signal (upper window: time signal, lower window: power density spectrum derived

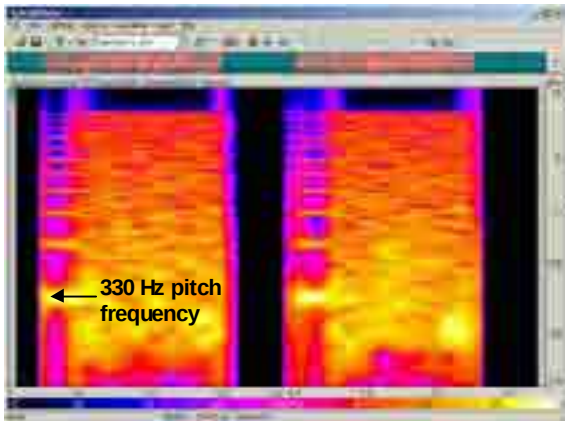


Figure 16: Spectral representation of the Composite Source Signal (time scale on the x-axis, frequency scale between 100 Hz and 5 kHz on the y-axis)

These two Composite Source Signals and its components (like the voiced part) can be combined to evaluate the switching characteristics by activating the two transmission paths. The following figures show one example. The time sequence is given in figure 17 whereas the two windows of figure 18 show the spectral characteristics of both signals. The red coloured signal activates one transmission direction applying the periodical repetition of the Composite Source Signal. After the last voiced part of this signal (see the red arrow), a second voiced sound (green) taken from the uncorrelated Composite Source Signal is fed in the opposite direction. This test signal is suited to determine the switching characteristics by appropriate level versus time analysis.

Figure 17: Combination of the Composite Source Signal in one transmission direction (red) and the uncorrelated voiced sound in the opposite direction (green) to switch between the two transmission directions

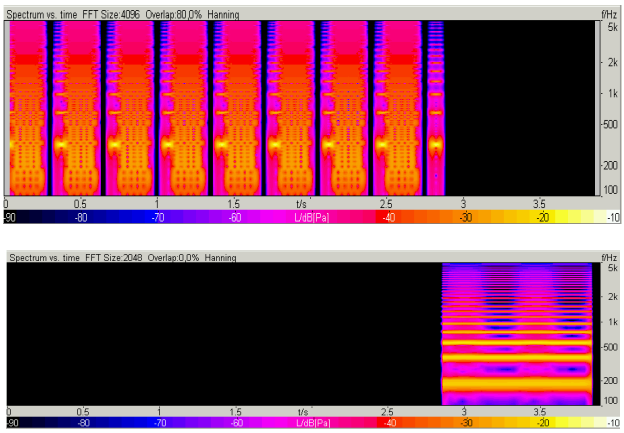


Figure 18: Spectral representation of the two Composite Source Signals (time scale on the x-axis, frequency scale between 100 Hz and 5 kHz on the y-axis)

If both signals (the one from figure 8 and the one from figure 15) are applied simultaneously to the measurement setup - this simulates a double talk period -, specific parameters determining transmission quality under double talk conditions can be analysed. These two Composite Source Signals can be combined in various ways to a two channel signal (including for example level variation or other signal characteristics) to simulate specific double talk situations during the tests. From the analysis of two Composite Source Signals those quality parameters that occur during periods of double talk can be determined (such as audible signal level variations or echo components during double talk).

Figures 19 and 21 give an example for a double talk signal. The signal construction is described in ITU-T Recommendation P.502 [4], clause 5.3.1. The Composite Source Signal on both channels is periodically repeated with a level variation of 20 dB in each transmission direction. The green test signal is the one shown in figure 8 and is fed into the measurement setup in the sending direction, the red signal is the one shown in figure 15 and simultaneously fed in the receiving direction. The yellow part shows the periods of "double talk", i.e. where both green and red signals are present. Note that the entire signal sequence has a duration of 32 s.

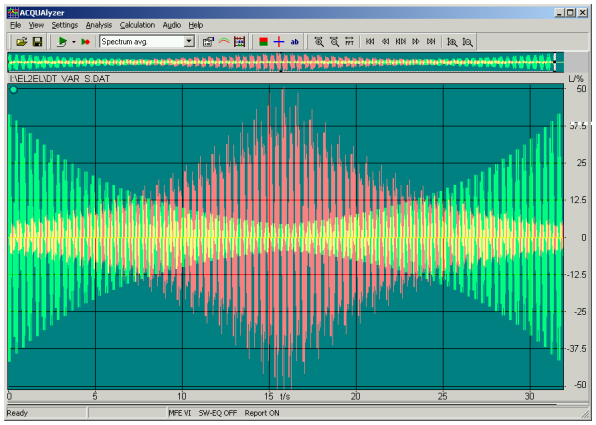


Figure 19: Test signal to simulate double talk based on the periodical repetition of the two Composite Source Signals (sequence length 32 s)

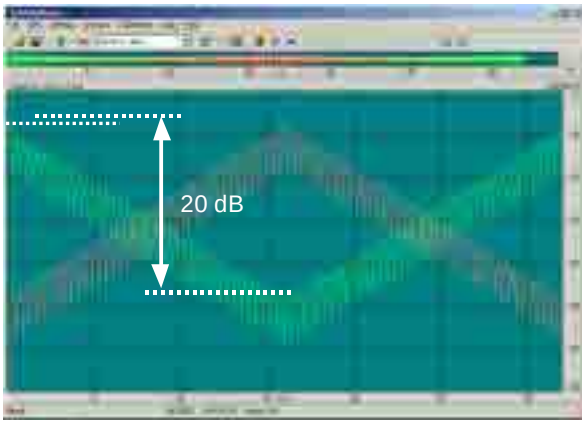


Figure 20: Level versus time analysis of the simulated double talk test signal

A sequence of 2 seconds from this double talk signal is shown in figure 21 in order to demonstrate the signal composition in detail. The two test signals reproduce short single talk periods in both transmission directions (only one signal is active, either green or red) and real double talk periods (both signals are active simultaneously as indicated in yellow). Figure 22 shows the level vs. time analysis of this enlarged part of the signal demonstrating that the signal energy partly overlaps.

Double talk sequences using real speech look just like this short periods of single talk in both directions and real double talk (overlap) as shown in figures 23 and 24. The green part is a male voice and the red part a female voice.

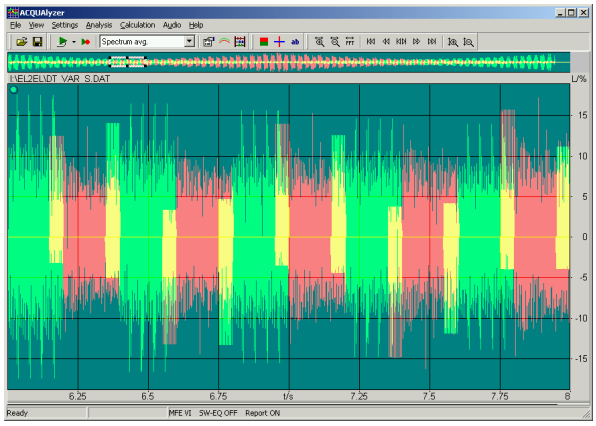
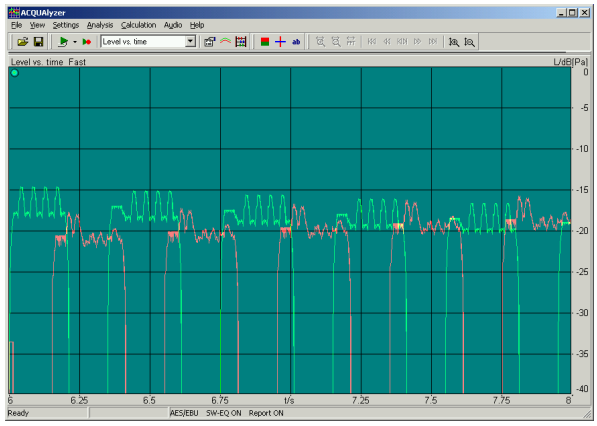


Figure 21: Enlarged time sequence from figure 19 showing in detail the periodical repetition of both CS signals



NOTE: Both signals partly overlap.

Figure 22: Level versus time analysis of both signals

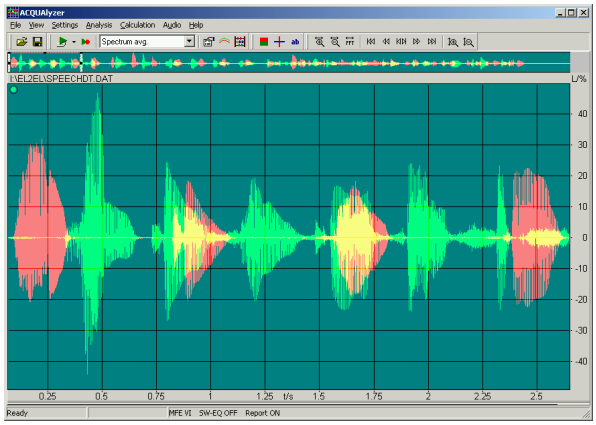


Figure 23: Enlarged time sequence from a double talk sequence (real speech sample, green: male voice, red: female voice)

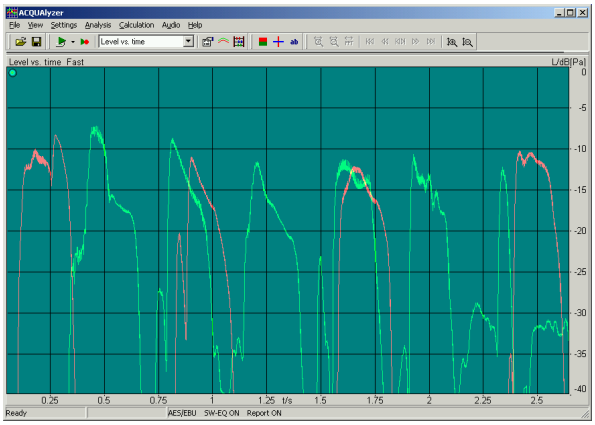


Figure 24: Typical double talk sequence using real speech (green: male voice, red: female voice)

Beside these speeches like test signals additional signals have been used to determine further parameter influencing the conversational quality. The background noise present at the location of both subscribers can not only be regarded as a disturbing factor in a real conversation. Moreover the transmitted background noise carries important information about the environmental conditions for the other subscriber and therefore plays an important role for conversational quality determination.

The transmission characteristic for a background noise signal can be determined using the test signal shown in figure 25. The sequence is a random noise signal with Hoth spectrum (see figure 26) according to ITU-T Recommendation P.800 [15] and is applied with increasing level to the measurement object. The level varies from infinite to -25 dB_V. If the transmitted signal is recorded and analysed the systems reaction on this background noise can be determined for different input signal levels.

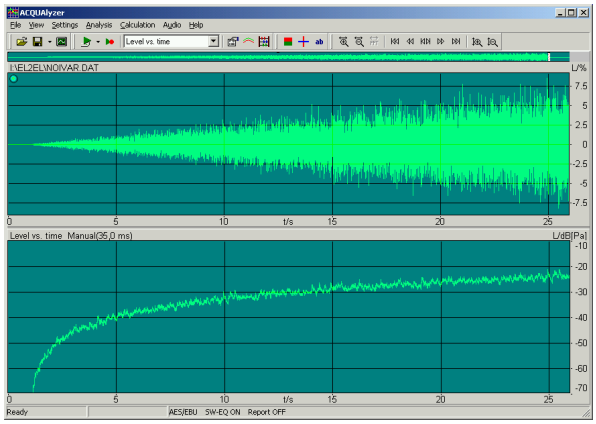


Figure 25: Noise signal with increasing level to determine quality of background noise transmission (lower window: level versus time analysis)

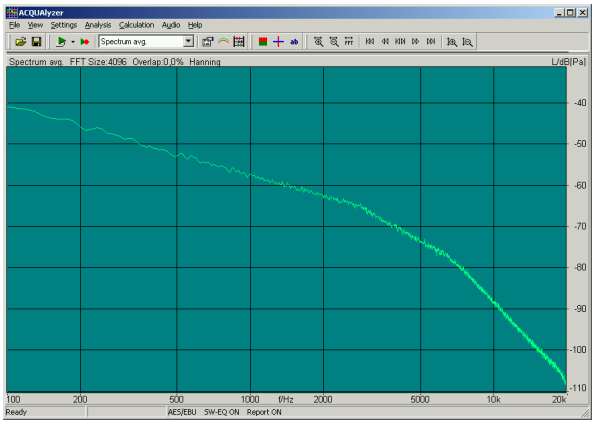


Figure 26: Hoth spectrum as defined in ITU-T Recommendation P.800 [15]

Moreover, several realistic background noise signals - recorded in a cafeteria and a student's pub - were used for the tests. These realistic signals have higher signal level variations vs. time due to voice babble in the background, laughing or other sounds from the cafeteria or the pub. Both signals are given in the following figures 27 and 28 together with level versus time analysis in the lower windows. Recordings were carried out using these test signals from realistic background noise scenarios to generate listening examples in order to compare it to the results obtained with the test signal is shown in figure 25.

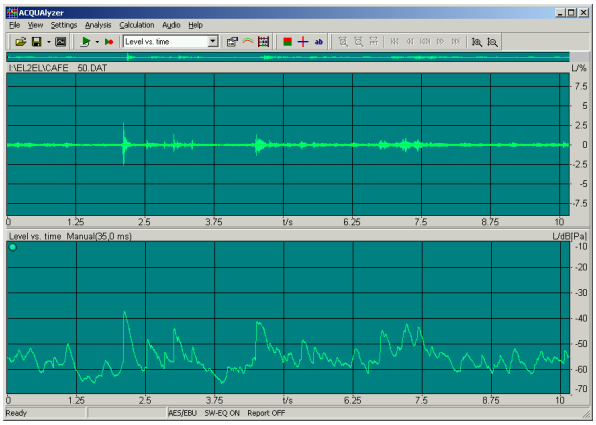


Figure 27: Cafeteria noise signal, average signal level -50 dB_{m0} (lower window: level versus time analysis)

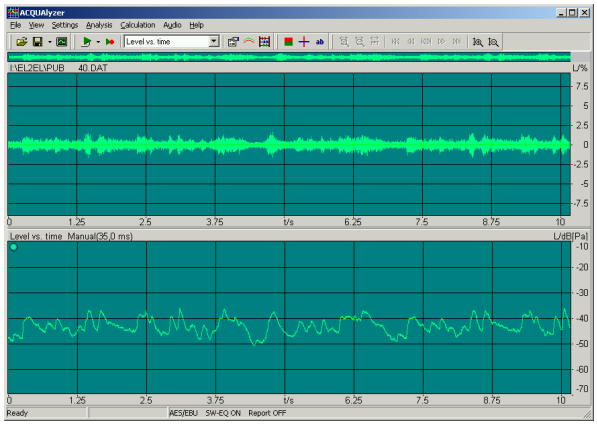


Figure 28: Pub noise signal, average signal level -40 dB_{m0} (lower window: level versus time analysis)

In addition these background noise signals were applied together with real speech and the Composite Source Signal in the opposite transmission path. The complete test signal is shown in figure 29. Typically the background noise signal (green) is applied at the near end, the speech signal and the Composite Source Signal at the far end. These sequences are used to determine level variations in the transmitted background noise signal, e.g. caused by the implemented non-linear processors working together with the echo cancellers to suppress the residual echo signals. The first part of the sequence including the speech signal can be used as a listening recording, the second part with the Composite Source Signal is used for detailed objective analysis.

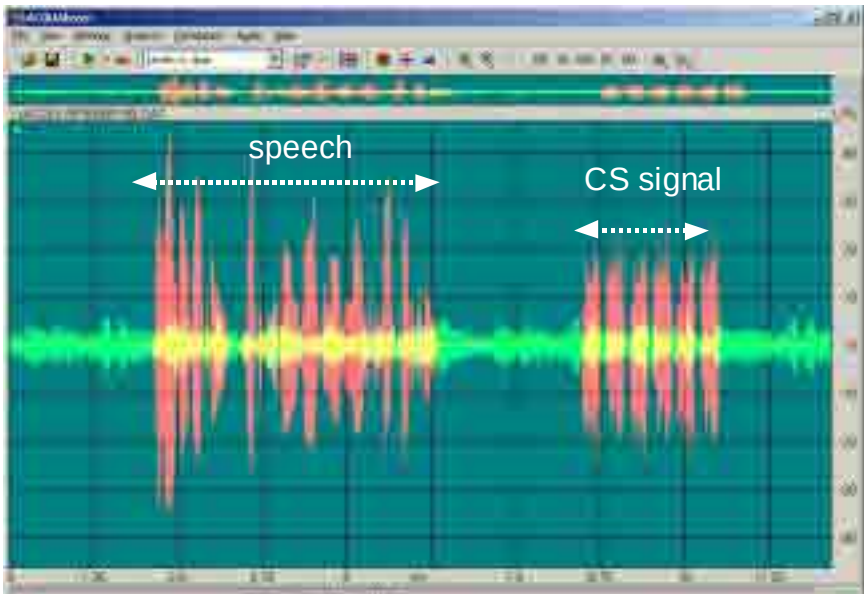


Figure 29: Realistic background noise (pub) with additional speech sequence and CSS applied in the same direction (green) or in the opposite transmission path (red)

The test signal for measuring echoes in the connections consists of the periodical repetition of the Composite Source Signal as shown in figure 30. The Composite Source Signal is repeated to achieve an appropriate signal length (12 seconds in figure 30).

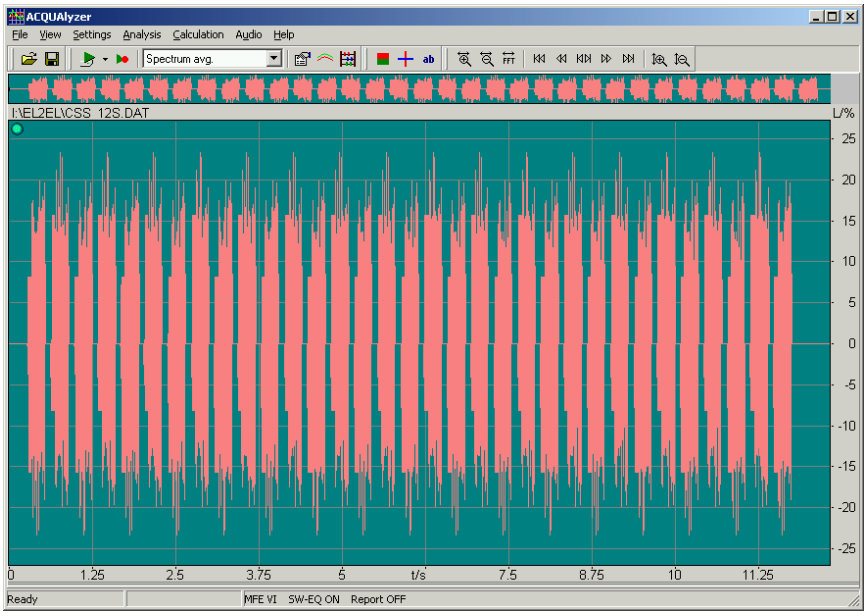


Figure 30: Periodical repetition of the Composite Source Signal to measure echo during single talk

An additional test signal to determine echo during double talk is represented in figure 31. The red signal is applied in one direction of the measurement setup and the green signal simulates the double talk (typically applied at the near end). The sequence shown in figure 31 in the upper window represents a single talk situation in receiving direction for about 2 s, then the double talk sequence (yellow colour) is applied for again 2 seconds and the sequence ends with another short single talk period. The power density spectrum calculated by Fourier transform is given in the lower window. The two signals show "comb-filter" spectra, which is necessary to distinguish between the double signal (coming from the near end) and the echo signal (coming from the echo path as a reaction on the receive signal) by appropriate filtering.

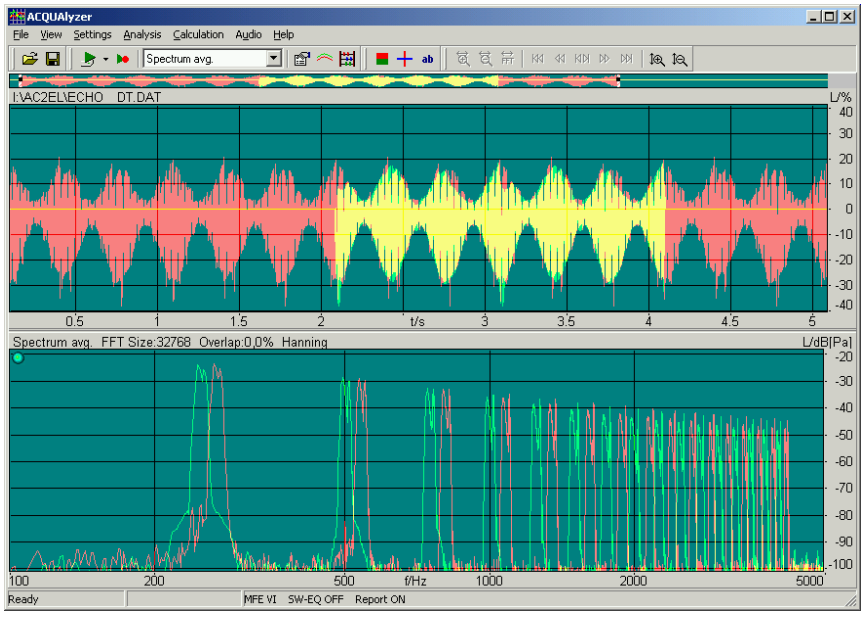


Figure 31: Two channel test signal with comb-filter structures to determine echo during double talk

7.5 Assessment methods

7.5.1 Instrumental assessment of listening quality

Latest psychoacoustic instrumental analyses using TOSQA [2] or PESQ (according to ITU-T Recommendation P.862) lead to an one dimensional "MOS-like" test result with a high correlation to quality scores gained by auditory listening only tests. These speech quality values describe listening quality and contain effects by one-way speech transmission which are perceived by a listener.

Both methods show a similar structure, because they are based on comparisons of the distorted signal with the undistorted reference input signal of the system. PESQ as well TOSQA2001 use psycho-acoustic models of the human speech perception.

Procedures for instrumental speech quality estimation usually work in several steps. The first step eliminates signal differences that are irrelevant in the modelled auditory test (e.g. total delay and level differences). The next stage transforms both signals to an "internal representation" using psycho-acoustic models for the human sound perception. The spread (including multidimensional aspects) between both pre-processed signals is computed and will be used for estimating a quality value.

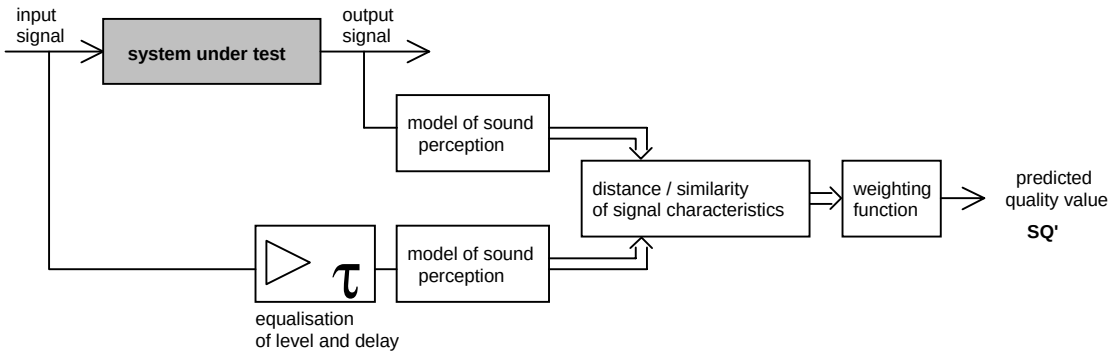


Figure 32: General structure of instrumental speech quality estimation approaches

PESQ as well TOSQA2001 follows this general structure but they differ mainly in:

- the kind of signal pre-processing;
- the used psycho-acoustical models;
- the determination of a value describing the speech quality; and
- the database for definition of the fitting function to transform the technical speech quality value to a predicted MOS value.

The methods are validated for VoIP transmission scenarios and therefore applicable for the scenarios to be tested during the event. TOSQA was used and validated by an auditory test in the 1st ETSI VoIP Quality Test Event. Here it demonstrated the same accuracy compared to ITU-T Recommendation P.862 [14].

For recordings at the acoustical interface, i.e. for electrical-acoustical as well as for acoustical-acoustical VoIP transmission scenarios, only TOSQA2001 was used because P.862 has not yet been verified for these conditions.

7.5.2 TOSQA and PESQ results and precision

For the ETSI TIPHON context two different VoIP speech quality tests can be used for demonstration of the accuracy of the PESQ and TOSQA results.

In 1999 the ETSI TIPHON SuperOP was carried in Hawaii. During this event speech recordings were made and evaluated subjectively. This database was part of the competition for the new ITU-T Recommendation P.862 in 2000. For this database results gained by PESQ as well as TOSQA shows correlations of 95 % and higher with the MOS scores from the auditory test.

During the 1st ETSI VoIP Quality Test Event only TOSQA was used for estimation of the one-way speech quality. A subset of the recorded data was also scored in a subjective evaluation by human listeners. For this subset the TOSQA results show a correlation of 92 %. The same speech material was also evaluated by PESQ for an ITU-T publication [14], also PESQ shows the same accuracy of about 92 % for this data.

In this 2nd ETSI VoIP Quality Test Event both methods were used in all scenarios with electrical measuring access at sending and receiving side. This on the one hand allows to compare the results of both test events directly on the basis of TOSQA's TMOS values. On the other hand the results of PESQ and TOSQA can be compared for measurements carried out in this 2nd ETSI VoIP Quality Test Event.

Although both methods show a similar accuracy in VoIP scenarios in comparison to MOS values, both methods are not identically. Therefore the methods could show different sensitivities for special impairments introduced by the VoIP system. The usage of both methods during the second VoIP ETSI speech quality test event allows a direct comparison of both methods under exactly the same conditions.

The transmitted 32 s speech files were divided in four sentence pairs 8 seconds each. These speech samples fulfil all requirements of ITU-T Recommendation P.800 [15] for conducting auditory tests. For each of these 8 s speech samples a TMOS as well as PESQ value were calculated. The resulting scores (16 for each measurement condition) were averaged and yields the instrumental score for the evaluated condition.

7.5.3 Instrumental computational assessment using speech-like (P.501) test signals

The auditory perceived quality for speech controlled, non-linear or time-variant systems like the VoIP equipment under test is influenced by various parameters like:

- one-way listening speech quality;
- echo disturbances;
- double talk performance;
- background noise transmission.

These parameters are subdivided in listening related disturbances (listening speech quality), talking-related impairments (e.g. echo) and conversational aspects (e.g. double talk performance, quality of background noise transmission). The overall quality of the complete system is determined by the combination of all these parameters.

Tests based on sophisticated test signals according to ITU-T Recommendation P.501 [3] and analysis methods ITU-T Recommendation P.502 [4] were developed to determine the corresponding instrumental parameters and provide tools for optimization. The following points indicate how the results of the measurements during the event can be interpreted and used.

The **one-way listening speech quality** is determined by analysis methods like PESQ and/or TOSQA expressed by estimated mean opinion scores. The results are influenced by the speech coder and decoder, frequency responses, loudness ratings, the Packet Loss Concealment (PLC) and Voice Activity Detection (VAD) implementation and the jitter buffer behaviour. But these methods PESQ and TOSQA can only express the current quality without giving detailed hints about the reasons for quality impairments or, in other words, how to improve the current implementation. Therefore specific tests based on the P.501 test signals are used to analyse the corresponding parameter for potential improvements. The following analyses are therefore carried out:

- determination of frequency responses and loudness ratings;
- detection of activation thresholds for VAD both for speech like test signals and background noise signals;
- evaluation of the quality of implemented PLC and jitter buffer design using cross-correlation analysis and the "Relative Approach" [16]. The cross-correlation analysis is suited to analyse and demonstrate the current implementation whereas the Relative Approach is a hearing model based method to determine audible disturbances introduced by PLC or the jitter buffer control in the time and frequency domain. Consequently the combination of both methods can be used to optimize the current implementation.

These analyses, combined with the results of the test methods PESQ and TOSQA may provide useful information for improvements.

The **talking-related impairments** (typically echo) and the **conversational aspects** (double talk performance, background noise transmission) require additional tests and analyses:

- The results of delay measurements are compared to recommended values, e.g. in current ITU-T Recommendations.
- The double talk performance is mainly determined by the implemented echo cancellers and/or the combined non-linear processors. The corresponding tests based on the test signals according to ITU-T Recommendation P.501 [3] are used to determine audible level variations, syllable clipping and echo during double talk. These tests were carried out with different realizations of return losses in the echo path of infinite attenuation, 40 dB and 6 dB.
- The quality of background noise transmission is determined by parameters like activation threshold or silence suppression and comfort noise injection (if implemented).

Moreover recordings were carried out using real speech samples under single and double talk condition and realistic background noise. These recordings are provided on CD.

8 Results

8.1 Estimation of one-way speech quality

This anonymous presentation contains test results measured on gateway configurations. The results were averaged separately for the different types of codecs (e.g. G.711, G.729) and for the different test conditions. The individual system settings (e.g. packet length, VAD on or off) may differ depending on the manufacturers' implementation. At least results of three participants were necessary for the average process in order to guarantee the anonymity of each manufacturer involved in the averaging process. Three or more IP gateway implementations were tested with G.711 and G.729 codec. Consequently only these results are analysed here. For the averaging procedure not only the results of the "standard" test sessions (session one and session two) were used but also - as far as possible - results of the so called free-style sessions were included.

The estimation of one-way speech quality is based on real speech samples using the methods PESQ according to ITU-T Recommendation P.862 and TOSQA. Note, that TOSQA has already been used during the 1st ETSI VoIP Speech Quality Test Event.

All results presented here were measured at the electrical interfaces. 16 speech samples (sentence pairs) of 8 s length each were transmitted and scored by PESQ and TOSQA under all test conditions. These 16 single scores are averaged and represented on a MOS-like scale by the PESQ or TMOS values respectively. These scores were derived separately for each manufacturer and - in a second step - combined to an average overall result. In addition the maximum and minimum PESQ and TMOS score (representing the quality score of two individual manufacturers) is given to represent the quality range.

8.1.1 G.711 Codec

The G.711 codec was evaluated in the first test session. This session was obligatory for all participants in the test event. The different G.711 codec implementations use a packet-length of 20 ms or -in a few cases- 10 ms. The voice activity detection (VAD) was switched on for some tests and switched off for others. All G.711 codec implementations under test use an implemented Packet Loss Concealment (PLC).

As described above, for each test condition the 16 speech samples were transmitted and scored by the instrumental methods PESQ and TOSQA. The resulting average scores PESQ and TMOS are shown in table 5. On the left hand side the test condition number taken from the Test Specification is given. These conditions were realized by different settings of the NIST-Net IP simulator. For comparison the results measured on a reference connection are also included in the table. This reference connection was realized by an ISDN loop through the PABX without any VoIP components (*Reference connection PABX*). These results therefore indicate the measured speech quality values determined by the measurement equipment, the ISDN access (containing A/D and D/A converters and G.711 codec) and the PABX itself. In addition the results derived from a pure G.711 coding is shown. For these references the speech samples were digitally processed (offline) using the ITU-T standard code (*G.711 reference, VAD off*).

Table 5: TOSQA and PESQ averaged scores for G.711 (Gateway)

G.711, 20/10 ms packet length, VAD on/off, PLC on							
	NIST-Net settings fixed delay / packet loss / delay-jitter	TOSQA (TMOS)			PESQ		
		Average	Min	Max	Average	Min	Max
1a	0 ms / 0 % / 0 ms	**3,92**	3,81	4,05	**4,05**	4,00	4,11
2a	0 ms / 1 % / 0 ms	**3,64**	3,23	3,85	**3,80**	3,61	3,89
3a	0 ms / 2 % / 0 ms	**3,60**	3,44	3,77	**3,70**	3,65	3,73
4a	0 ms / 3 % / 0 ms	**3,46**	2,95	3,69	**3,58**	3,43	3,66
5a	0 ms / 5 % / 0 ms	**3,46**	3,13	3,76	**3,40**	3,31	3,48
6a	50 ms / 1 % / 20 ms	**3,49**	3,06	3,85	**3,66**	3,42	3,85
	Reference connection PABX	**4,07**			**4,14**		
	G.711 reference, VAD off	**4,20**			**4,39**		

Under test condition 1a (no IP network impairments introduced by NIST-Net) the quality scores are relatively close to the maximum possible quality score measured for the PABX reference connection. The slight differences are mainly caused by some implementations with jitter buffers reacting sometimes very sensitive and slightly unstable. This behaviour produces noticeable artefacts even under the test condition 1a.

Under the influence of packet loss (test condition 2a to 5a), the quality scores decrease, but the results can still be regarded as relatively high. This clearly demonstrates the influence of the implemented PLCs. These results can be compared to the results of the first VoIP test event (note that only TOSQA has been used during the first test event). In order to provide an additional result for comparison, it should be noted that TOSQA rates an error free transmission through the well-known GSM Full Rate codec with 3,4 TMOS.

Condition 6a covers the same packet loss rate as condition 2a, but considers an additional delay jitter (standard deviation 20 ms) by the IP network simulation NIST-Net. Assuming an ideal jitter buffer management the one-way speech quality should be the same under these two test conditions 2a and 6a. As shown in figures 33 and 34 differences between the test results under both test conditions occur. This is clearly noticeable and demonstrates that the jitter buffer obviously does not always cover the delay variations in an appropriate way in the tested implementations.

8.1.2 G.729 Codec

The G.729 codec was chosen as a second common implementation by the most of the participants. Consequently sufficient measurements data are available to provide averaged and anonymous analyses.

The codec implementations used a 20 ms packet length and -in some very few cases- a 10 ms packet length. Packet loss concealment was included, the voice activity detection was switched on or off respectively depending on the individual setting chosen by the manufactures in the specific test sessions.

Table 6: TOSQA and PESQ scores for G.729 (Gateway)

G.729, G.729A, 20/10 ms packet length, VAD on/off, PLC on							
	NIST-Net settings *fixed delay / packet loss / delay-jitter*	TOSQA (TMOS)			PESQ		
		Average	Min	Max	Average	Min	Max
1a	0 ms / 0 % / 0 ms	**3,49**	3,44	3,56	**3,59**	3,51	3,67
2a	0 ms / 1 % / 0 ms	**3,33**	3,11	3,45	**3,44**	3,26	3,54
3a	0 ms / 2 % / 0 ms	**3,26**	2,92	3,44	**3,33**	3,06	3,51
4a	0 ms / 3 % / 0 ms	**3,13**	2,86	3,28	**3,24**	3,02	3,36
5a	0 ms / 5 % / 0 ms	**2,92**	2,51	3,11	**3,10**	2,73	3,37
6a	50 ms / 1 % / 20 ms	**3,04**	2,58	3,35	**3,21**	2,91	3,45
	Reference connection PABX	**4,07**			**4,14**		
	G.729 reference, VAD off	**3,62**			**3,78**		

The G.729 codec itself introduces its specific, audible distortions, leading to lower PESQ and TMOS scores compared to G.711. Under the test conditions introducing different rates of packet loss these distortions are additionally superset by artefacts caused by lost and replaced packets. Moreover the comparison of the test results for condition 2a and 6a again shows the influence of the jitter buffer management on listening speech quality.

Figure 33 shows the PESQ scores for the G.711 and G.729 codec evaluation. Figure 34 shows the TOSQA results for the same conditions. The results are taken from the upper tables. In addition to results given in table 5 and table 6 the diagrams show the measured minimum and maximum scores for each condition. At the right hand side of each diagram some reference values are given. These scores represent the reference PABX connection (as described above) as well as the ITU-T reference codecs in a pure off-line simulation without physical measurement access or VoIP components.

The average values of PESQ and TMOS show nearly the same rank order (see also figure 35). Only in case of the jitter condition for the G.711 implementations the average TMOS scores are slightly lower than the average PESQ scores. Summarizing these differences between the average TMOS and PESQ scores can be described as a nearly constant offset. PESQ scores in general slightly more optimistic due to different implemented fitting function translating the internal results into the resulting values on the MOS-like scale. This can also be clearly pointed out by the comparison of the test results for the G.729 codec evaluation and the for reference conditions given at the right hand side of the diagram.

As an interesting matter of fact the quality assessment derived by PESQ and TOSQA differ more for the G.711 implementations. PESQ scores the different implementations of packet loss concealment in a closer quality range than TOSQA does. TOSQA distinguishes more between the different PLC algorithms and reacts much more sensitive than PESQ. For the G.729 implementations the quality range is comparable for PESQ and TOSQA.

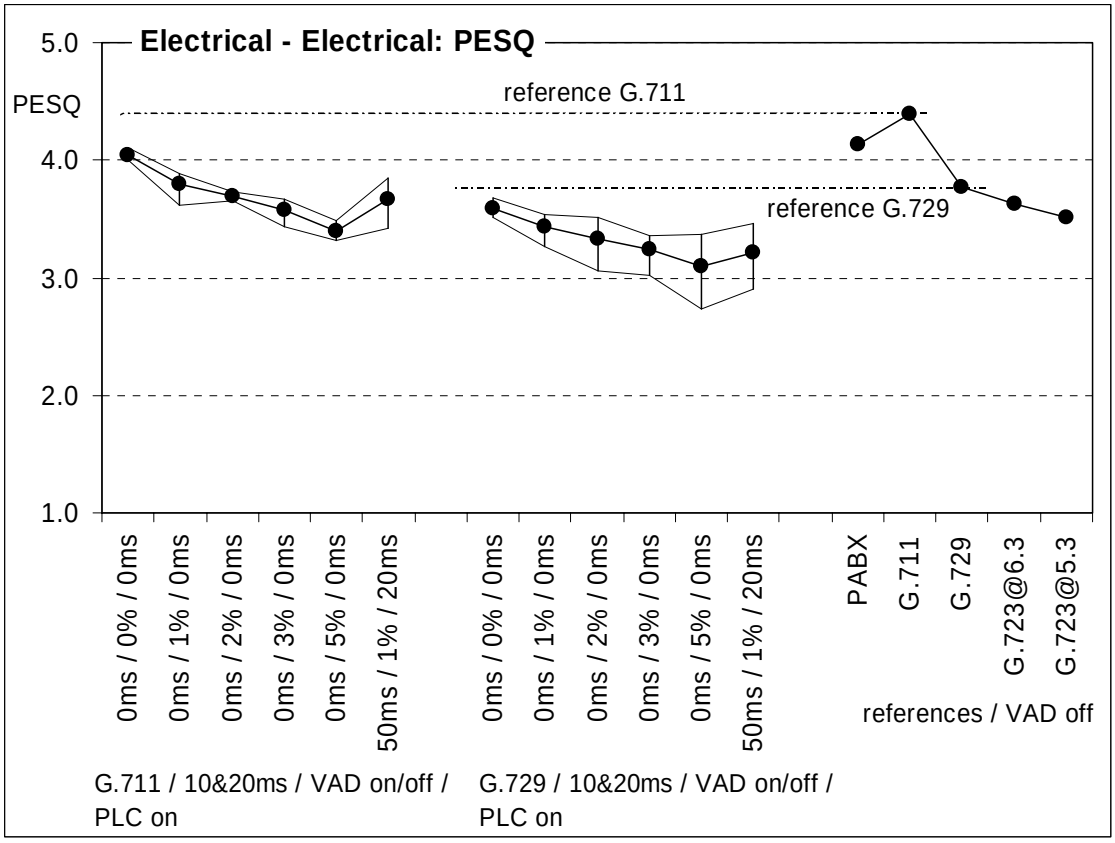


Figure 33: PESQ averaged scores for G.711 and G.729 (Gateway)

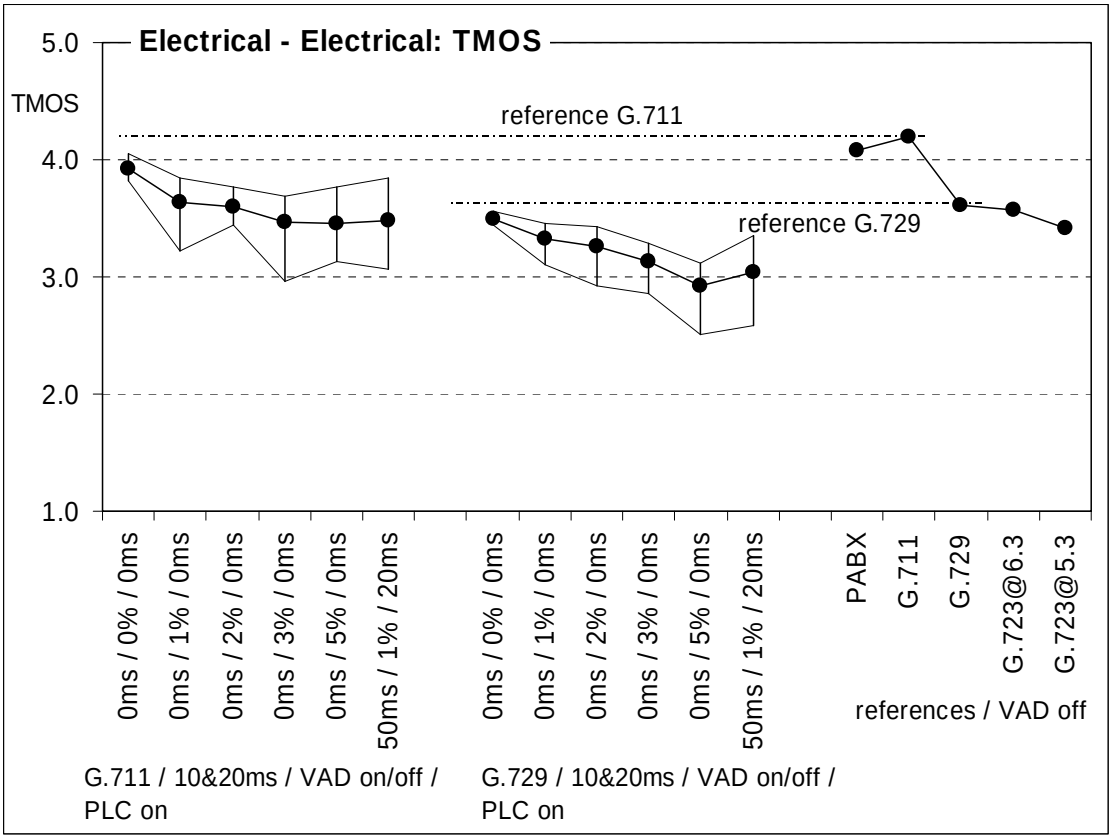


Figure 34: TOSQA averaged scores for G.711 and G.729 (Gateway)

Figure 35 shows first the averaged PESQ and TOSQA scores in direct comparison. This diagram again clearly demonstrates the small, constant offset between PESQ and TMOS and also the high similarity of the results.

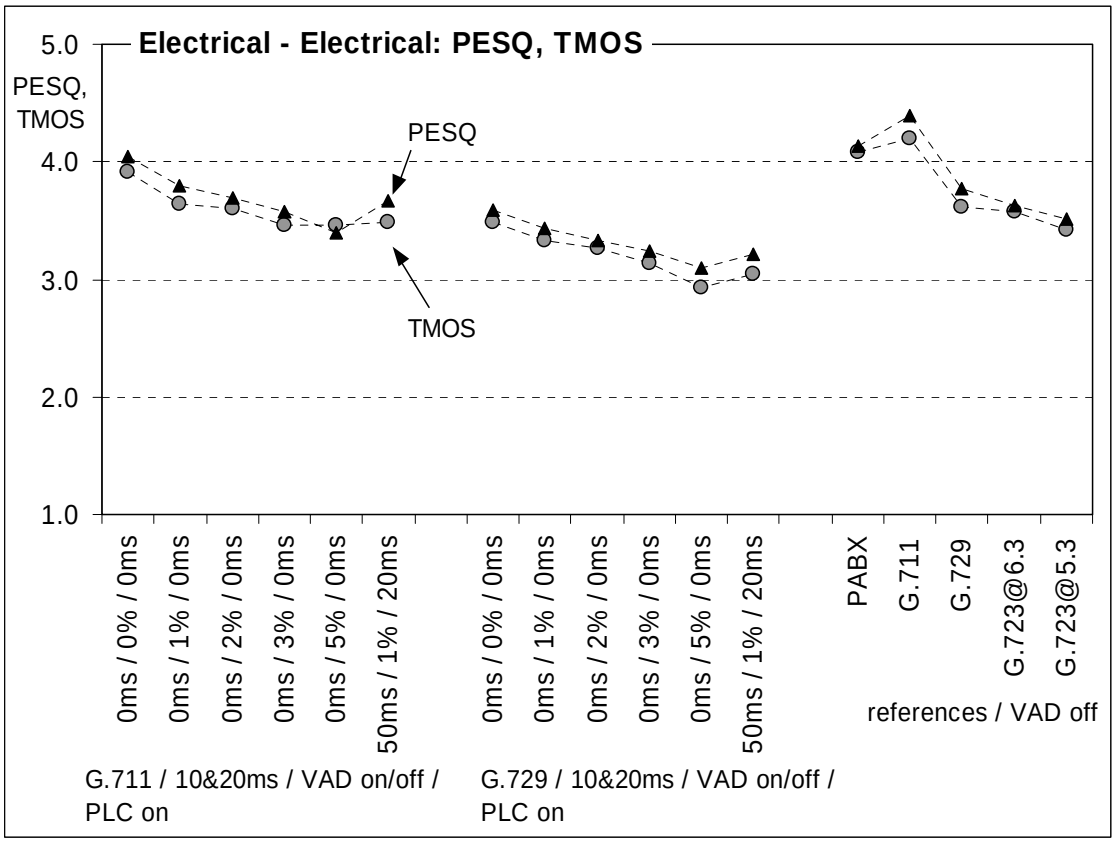


Figure 35: PESQ and TOSQA averaged scores for G.711 and G.729 (Gateway)

8.2 Delay measurements

The delay measurements for all codecs and under all test conditions lead to results between 62 ms and approximately 390 ms as the highest value.

The wide range can be explained by the different codec implementations but mainly by jitter buffer design and the jitter buffer control.

According to TS 101 329-2 these measured delays are between 2H (high quality) and 2A (acceptable quality). ITU-T Recommendation G.114 [22] recommends mean one-way delays less or equal 150 ms. Delays between 150 ms and 400 ms can be regarded as acceptable but conversation dynamics impairments can already be expected.

8.3 Transmission parameters, double talk performance and background noise transmission

The following analysis of transmission parameters cover all conversational aspects, the one way speech transmission, the echo performance, the double talk performance and the quality of background noise transmission. The description of some of the results are subdivided according to tests and test conditions providing enough single results for an anonymous representation:

- Tests with G.711 codec.
- Echo cancellers performance test with G.711 codec.
- Tests with G.729A codec.

Note, that the same tests as carried out for the VoIP gateways have been carried out with the local PBX. For better comparison some of the test results are also analysed for the local PBX without gateways connected. These test results demonstrate the test conditions during the test event and can be seen as references.

In the following, the order in which the figures are presented for each specific analysis is randomized.

Besides the gateway implementations measured also terminal implementations were measured. However the number of participants providing IP-terminals was not high enough to allow a summary presentation of the results without violating the confidentiality guarantee given to the manufacturers.

In general it can be stated that the effects and impairments explained below can be observed in a similar way for terminal implementations. One important difference however is the measured frequency response from mouth to ear where the typical leakage behaviour of handset in receiving can be observed. Depending on pressure force a high pass frequency characteristics which various due to the acoustical leakage between handset and ear is measured: for low pressure forces (2 N) applied to the handset a stronger high pass filtering can be observed compared to high pressure forces (13 N). For some implementations the measured frequency response at 13 N was close to ideal. It should be noted that for an excellent coupling all artefacts introduced by packet loss concealment algorithms are more audible compared to low pressure force coupling since the audible distortions are mostly present in the low frequency domain (see e.g. figures 48 to 51**)**. Although the speech sound quality for low pressure force coupling is worse than for high pressure force coupling, this effect may help to minimize the audible distortions of the PLC algorithms.

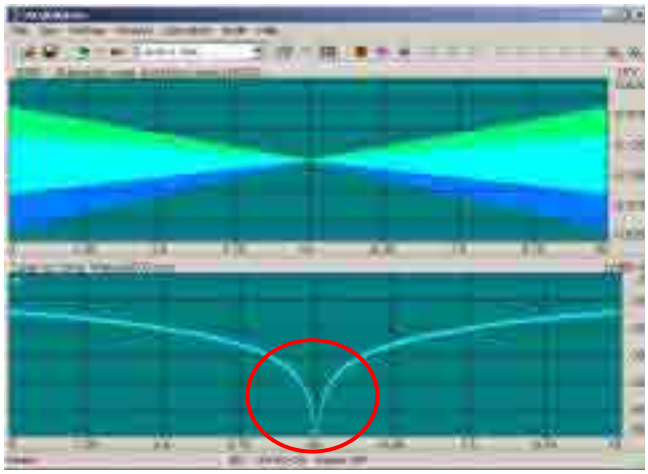
8.3.1 Tests with G.711 Codec

These tests were carried out with the following settings:

- Packet length 20 ms.
- PLC on.
- VAD on or off.

The behaviour of the implemented VAD and automatic gain control (AGC) - if implemented- can be analysed using the test signal consisting of the periodical repetition of a voiced sound with decreasing and increasing level vs. time (see clause 7.4.2). The measured result for the PBX as reference connection is shown in the figure 36. Similar results can be measured for some of the VoIP systems if VAD is switched off and comfort noise insertion is disabled. However it is clear, that with these settings no silence suppression and consequently no bandwidth reduction can be achieved.
Figure 36 shows the sequence for the complete test signal length of 10 seconds. The original test signal is shown in blue in the upper window (time sequence), the transmitted (measured) signal is displayed in green. The lower window represents the level vs. time analysis. Again the curve for the original test signal is displayed in blue and the one for the measured signal in green.

Figure 36 demonstrates that the level of the transmitted signal follows the original test signal level over the complete range independent of the actual signal level. The connection via the local PBX is linear and transparent, as it could be expected. The figures 37 to 41 show typical results for different gateway connections measured during the event. Note, that for better comparison, these figures only show the enlarged sequence displaying the analysis results for the lower signal levels (taken from the middle of the test sequence as indicated by the red circle in figure 36).



NOTE: The red circle shows the enlarged sequence for the analyses shown below.

Figure 36: PBX reference connection, analysed over the complete signal length (10 s)

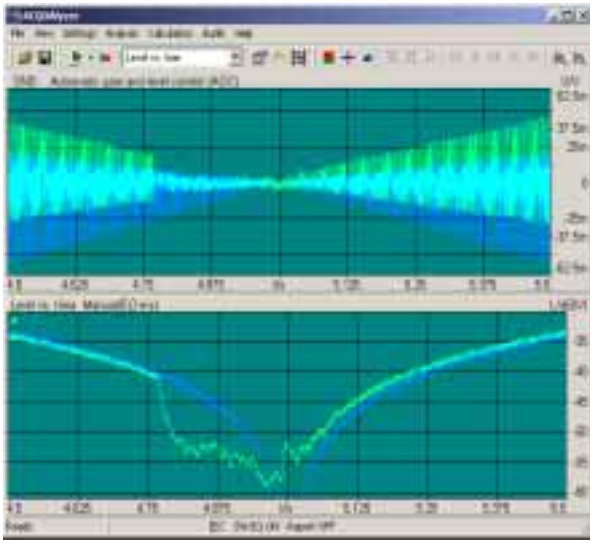


Figure 37: Implementation 1, VAD on

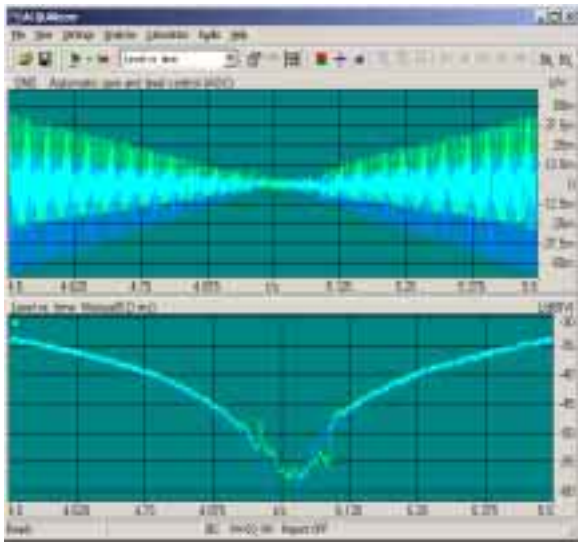


Figure 38: Implementation 2, VAD on

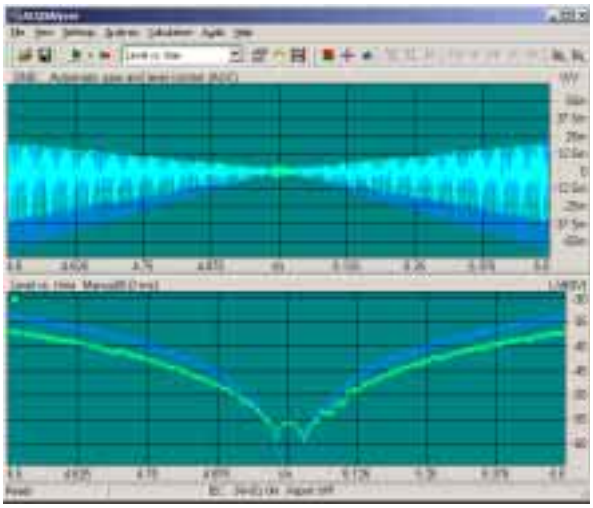


Figure 39: Implementation 3, VAD off

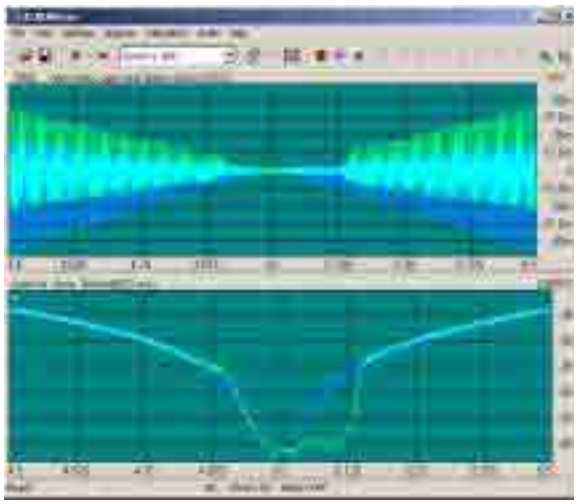


Figure 40: Implementation 4, VAD on

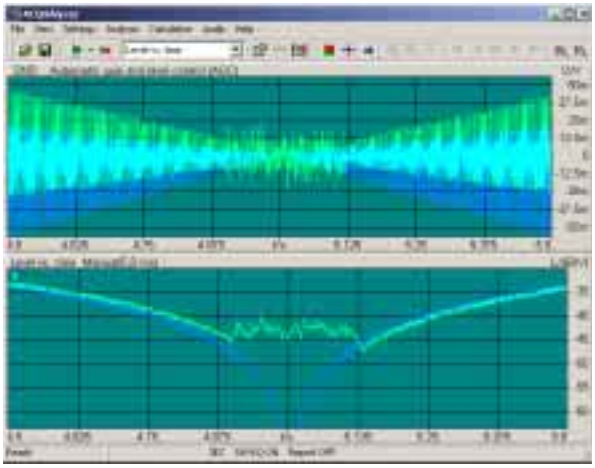


Figure 41: Implementation 5, VAD on

The two implementations 1 and 4 are in principle comparable. The original test signal is not transmitted if the signal level decreases below a certain threshold. Note that this threshold is slightly different for these two implementations 1 and 4 (see figures 37 and 40). If the test signal level increases again, the signal is immediately transmitted for the implementation 1 (see figure 37). Figure 39 demonstrates that the activation threshold seems to be identical for the decreasing and the increasing signal level.

The example of the implementation 2 is shown in figure 38. Obviously the original test signal is replaced by another signal. A deactivation and activation threshold can be determined, but the measured signal between these thresholds seems to be adapted to the original -not transmitted- signal.

The two other implementations 3 and 5 as they are analysed in figures 39 and 41 show, in principal, a comparable transmission behaviour. Obviously, the original test signal is replaced by a noise signal generated at the receiving side. The difference compared to the other implementations as described above, is the constant signal level of the generated noise. This signal seems not to be adapted to the original test signal. But the level of the generated signal seems to be low enough for the implementation 3 (see figure 39) not to cause audible disturbances. The signal level analysed for the implementation 5 (see figure 41) is too high. It should be noted, that the level does not correspond neither to the actual test signal level, nor to the idle channel noise measured during signal causes (e.g. before the application of this test signal). The level of this generated noise seems to correspond to the activation threshold.

One way speech quality assessment using the methods PESQ according to ITU-T Recommendation P.862 [19] and TOSQA cover the influence of speech coding and decoding, the quality of the implemented PLC and VAD and the jitter buffer behaviour. The results are shown in clause 8.1 - Estimation of One-way Speech Quality. The following analyses concentrate on the **PLC and jitter buffer implementation** in order to analyse current implementations with its specific, audible disturbances.

The test signal used is again the periodical repetition of the voiced sound as introduced in clause 7.4.2, Artificial Test Signals. Only the first 5 seconds of the test signal with decreasing level vs. time are used. Two kind of analysis are applied:

- cross correlation analysis between the transmitted signal and the original test signal to show the technical implementation of PLC and jitter buffer design; and
- Relative Approach analysis, a hearing model based psychoacoustic analysis method to analyse audible disturbances in the time and frequency domain.

The basis for the analysis method Relative Approach [16] is a hearing model according to [17]. In contrary to other analysis methods the Relative Approach does not use any reference signal. The nonlinear relationship between sound pressure level and loudness perceived subjectively is taken into account by time/frequency warping in a Bark filter bank and proper integration of the individual outputs. In the implemented realization of the Relative Approach a forward estimation based on the signal history is calculated in order to predict the new - expected - signal value. Values between critical bands are interpolated. This predicted value is compared to the actual signal value and the deviation in time and frequency is displayed as an "estimation-error". Thus instantaneous variations in time and dominant spectral structures are found based on the human ear sensitivity on these parameters. The Relative Approach, applied on the transmitted signal consisting of the periodical repetition of the voiced sound is suited to evaluate signal discontinuities typically introduced by packet loss, packet loss concealment or jitter buffer adaptation.

Figure 42 shows the analysis result for the PBX as reference. In the right hand figure 43 one gateway implementation is analysed for the G.711 codec. This results from a measurement under "clean" network conditions (test condition 1b), i.e. without impairments introduced by the IP network simulation.

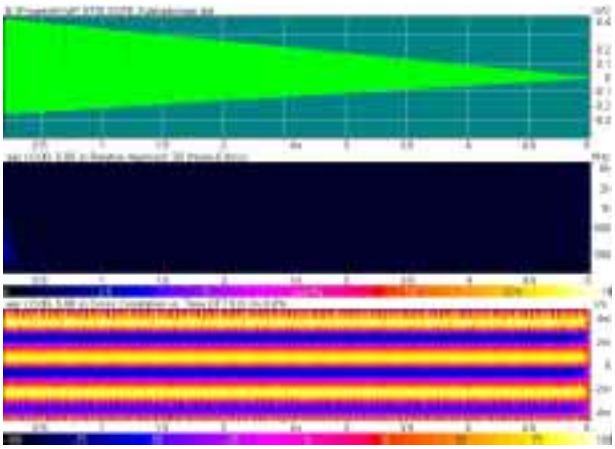


Figure 42: Reference connection via PBX

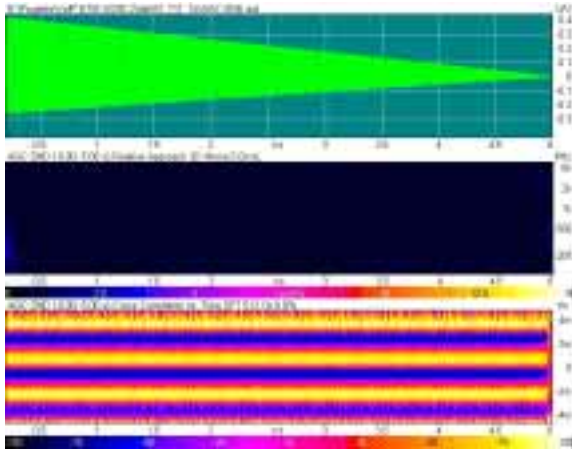


Figure 43: Implementation 1, G.711, clean network condition

The upper windows in both figures show the measured time sequence in green. The sequence length is 5 seconds. The middle window shows the Relative Approach analysis result vs. time (x-axis) and frequency (y-axis between 100 Hz and 5 kHz). As described above this analysis calculates a forward estimation and compares the actual signal. The estimation error is then colour coded and displayed. The warmth of the colour correlates to the estimation error, small errors are displayed in dark colours. As demonstrated in both figures no audible disturbances in the time or frequency domain are detected by this algorithm. This could be expected for both, the PBX connection (see figure 42) and the gateway connection due to the clean network conditions and the speech coding introduces by the G.711 codec (see figure 43).

The lower window shows the cross correlation analysis vs. time (x-axis). The phase shift between the transmitted signal and the original test signal is shown on the y-axis between -4,5 ms and +4,5 ms. It should be noted that the periodical structure of the test signal leads to a periodical pattern in the cross correlation analysis. The peaks show a harmonic structure of approximately 3 ms which corresponds to the duration of the voiced sound. It can be seen that the signal transmission is constant, no disturbances or signal discontinuities can be detected with the cross correlation analysis.

Different kinds of network impairments were introduced in test conditions 2b, 3b and 4b according to the test specification:

- 5 % packet loss (2b);
- 20 ms jitter (3b); and
- 5 % packet loss and 20 ms jitter (4b).

The **reaction on packet loss** introduced by the IP network simulation was covered by test condition 2b. The one-way listening speech quality is mainly determined by the quality of the implemented packet loss concealment algorithm. Differences occur for the implementations under test as expressed by the results analysed in clause 8.1, Estimation of One-way Speech Quality. These differences are clearly shown by PESQ and TOSQA for the test condition of high packet loss rates (5 %). For lower packet loss rates, TOSQA reacts more sensitive than PESQ.

The detailed tests carried out during the test event based on the test signals according to ITU-T Recommendation P.501 [3] and P.502 [4] together with the analysis methods described above were specially designed to analyse the current implementations and the corresponding parameters in the gateways. The application of these analysis methods to evaluate the influence of packet loss (without jitter) is compared for four different implementations in figures 44 to 47. Again the time signals are shown in the upper window of each figure, the results derived from the Relative Approach analysis in the middle window and the cross correlation in the lower window.

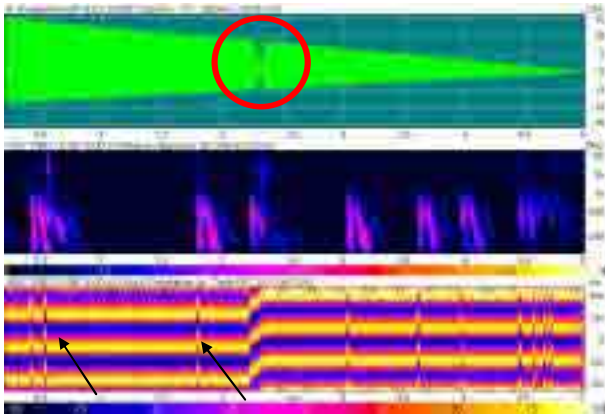


Figure 44: Implementation 1, packet loss

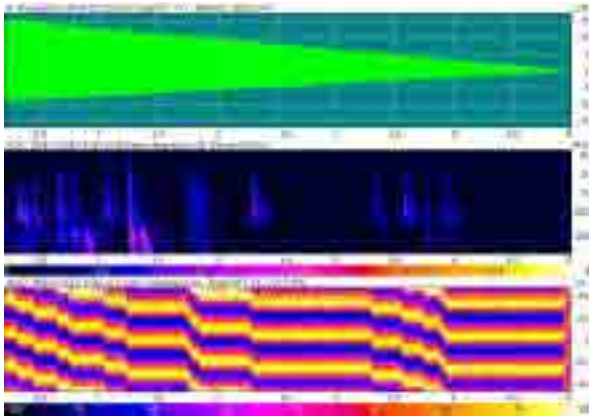


Figure 45: Implementation 2, packet loss

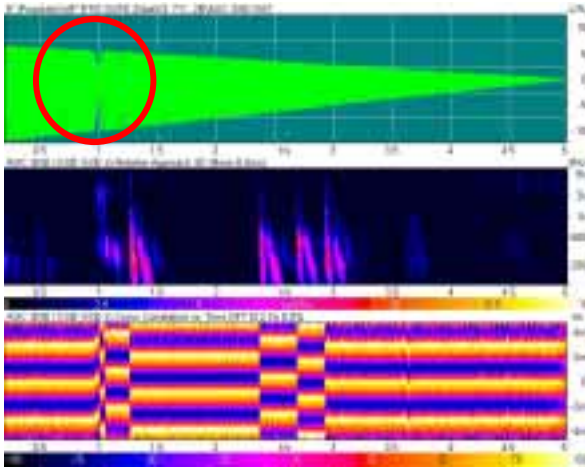


Figure 46: Implementation 3, packet loss

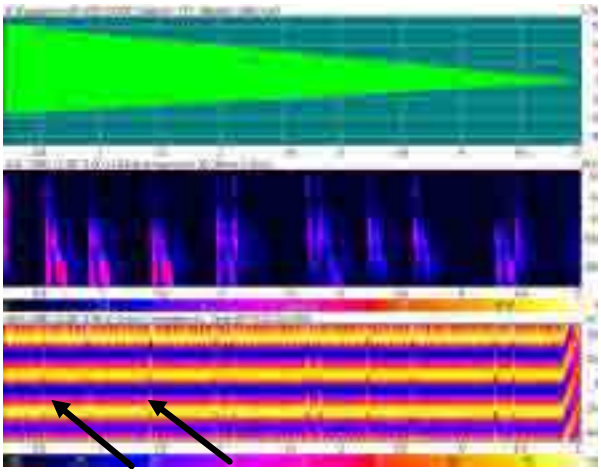


Figure 47: Implementation 4, packet loss

NOTE: The y-axis for the Relative Approach analysis in the middle window is scaled between 100 Hz and 5 kHz. The white, dotted line indicates the 300 Hz frequency.

In the example given in figure 44 packet loss and its concealment can be detected by the cross correlation analysis in the lower window. The occurrence of packet loss and its "substitution" leads to "visible" short signal discontinuities in this analysis. The information of the signal phase is detected and displayed by the cross correlation analysis. The short signal discontinuities in the cross correlation analysis indicate that the lost packet are substituted without changing the characteristic of the signal before and after this "disturbance" (see the black arrows in figure 44). A comparable result is demonstrated in figure 47.

The -probably most important- information about the audible disturbances introduced by these PLC algorithm implementations can be derived from the Relative Approach analysis as indicated in the middle window. For the two examples in figures 44 and 47 it can be noticed that the substitution of the - lost - packets leads to audible disturbances mainly in the lower frequency range compared to the occurrence of packet loss in the middle or and the end of the test signal. This may be partly caused by the test signal level which is higher at the beginning and decreasing vs. time. Moreover, the disturbances occur mainly in the lower frequency range below 1 kHz. This indicates that the signal interpolation between the transmitted test signal and the substituted packet could be optimized.

The example in figure 44 shows a short signal gap at 2,2 seconds on the time axis (see the circle in the upper time sequence window). This leads to a significant disturbance not only in the lower frequency range as analysed by the Relative Approach. The cross correlation analysis in the lower window demonstrates a continuous phase shift versus time for about 1 ms. One explanation for this could be the re-adaptation of a signal buffer in re-synchronizing the phase of the original, transmitted signal. A similar effect was also analysed for the example given in figure 46 (see the red circle).

The result shown in figure 45 shows a "smooth" implementation of PLC. As indicated in the middle window in figure 45 the Relative Approach detects unexpected disturbances for the human ear but these disturbances in the audible frequency range, typically used for telephone band limited signals (see the dotted line indicating the 300 Hz frequency) are minimized. The occurrence of packet loss is detected as signal discontinuities in the cross correlation function given in the lower window. The phase changes coincident with the occurrence of packet loss.

A comparable implementation adapting the phase of the substituted and transmitted signal is analysed in figure 46. The cross correlation analysis in the lower window detects the lost packets which are substituted by the implemented PLC algorithm. The signal phase obviously changes when a loss packet occurs and has to be substituted by the PLC. It is constant then until the next lost packet has to be substituted. The audible, residual disturbances as analysed by the Relative Approach in the middle window are mainly detected below 800 Hz instead of 1 kHz as analysed for other implementation.

The combination of these analysis methods (cross correlation, Relative Approach) with its specific results together with the listening examples provided on CD for each manufacturer may help to improve the current implementations. The quality of the implemented packet loss concealment algorithms can be evaluated using the Relative Approach as a hearing model based and therefore aurally adequate analysis. Combined with the cross correlation analysis between the measured (disturbed) signal and the original test signal details of the implementation can be evaluated.

The occurrence of **5 % packet loss and 20 ms jitter** (test condition 4b) is analysed in figures 48 to 51.

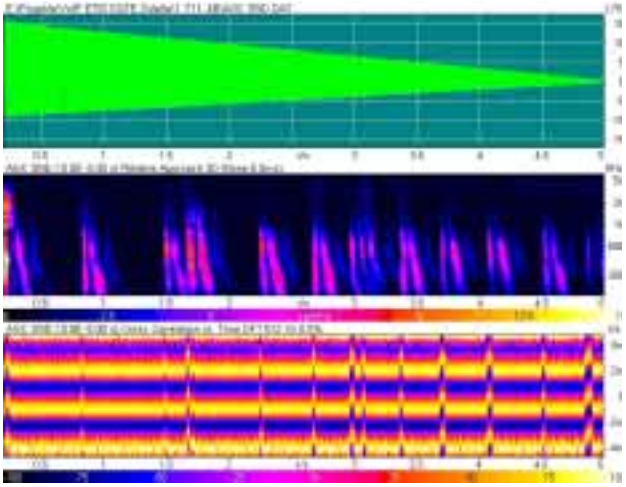


Figure 48: Implementation 1, packet loss and jitter

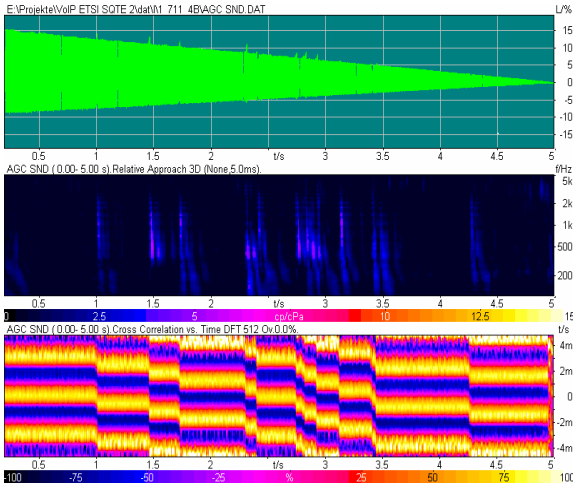


Figure 49: Implementation 2, packet loss and jitter

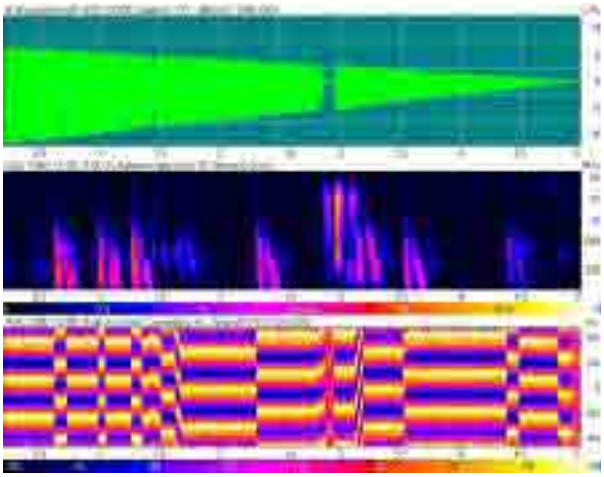


Figure 50: Implementation 3, packet loss and jitter

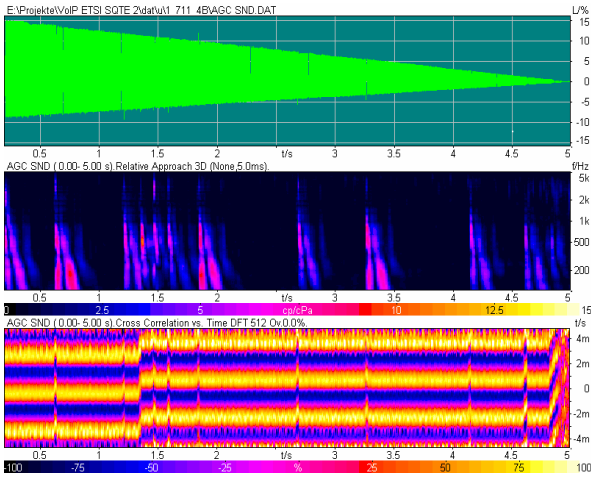


Figure 51: Implementation 4, packet loss and jitter

The analysis in figure 48 for the implementation 1 shows the substitution of lost packets in both, the Relative Approach and the cross correlation analysis. The residual disturbances for the human ear are again concentrated in the lower frequency range.

The example analysed in figure 51 is comparable except for one discontinuity detected by the cross correlation analysis (around 1,4 seconds on the time axis). Obviously the signal phase changes. This behaviour could only be detected once during this sequence of 5 seconds. Therefore perhaps is can be assumed that the jitter buffer is changed, e.g. in terms of one packet length (or multiples). Another explanation would be the slight adaptation of an additional signal buffer in order to re-synchronize the phase of the original signal.

The two implementations 2 and 3 analysed in figures 49 and 50 show in principle a different behaviour. The cross correlation analysis in both lower windows detects frequent phase shifts. The Relative Approach indicates an optimized performance for the implementation 2 in figure 49. The residual, audible disturbances are significantly lower than for the three other implementations.

The **quality of background noise transmission** is influenced by parameters like activation thresholds or the generation of Comfort Noise. Tests have been carried out using noise signals and realistic background noise scenarios as described in clause 7.4.2, Artificial Test Signals.

The application of the noise test signal with increasing level vs. time (Hoth spectrum) leads to the result shown in figures 53 to 57. Figure 52 shows the reference measurement for the PBX.

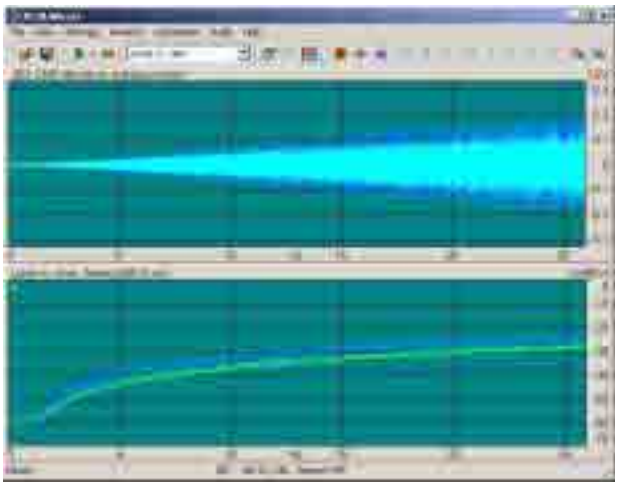


Figure 52: PBX Reference connection

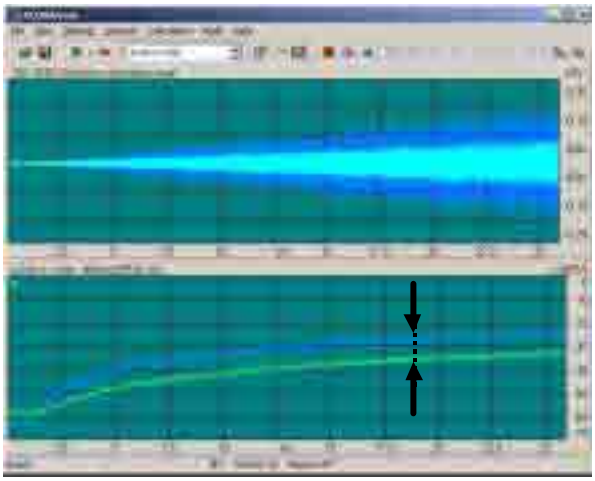


Figure 53: Example 1, VAD disabled

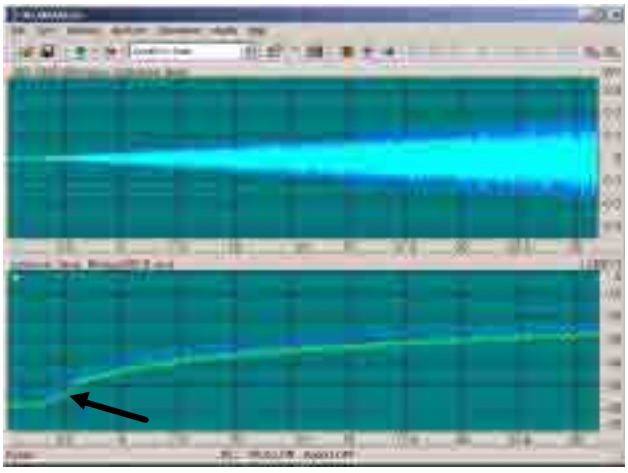


Figure 54: Example 2, VAD enabled

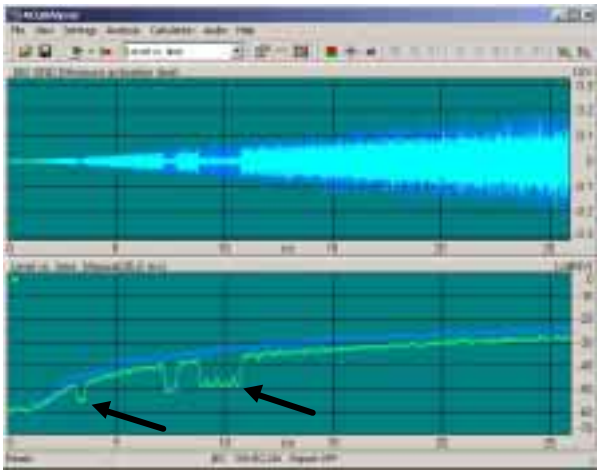


Figure 55: Example 3, VAD enabled

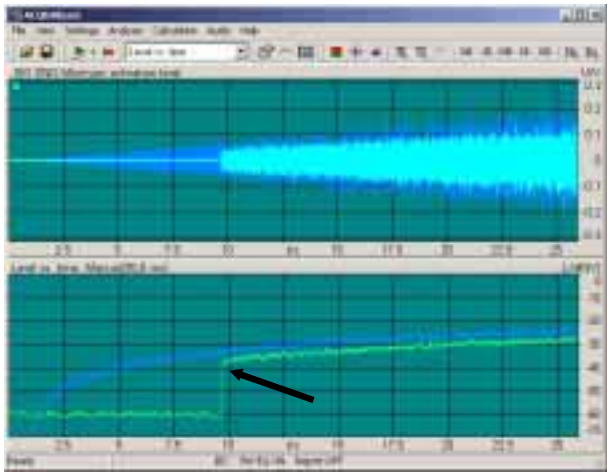


Figure 56: Example 4, VAD enabled

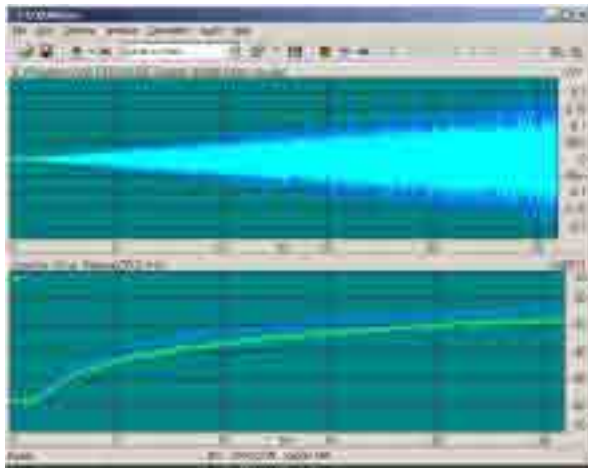


Figure 57: Example 5, VAD disabled

NOTE: The y-axis scaling is slightly different in some of these figures. This only effects the representation but not the results or the comparison between the implementations.

The level of the noise test signal increases vs. time over a complete length of approximately of 25 seconds up to -25 dBV. The upper window of each figure demonstrates the time sequence. The original test signal is displayed in blue, the measured signal after transmission over the two gateways in green. The lower window shows the analysis of level vs. time. Again the curve for the original test signal is represented in blue, the curve for the measured signal in green. The differences between these two curves, the one for the original test signal and the one for the measured signal characterize the transmission behaviour of the connection.

The corresponding result for the PBX stand alone is shown in figure 52. As expected no activation threshold is implemented. The analysis curve for the transmitted signal follows the original test signal level. The noise floor can be determined to approximately -60 dBV.

The example 1 analysed in figure 53 was measured for a gateway connection with VAD disabled. The measured curve in the level versus time analysis window follows the corresponding curve for the original test signal. But the signal is attenuated. The distance between the two curves is significantly higher compared to the result for the PBX reference connection. This attenuation is constant and independent of the kind of test signal (noise signal, speech, speech-like test signals).

The two examples analysed in figures 54 and 55 demonstrate different implementations of noise generation at the receiving side. Note, that VAD was enabled in these examples. The noise signal level is properly adapted to the original test signal level for the example 2 (see figure 54), before the original signal itself is transmitted. An adaptive noise generation can also be measured for example 3 (see figure 55), but the noise level is approximately 10 dB lower compared to the original test signal level. Moreover, the listening example demonstrates that the generated noise does not match the background noise test signal characteristic. The control mechanism in this implementation seems to switch frequently between the transmission of the original background noise signal and the generated noise.

A very high activation threshold for this background noise test signal was determined for the example analysed in figure 56. This implementation does not generate Comfort Noise at the receiving side, which may lead to quality degradation in telephone conversations under realistic noisy conditions.

Another example with a disabled VAD is given in figure 57. The level of the transmitted signal is comparable to the original test signal level, indicating that the signal is completely transmitted.

In general it can be assumed that a high activation threshold of an implemented VAD may lead to clipping in the presence of **realistic background noise signals**. Typically realistic background noise scenarios can not be described as continuous signals. Quite frequently a significant level variation vs. time including signal peaks occurs. An implemented activation threshold may lead to the transmission of the high level signal peaks in the background noise and the suppression of the lower level signal parts. Consequently the signal transmission is not continuously and interrupted by audible gaps. Moreover an injected Comfort Noise should be adapted to the original background characteristics.

Figure 58 shows a recording carried out with a realistic background noise signal originally recorded in a student's pub. The original signal used as the measurement signal is displayed in blue in the upper window, the transmitted and recorded signal in green. The time sequence in the upper window already shows interruptions, the background noise signal is not completed transmitted, the signal parts with lower levels are suppressed. The lower window shows the spectral analysis vs. time, using a frequency range between 100 Hz and 5 kHz (y-axis). The intensity of the transmitted signal is colour coded, gaps are indicated by dark colours.

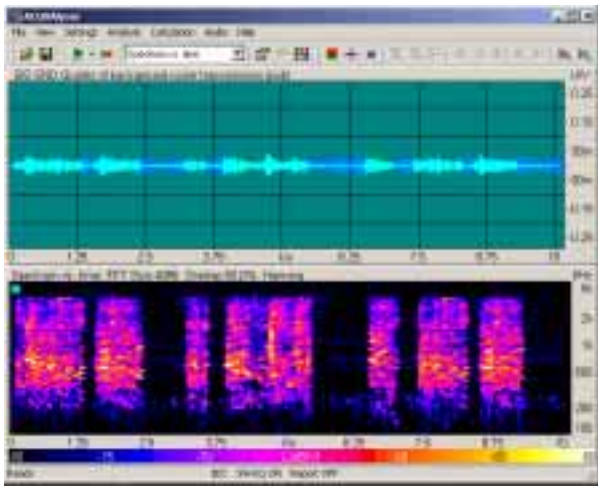


Figure 58: Interrupted background noise, student's pub

Figure 59: Comfort Noise generation

The Comfort Noise feature was enabled in this example. It can be seen clearly that the gaps are now filled up by the generated noise at the receiving gateway. The power density spectrum matches the original signal spectrum in the higher frequency range, but not in the lower frequency range. Although the spectral analysis in figure 59 displays also the lower frequency range below 300 Hz (see the dotted line), which is typically not transmitted over a standard, bandlimited handset used for telephoning, care should be taken to take the spectral characteristics **and** the actual level of the original background noise signal into account.

8.3.2 Echo canceller performance test with G.711 Codec

The **tests of the implemented echo cancellers** under single and double conditions were carried out in different setups: a complete 4-wire scenario with infinite ERL and with a simulated echo path attenuation of 40 dB ERL and 6 dB ERL. The echo path simulation was provided by the measurement front-end MFE VI controlled by the test system ACQUA. The setup together with a description of some definitions (near end, far end) is shown in figure 60. As a reference connection the same test were carried out with the PBX stand alone. This setup is shown in figure 61.

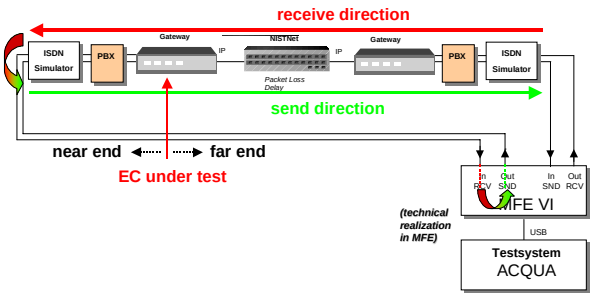


Figure 60: Test setup and definitions for the echo canceller used for all EC implementation tests

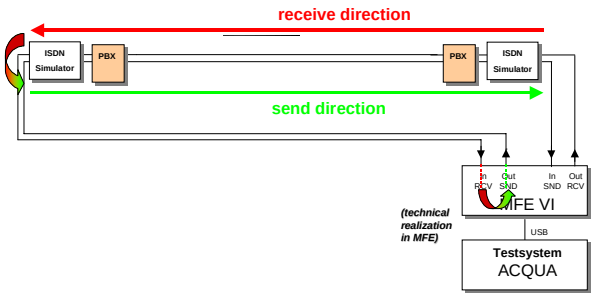


Figure 61: Test setup for the PBX as reference

Subjective tests indicate that, specially the transmission performance under double conditions, determine the quality of echo cancellers. Tests under double talk conditions were carried out with the periodical repetition of 2 Composite Source Signals applied in receiving direction and in sending direction at the near end side (for the definitions see figure 60**)**. The receive signal level for the echo canceller varies between -30 dB_{mo} (corresponding to -32,2 dBV) in the beginning, increasing up to -10 dB_{mo} (corresponding to -12,2 dBV) and decreasing to -30 dB_{mo} at the end of the test signal. The signal itself it is described in clause 7.4.2, Artificial Test Signals. The near end signal (double talk signal for the echo canceller under test) is applied with -10 dB_{mo} in the beginning decreasing to -30 dB_{mo} and increasing again to -10 dB_{mo}. This double talk signal should be completely transmitted and should not be disturbed by echo components. In the following figures the measurement result for the gateway connections and for the connection via the PBX are compared. The left hand figure 62 shows the result for PBX. The test setup was a complete **4-wire connection**. This implies that - physically - no echo exists, the connection is echo-free.

The test signal applied at the near end is shown in blue colour, the transmitted signal measured in sending direction (see the definitions in figure 60) is represented in green.

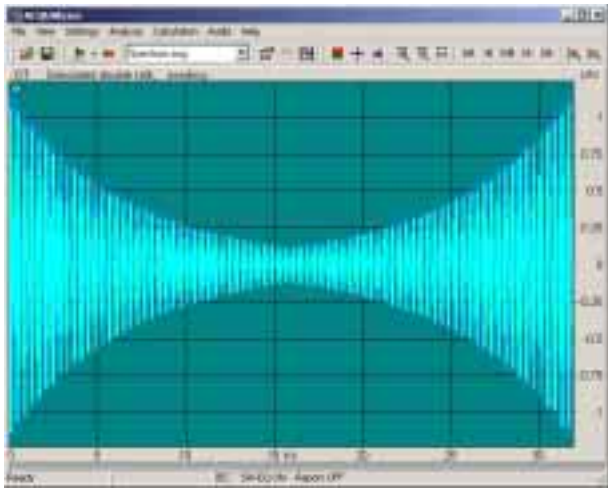


Figure 62: PBX Reference connection

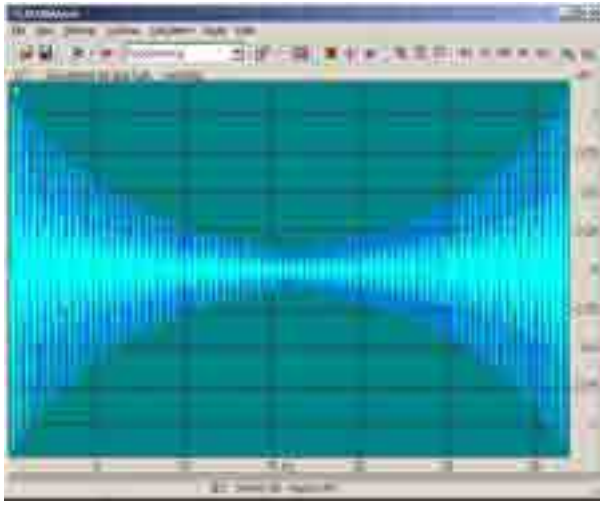


Figure 63: Example 1

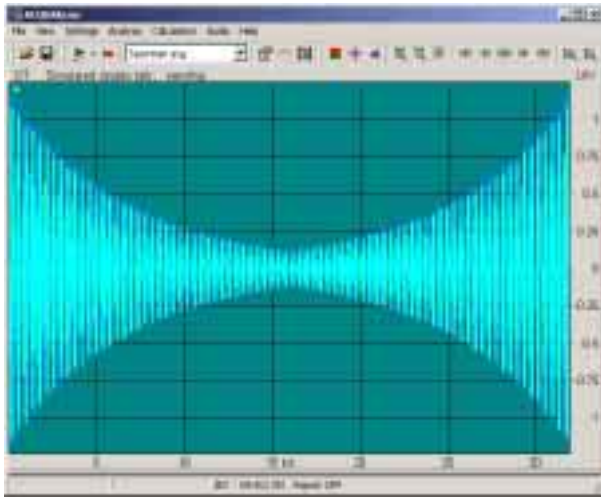


Figure 64: Example 2

Figure 65: Example 3

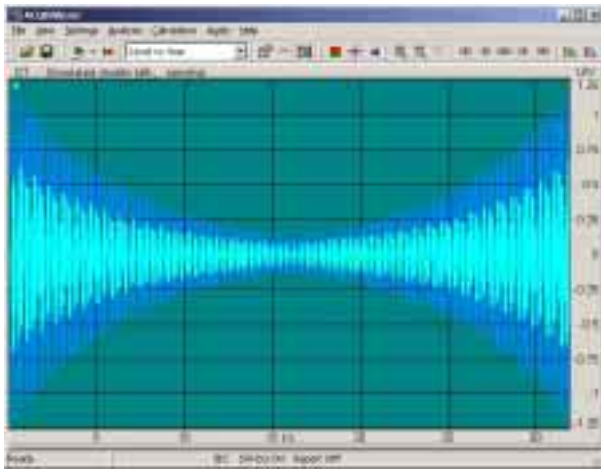


Figure 66: Example 4

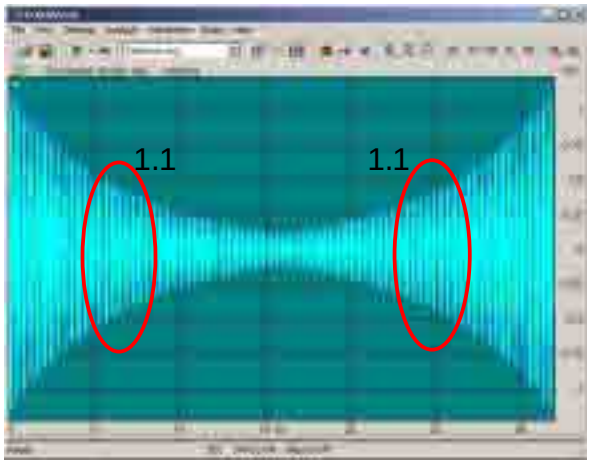
The analyses result for the PBX reference connection demonstrates that all signal bursts applied at the near end, are completely transmitted. There is not level difference between the original signal and the transmitted signal. This could be expected for the reference connection.

A comparable result can be seen for the example 2 analysed in figure 64, whereas the two implementations from figures 63 and 66 behave different: The original test signal applied at the near end (blue colour) is significantly attenuated. Some very short signal components (peaks) are transmitted with the original signal level for the example analysed in figure 63, but this level variation reacts very fast, so that nearly the complete signal bursts are attenuated.

The example shown in figure 65 again shows a different implementation. At the beginning and at the end of the analysed double talk sequence, the signal bursts are completely transmitted or attenuated by approximately 6 dB. In the middle sequence the signal bursts applied with the lower levels are clipped. Note that the test signal level in receiving direction is very high during this middle part, whereas the near end signal level is low. It can be expected, that clipping occurs also during the application of real speech for this implementation. Moreover, the two examples, analysed in figures 63 and 66, inserting a significant attenuation to the near end signal will also lead to the same behaviour during the application of real speech. The near end speech is attenuated.

It should be remembered, that these tests were carried out in a complete 4-wire-scenario. Echo disturbances can physically not occur. The implemented echo cancellers are probably responsible for the level variations or signal clipping as demonstrated above. But they could be completely disabled in this scenario or at least, should not influence neither the receive nor the send direction.

The same tests have been carried out with an echo path realization of 40 dB ERL and 6 dB ERL. The 40 dB ERL echo path realization leads to similar results compared to the infinite ERL test condition, because the 40 dB ERL typically exceed the possible echo attenuation provided by the implemented adaptive filters in the echo cancellers. Therefore, some detailed analyses are given below measured with the **6 dB ERL echo path realization**. The following analysis shows two enlarged sequences taken from the measured signal in order to demonstrate analysis details.



NOTE: This figure shows the definition of two sequences A and B from the double talk sequence. The signal level of the original test signal in receiving and sending direction is comparable during these periods (see the description of the test signal in clause 7.4.2, Artificial Test Signals).

Figure 67: Enlarged sequences A, B (PBX)

Figure 68: Example 1,6 dB ERL (A)

Figure 69: Example 1,6 dB ERL (B)

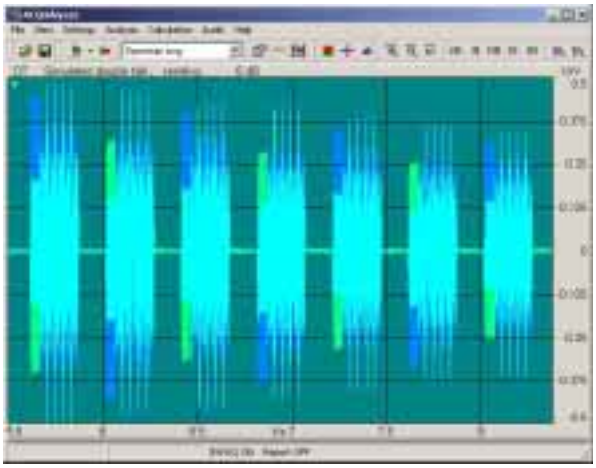


Figure 70: Example 2,6 dB ERL (A)

Figure 71: Example 2,6 dB ERL (B)

Figure 72: Example 3,6 dB ERL (A)

Figure 73: Example 3,6 dB ERL (B)

The example analysed in figures 68 and 69 demonstrates a high echo attenuation during the double talk sequence. Note, that the echo can be determined during the pauses between two CS-signal bursts under the double talk condition (see the red arrows). But the signal bursts applied at the near end are attenuated. It should be noted, that any attenuation inserted in sending direction also lowers the residual echo level.

A different result can be seen in figures 70 and 71. The signal bursts are not attenuated, but the echo attenuation during the pauses of two CS-signal bursts is still very high.

Again a different behaviour is shown in figures 72 and 73. For this implementation a level variation in the transmitted near end signal appears during sequence A (compare the two bursts indicated by the black arrows). Moreover, the echo canceller seems to diverge during the double talk sequence. The echo level determined between two CS-signal bursts during sequence A is significantly lower compared to sequence B.

8.3.3 Tests with G.729A Codec

For the G.729A - codec the Relative Approach and the cross correlation analysis were applied. Again these measurements were carried out under the different network conditions. The implementations of the G.729 - codecs are first analysed under the clean network condition. For comparison the reference connection via the PBX (without a G.729 codec) is analysed in figure 74. The other results shown in figures 75, 76, 77 and 78 are derived from tests with the G.729 codec for 4 manufacturers.

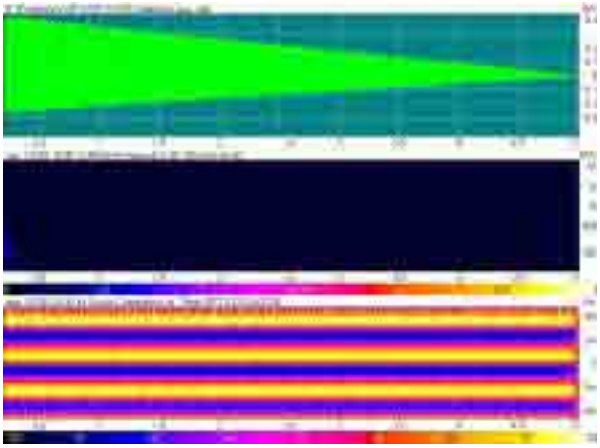


Figure 74: Reference connection via PBX

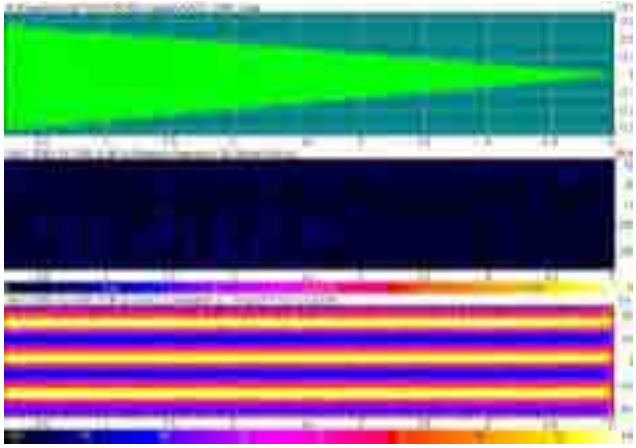


Figure 75: Implementation 1, G.729

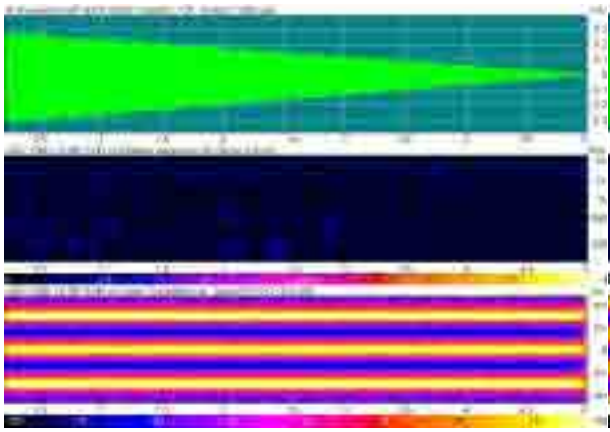


Figure 76: Implementation 2, G.729

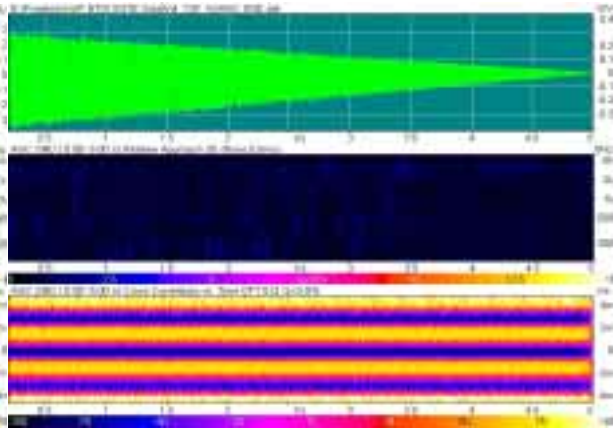


Figure 77: Implementation 3, G.729

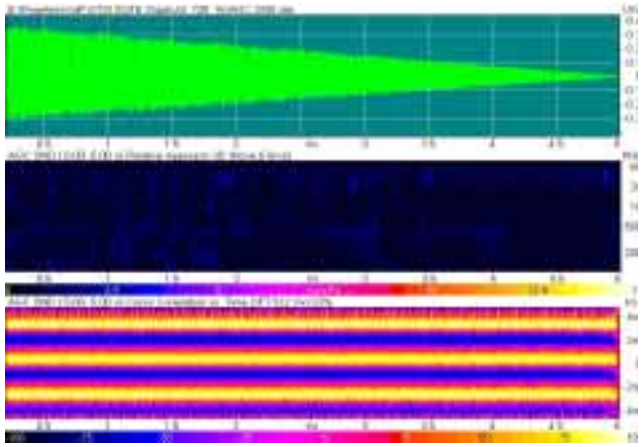


Figure 78: Implementation 4, G.729

It can be noted that the cross correlation analyses leads to comparable results for all implemented G.729 - codecs and the PBX reference connection. The Relative Approach analysis detects slight differences between the four codec implementations and the reference connection, but no significant differences between the implementations itself.

The **reaction on packet loss** introduced by the IP network simulation was covered by test condition 2b. The cross correlation analyses in the figures 79, 80 and 81 demonstrate the re-synchronization on the original test signal. The application of this analysis on the transmitted test signal consisting of the periodical repetition of a voiced sound, leads to harmonic structures as analysed above for the clean network condition. The results given in the examples in figures 79, 80 and 81 therefore demonstrate the performance of the implemented packet loss concealment.

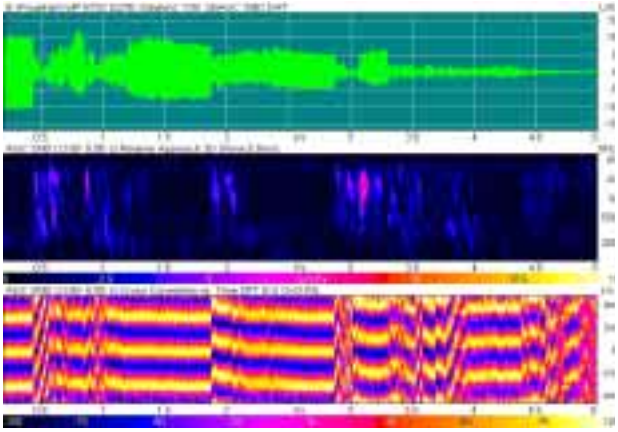


Figure 79: Implementation 1, packet loss

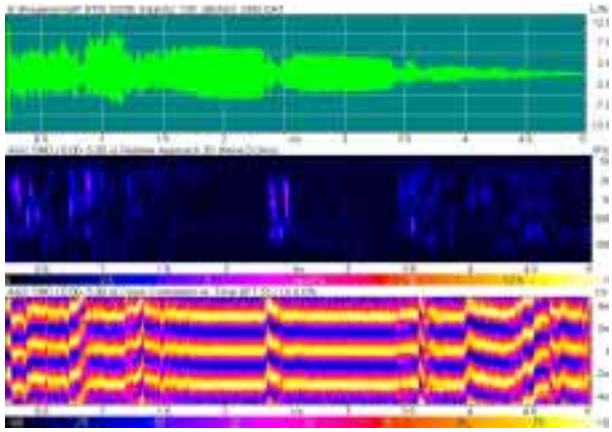


Figure 80: Implementation 2, packet loss

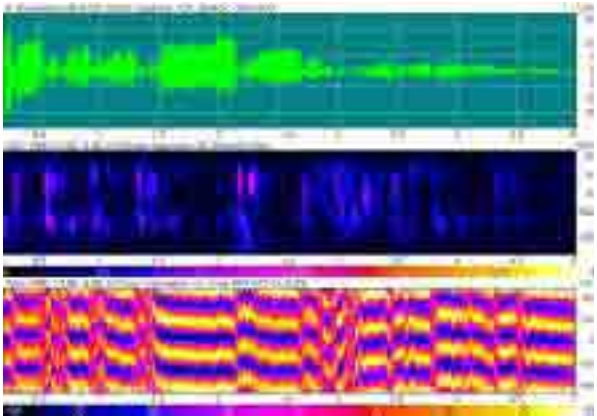


Figure 81: Implementation 3, packet loss

The Relative Approach analysis demonstrates audible disturbances introduced by the implemented PLC with significant contributions over the complete frequency range. Obviously the different types of packet loss concealment, implemented in the G.711 codec and the G.729 codec lead to a different kind of residual disturbances, auditory perception and annoyance. The comparison of the three implementations shows slight advantages for implementation 2 as shown in figure 80.

The occurrence of **5 % packet loss and 20 ms jitter** (test condition 4b) is analysed in figures 82 to 85.

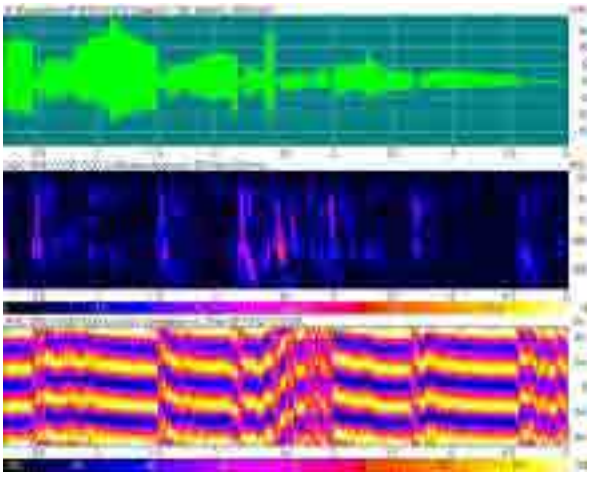


Figure 82: Implementation 1, packet loss and jitter

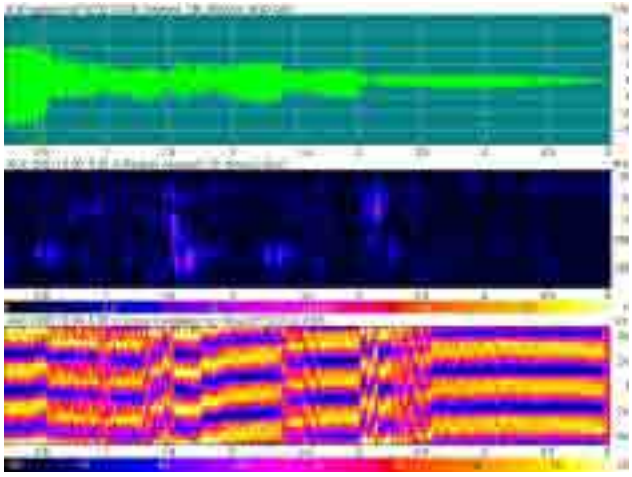


Figure 83: Implementation 2, packet loss and jitter

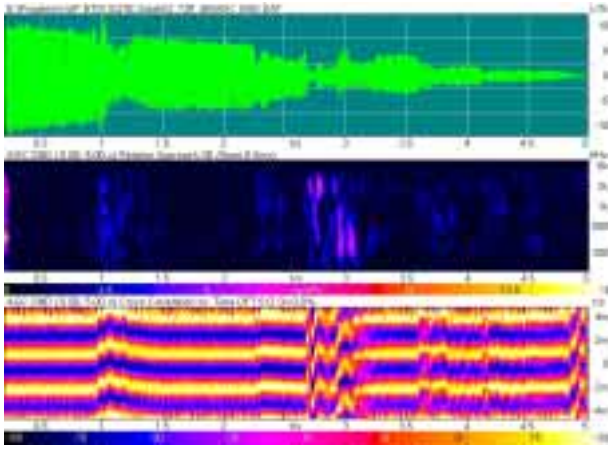


Figure 84: Implementation 3, packet loss and jitter

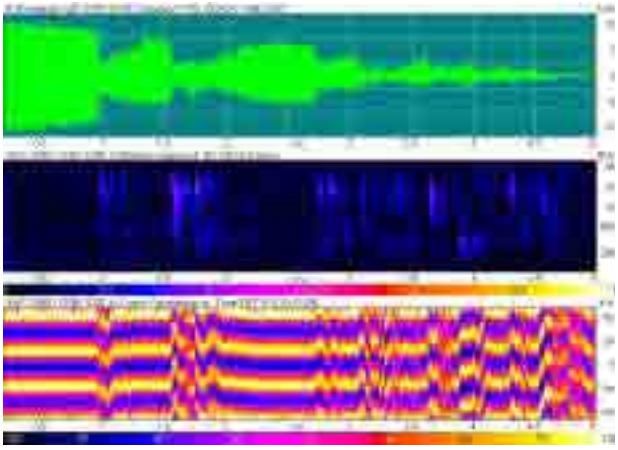


Figure 85: Implementation 4, packet loss and jitter

The comparison between the four implementations demonstrates slight advantages for implementations 2 (see figure 83) and 4 (see figure 85). The Relative Approach indicates significant disturbances distributed over the complete frequency range as analysed already above for the occurrence of packet loss in isolation.

History

Document history		
V1.1.1	February 2007	Publication