HEAD acoustics

Factors impacting the speech quality in VoIP scenarios – and how to assess them

Dr.-Ing. H.W. Gierlich
HEAD acoustics GmbH
Ebertstraße 30a D-52134 Herzogenrath, Germany
Tel: +49 2407/577 0 • Fax: +49 2407/577 99

http://www.head-acoustics.de

Overview

- ***Speech Quality in VoIP***
- ***Parameters influencing speech quality – measurement procedures***
  - Single talk - listening
  - Single talk – talking
  - Double talk
  - Background noise
- ***Summary***

Speech Quality Parameters

… from the user's perspective

talking situation

sidetone
echo
delay

speech quality

listening situation

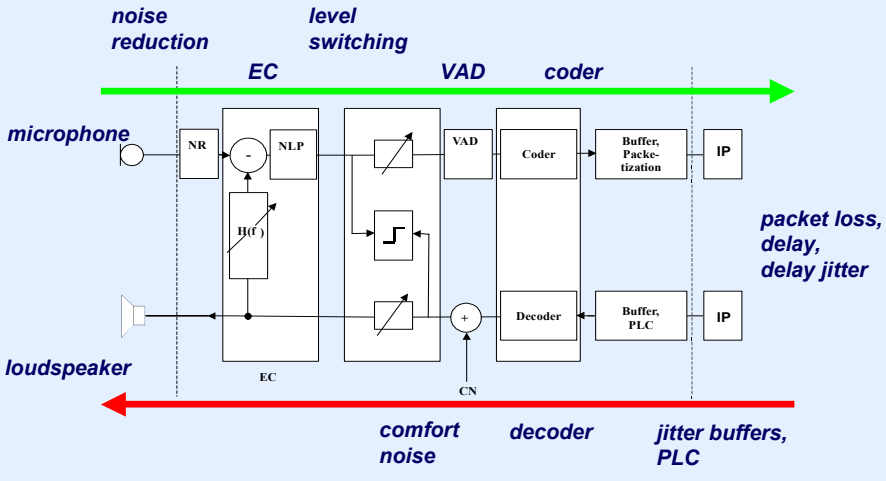
sound quality
naturalness
intelligibility

conversational situation

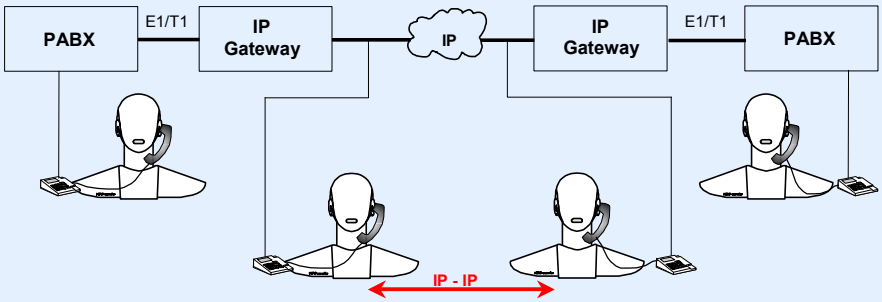
delay
echo (during double talk)
duplex capability

Typical Signal Processing

... in terminals and gateways

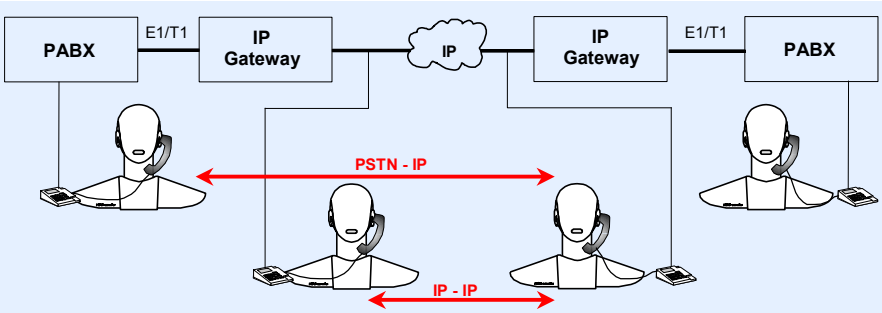


Typical IP Scenarios



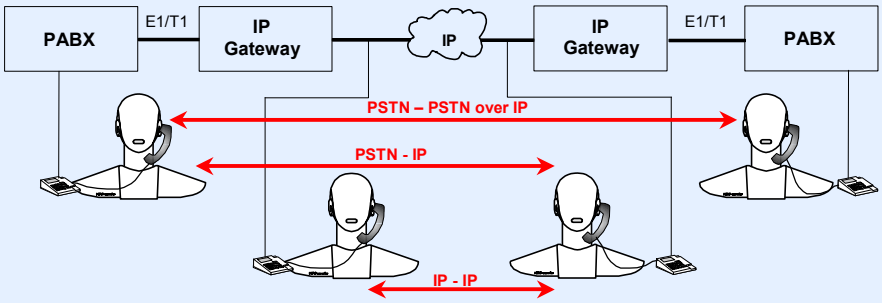
- connection between two IP phones
- all kinds of signal processing in the terminal

Typical IP-Scenarios



- connection between the PSTN and one IP phone
- all kinds of signal processing in gateway and terminal

Typical IP-Scenarios



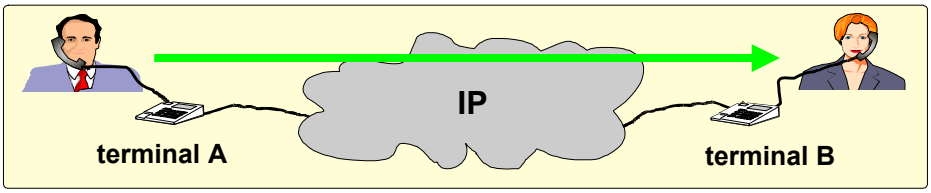
- connection between two PSTN subscribers over IP
- all kinds of signal processing in gateways

Overview

- ***Speech Quality in VoIP***
- ***Parameters influencing speech quality – measurement procedures***
  - Single talk - listening
  - Single talk – talking
  - Double talk
  - Background noise
- ***Summary***

Speech Quality Parameters

Listening Situation

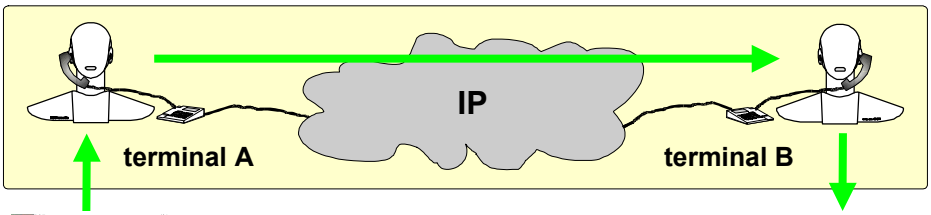


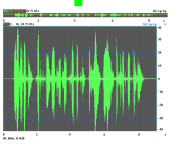
Sound quality, intelligibility, naturalness:

- voice activity detection
- speech coders
- packet loss
- packet loss concealment
- noise reduction
- ...

Testing Techniques

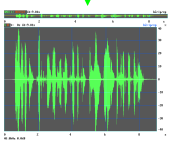
Listening Situation





feeding

recording and analysis



- ***Analysis methods: "traditional" analysis & perceptual speech quality models***

Testing Techniques

The „traditional“ numbers -

to be determined under realistic use conditions

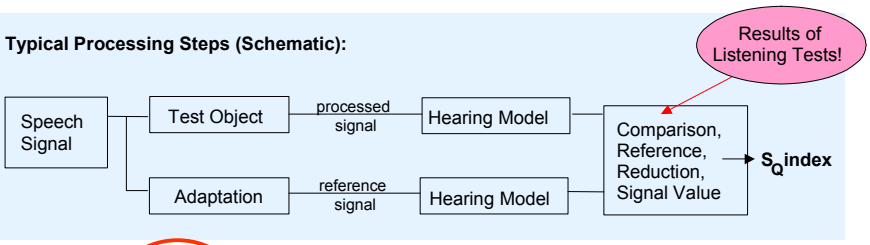
- loudness ratings (SLR, RLR)
- frequency responses
- listener sidetone (LSTR)
- listener echo

Testing Techniques

Instrumental Measures based on Hearing Models:

Modeling the Results of Auditory Tests by Comparison of Reference Speech Signal with Processed Speech Signal

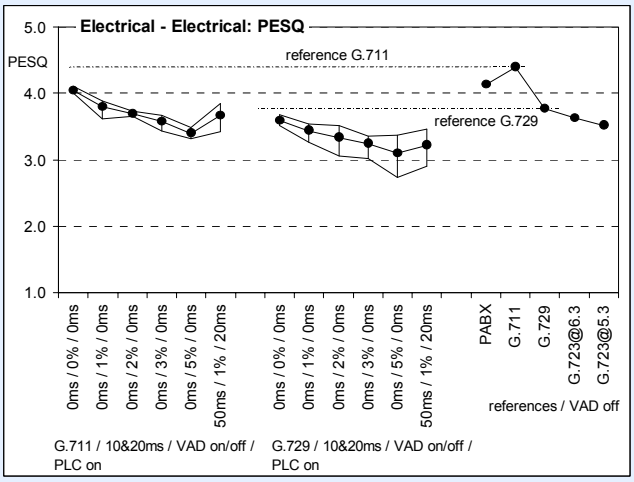
Typical Processing Steps (Schematic):



PSQM	PESQ	PSQM99	PAMS	TOSQA	PACE	VQI
P.861	P.862	KPN	BT	T-Systems	Ascom	Ericsson

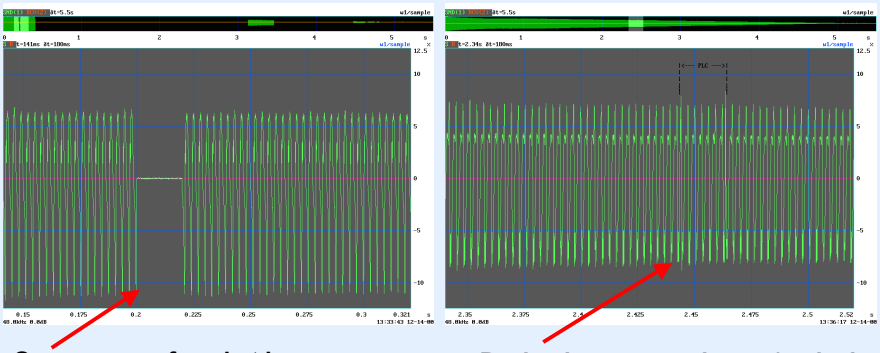
The ITU standard for electrical access

Typical Results 2nd ETSI VoIP



Test Signals and Analysis Methods

Analysis of packet loss and PLC implementation (example)



Occurrence of packet loss (20 ms packet length)

Packet loss concealment (typical implementation), signal discontinuities

Relative Approach

- *Approach: forward estimation based on signal history, comparison with actual signal value*
  - Hearing model
  - Extrapolation in the time domain
  - Interpolation between critical bands
  - Display of estimation error = audible degradation
    - **Relative Approach**
- *Relative Approach takes into account the sensitivity of the human ear*
  - on instantaneous signal variation in time
  - on dominant spectral structures
  - Relative Approach needs no reference signal

Relative Approach

$$Q = f(N,S) + f\left(\sum_{i=1}^{24}\left[\left|F_G(i-1) - F_G(i)\right| \cdot w_1(iF_G(i)) + \sum_{n=1}^{T}\left|F_G(i,n) - F_G(i,n+1)\right| \cdot w_2(i,F_G(i))\right]\right)$$

Basic principle of the Relative Approach:

Comparison between short term and long term averaging of signal energies in critical bands (app. 2 s vs. 2 ms) based on a hearing model [Sottek]

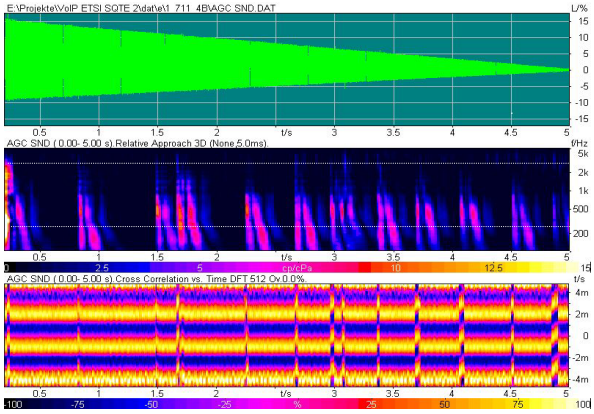
Packet Loss and Concealment

The Reasons Behind the Scores... ?

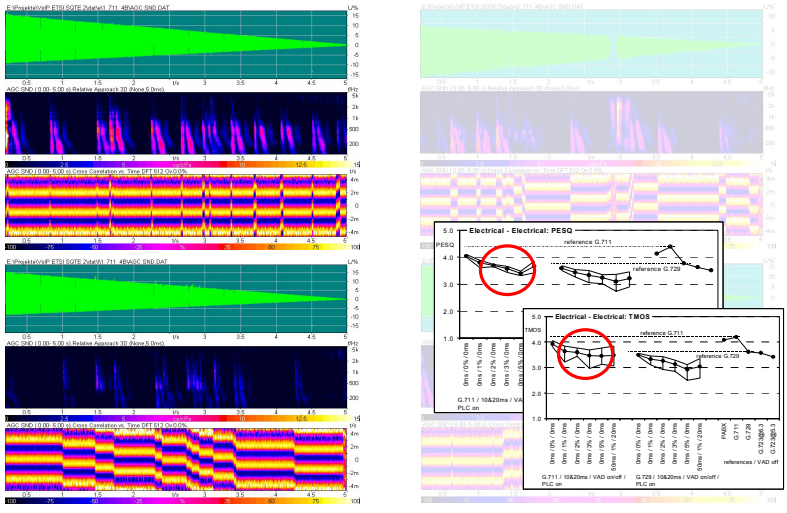
Transmitted time signal (5s)

Relative Approach

Cross correlation analysis vs. time



Packet Loss and Concealment



Testing Techniques

Background noise transmission - Influence of VAD

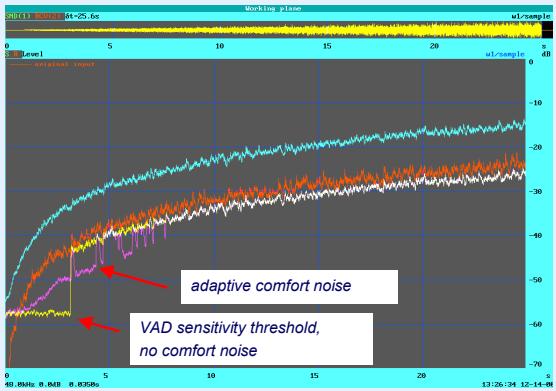
Example: Gateway - IP sim.- Gateway

- *G.711 codecs*
- *no packet loss, no jitter*
- *no additional delay*

red: original test signal

- *noise sequence*
- *Hoth spectrum (P.800)*
- *increasing level vs. time*

yellow, magenta, cyan: different implementations

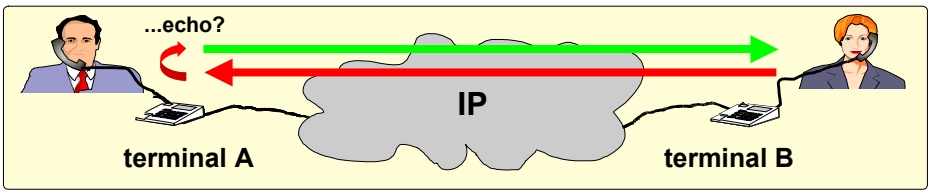


Overview

- ***Speech Quality in VoIP***
- ***Parameters influencing speech quality – measurement procedures***
  - Single talk - listening
  - Single talk – talking
  - Double talk
  - Background noise
- ***Summary***

Speech Quality Parameters

Talking Situation

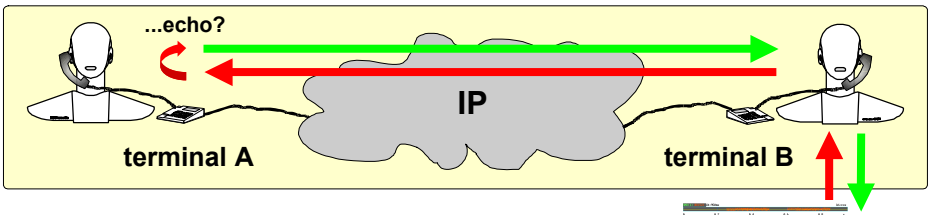


Delay and echo – also with background noise

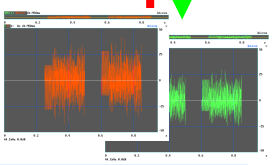
- round trip delay
- echo level and echo characteristic
- implementation of speech echo cancellers
- quality of background noise transmission

Testing Techniques

Talking Situation



feeding, recording and analysis

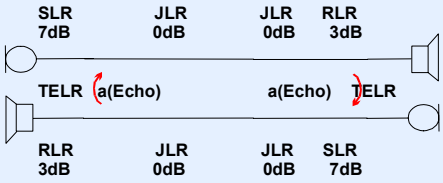


- ***Analysis methods: echo measurements based on Composite Source Signals (CSS) under single and double talk conditions***

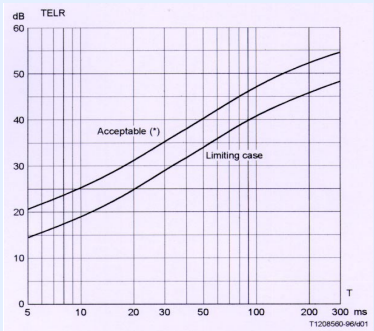
Testing Techniques

Echo under single talk conditions - ITU-T G.131

Requirement on echo loss depends on transmission delay



OLR = 10dB

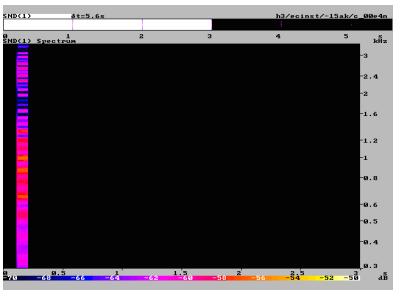


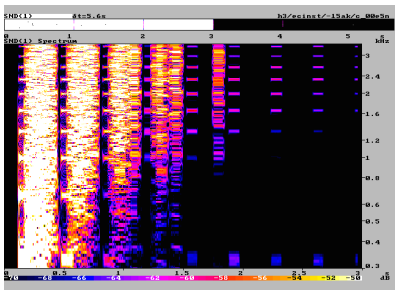
TELR Talker Echo Loudness Rating
T Mean one-way transmission time
(*) The "Acceptable" curve is equivalent to the curve with "1%" probability of encountering objectionable echo.

FIGURE 1/G.131
Talker echo tolerance curves

Echo Tests according to G.168

- *Convergence and steady state residual and returned echo level tests*
- *Convergence test in the presence of background noise*
- *Leak rate test*
- *Infinite return loss convergence test*
- *Non-divergence on narrow-band signals*



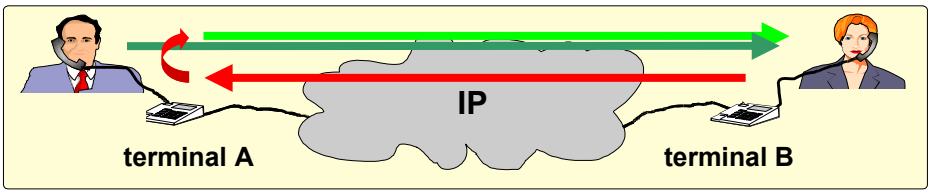


Overview

- *Speech Quality in VoIP*
- *Parameters influencing speech quality – measurement procedures*
  - Single talk - listening
  - Single talk – talking
  - Double talk
  - Background noise
- *Summary*

Speech Quality Parameters

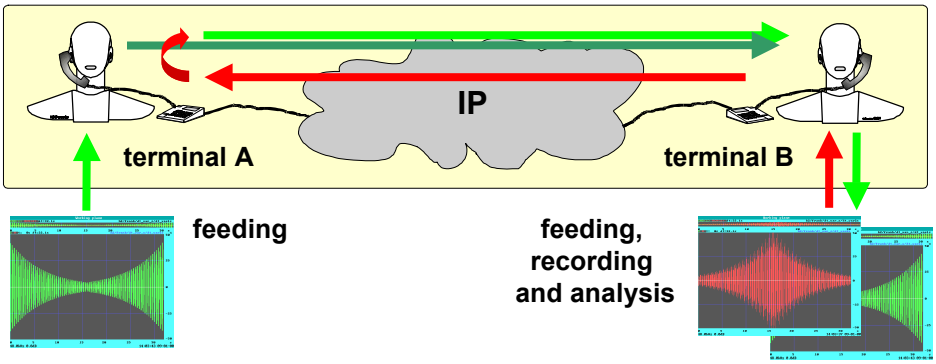
Conversational Situation



- *propagation delay - conversation dynamics impairments*
- *double talk detection (EC implementation)*
- *echo during double talk*
- *level variations during double talk*

Testing Techniques

Conversational Situation



- *Analysis methods: double talk measurements based on two uncorrelated Composite Source Signals*

Testing Techniques

Requirements on echo and switching during double talk:

MOS	≥4.0	4.0-3.5	3.5-3.0	3.0-2.5	2.5-2.0	≤2.0
$TELR_{DT}$ [dB]	≥37	≥33	≥27	≥21	≥13	<13
a_{Hsdt} [dB]	≤3	≤6	≤9	≤12	≤15	>15
a_{Hrdt} [dB]	≤3	≤5	≤8	≤10	≤12	>12

$TELR_{DT}$: talker echo loudness rating during double talk
a_{Hsdt} : attenuation range sending during double talk
a_{Hrdt} : attenuation range receiving during double talk

Testing Techniques

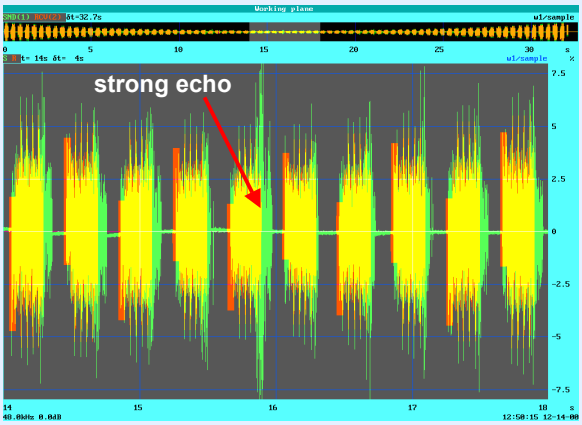
Double talk evaluation: Test result 1

Note strong echo components which occurred only during double talk but not under single talk conditions

green: measured signal

red: original test signal

yellow: overlap



Testing Techniques

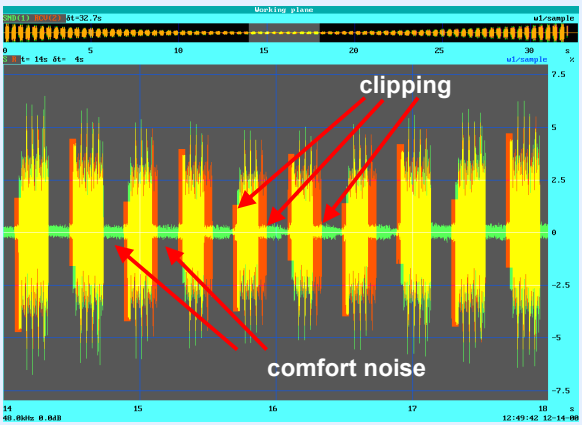
Double talk evaluation: Test result 2

Note clipping at beginning and end of most bursts and comfort noise during pauses

green: measured signal

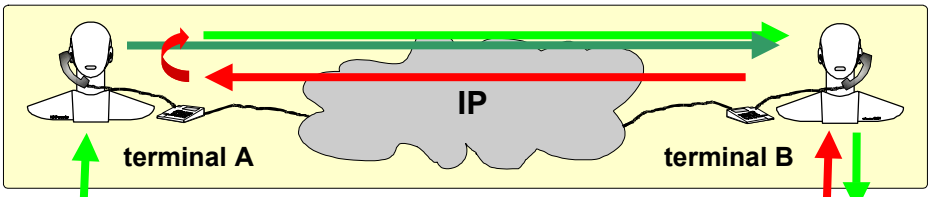
red: original test signal

yellow: overlap



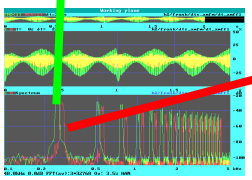
Testing Techniques

Analysis of echo during double talk



feeding

feeding, recording and analysis



- *Analysis method:*
  Separation of echo and near end signal by appropriate filtering

Testing Techniques

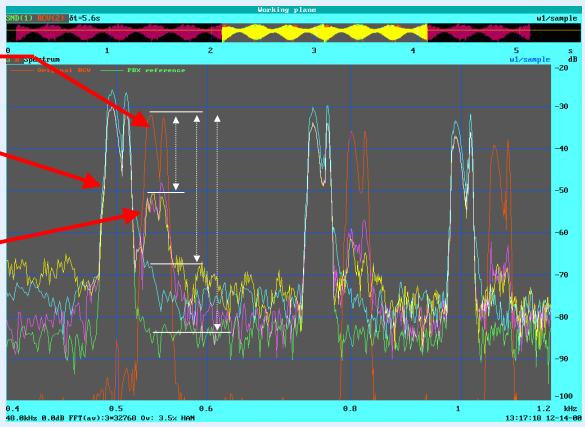
Analysis of echo during double talk

Original test signal in receiving direction

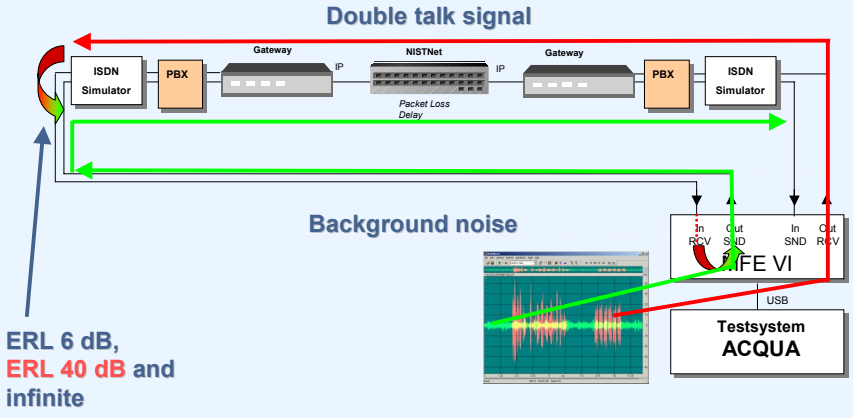
Near end signal

Echo components

Example:
Measured power density spectra between 400 Hz and 1.2 kHz



Background Noise – Double Talk

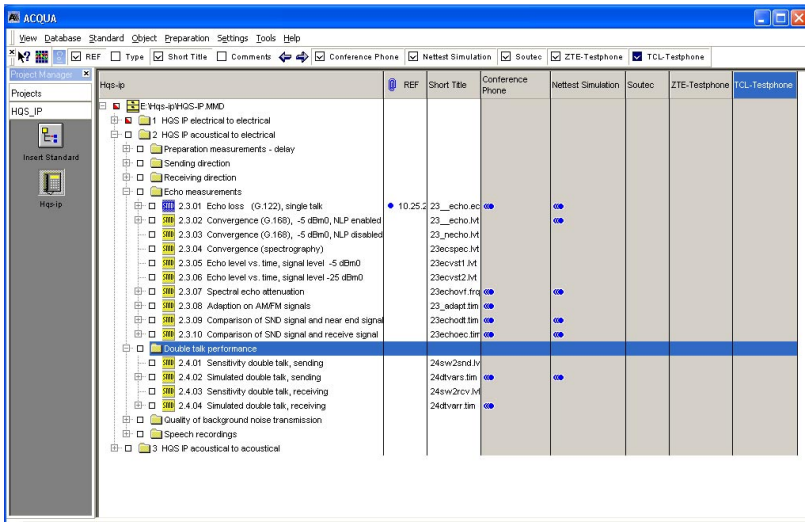


The work in ETSI STQ WI 011

Mouth to Ear Speech Transmission Quality Including Terminals

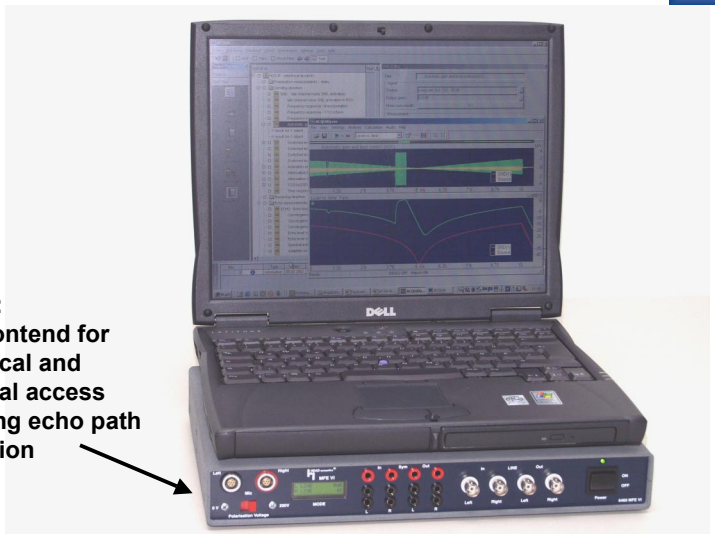
- **Advanced measurement procedures, taking into account the conversational situation**
  - **Quality of background noise transmission**
  - **Double talk performance**
  - **Switching characteristics**
  - **Level Adjustments by Companding or AGC**
  - **Additional Echo disturbances**
  - **Speech Sound Quality**
  - **Loudness and Noise**

Automated Testing with ACQUA

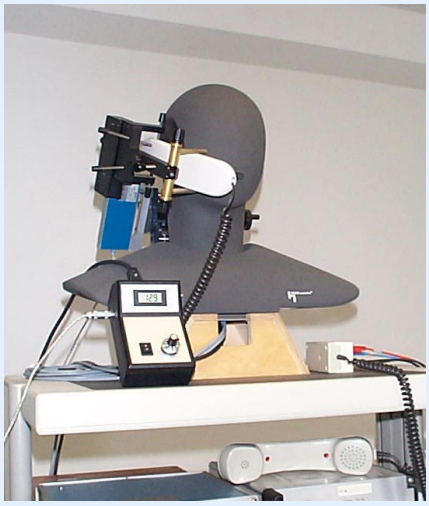


ACQUA with MFE VI

MFE VI:
USB-frontend for acoustical and electrical access including echo path simulation



Acoustical Access



Summary

- ***Listening situation***
  - „Overview“ speech quality tests => P.862 PESQ or TOSQA with acoustical components
  - Detailed investigations => Specific test signals and analysis procedures e.g. „Relative Approach“
- ***Talking situation***
  - G. 168 Tests
  - Additional background noise tests
- ***Double talk/conversational tests***
  - Delay tests
  - Double talk echo tests using specific test signals and analysis procedurs
  - Switching tests using specific test signals and analysis procedurs
  - Background noise tests during double talk
- ***To do:***
  - Tests methods for noise reduction, single number for speech quality,..