Speech Quality in modern Network-Terminal Configurations

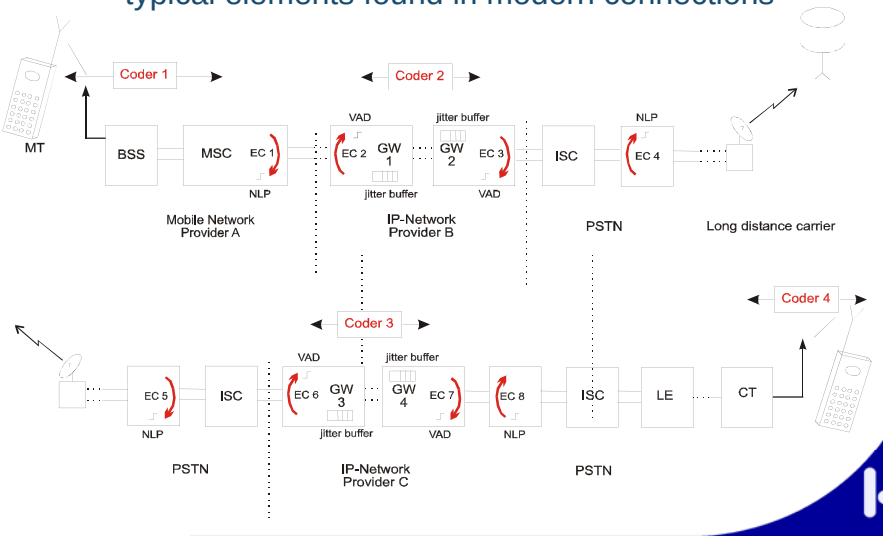
H. W. Gierlich
HEAD acoustics GmbH

Outline

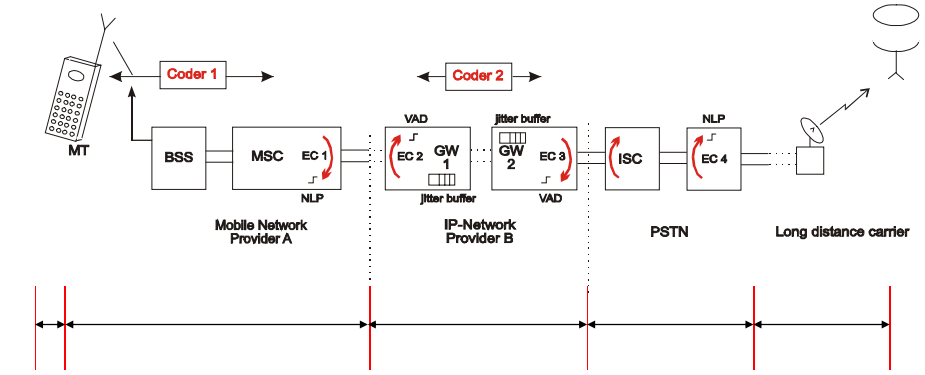
- Overview of configurations and their impact on quality
- Impact of
  - Delay
  - Coding
  - Noise and noise cancellation
  - Insufficient echo loss/improper echo cancellation
  - Cascading of signal processing
  - Missing signaling information between endpoints and network elements
- Summary

Terminal - Network

- typical elements found in modern connections



Delay aggregation (I)



t_{MT1} (10-50ms)

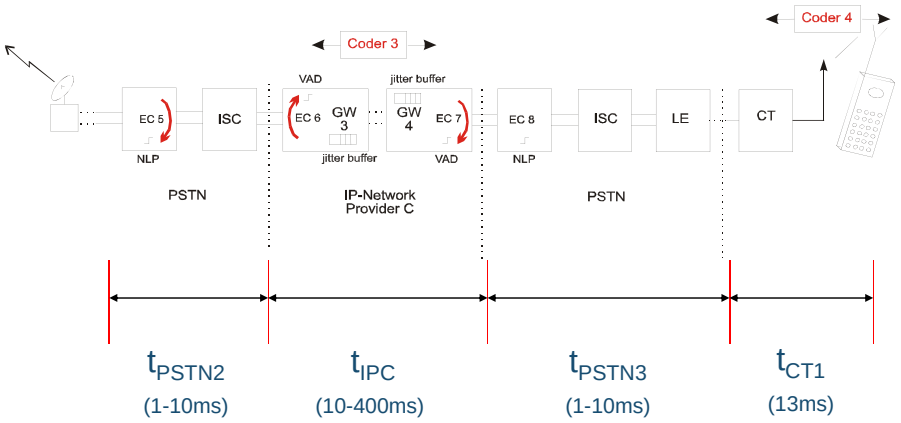
t_{MNA} (≈ 80ms)

t_{IPB} (10-400ms)

t_{PSTN1} (1-10ms)

t_{LDC} (≈ 240ms, Satellite) (≈ 20 ms, fiber)

Delay aggregation (II)



t_{PSTN2} (1-10ms) t_{IPC} (10-400ms) t_{PSTN3} (1-10ms) t_{CT1} (13ms)

Delay aggregation (III)

$$t_{oneway} = t_{MT1} + t_{MNA} + t_{IPB} + t_{PSTN1} + t_{LDC} + t_{PSTN2} + t_{IPC} + t_{PSTN3}$$

Best case

= (10 + 80 + 10 + 1 + 20 + 1 + 10 + 1 + 13)ms = ***146ms***

Worst case

= (50 + 80 + 400 + 10 + 240 + 10 + 400 + 10 + 13)ms = ***1113 ms***

E-model: Influence of delay

If no terminal impairment except TCLw:

146 ms/55 dB TCLw

R_	87,9
MOS (4.5)	4,2

1113 ms/55 dB TCLw

R_	26,6
MOS (4.5)	1,5

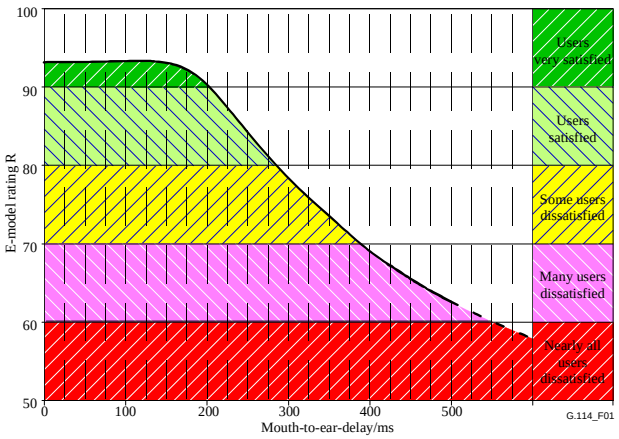
146 ms/35 dB TCLw

R_	82,1
MOS (4.5)	4,1

1113 ms/35 dB TCLw

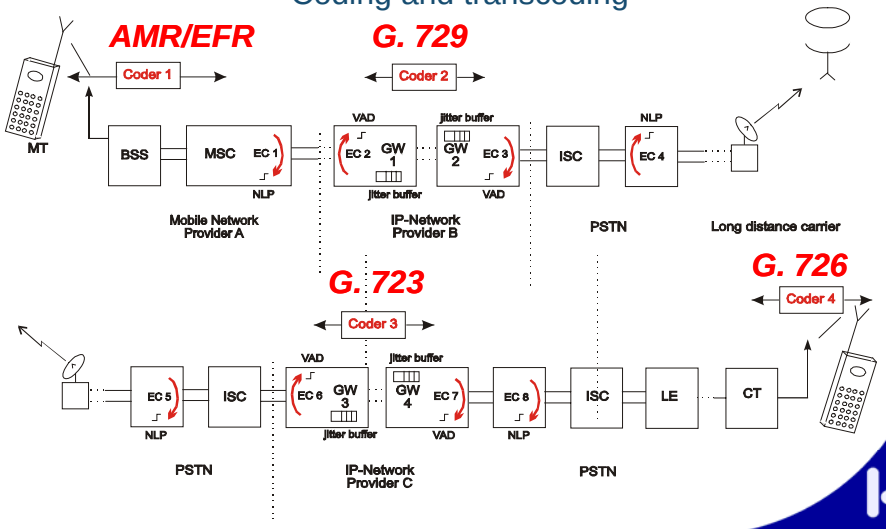
R_	4,3
MOS (4.5)	1

ITU-T G.114: Effect of transmission time



Terminal - Network

- Coding and transcoding



E-Model: The effect of codec tandeming

	Type	Ie (G.113)
Coder 1	EFR (ACELP)	5
Coder 2	G. 729 (CS-ACELP)	10
Coder 3	G. 723 (ACELP, 5.3 kbit/s)	19
Coder 4	G. 726 (ADPCM, 32 kbit/s)	7

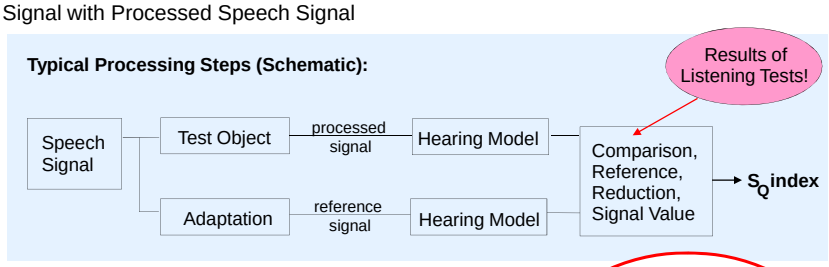
=> Ie(tot) = 5+10+19+7 = 46

*R_*	47,2
MOS (4.5)	2,43

Objective Testing based on ITU-T P.862 and others

Instrumental measures based on hearing models - Perceptual models:

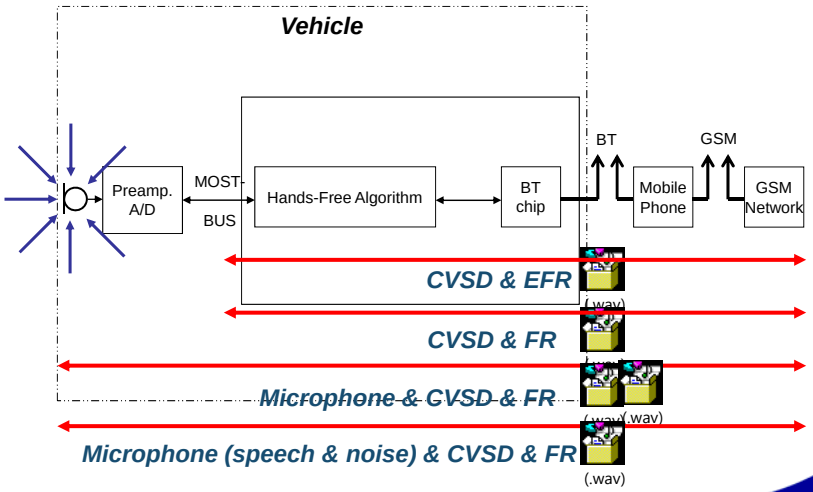
Modeling the Results of Auditory Tests by Comparison of Reference Speech Signal with Processed Speech Signal



PSQM *P.861* ***PESQ*** *P.862* ***PSQM99*** *KPN* ***PAMS*** *BT* ***TOS*** *T-Syst*

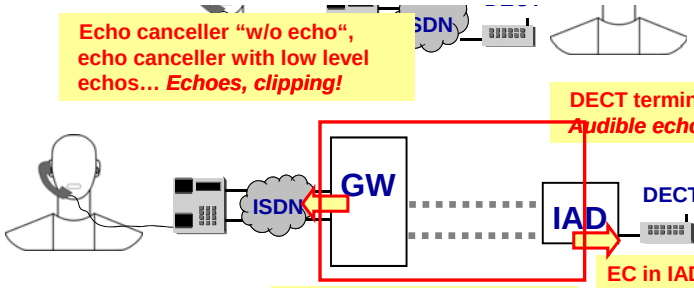
ITU-T P.OLQA

Example: Tandeming of Codecs



Echo loss in IP – DECT configurations

- Assuming… G.711 speech coder, no compression
- Assumimg… no transcoding
- Assuming no VAD, sufficient bandwidth



Echo canceller "w/o echo", echo canceller with low level echos… ***Echoes, clipping!***

DECT terminals, delay… . ***Audible echoes!***

EC in IAD!... ***DECT delay, DECT distortions!***

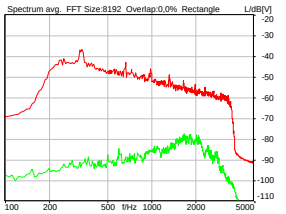
Coding, packetization, buffer,… ***Delay! …Degrades conversational quality!***

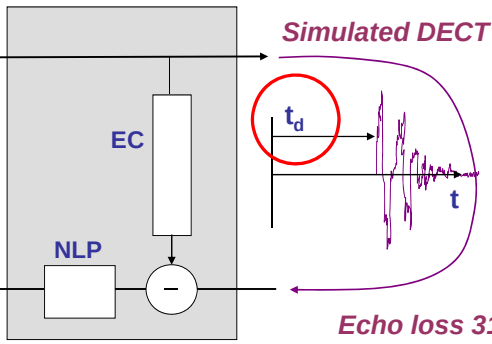
Packet loss and jitter under network load, PLC, jitter buffer… ***Artifacts, delay!***

Echo Cancellation in IAD

Echo Performance

Echo loss 36.7 dB





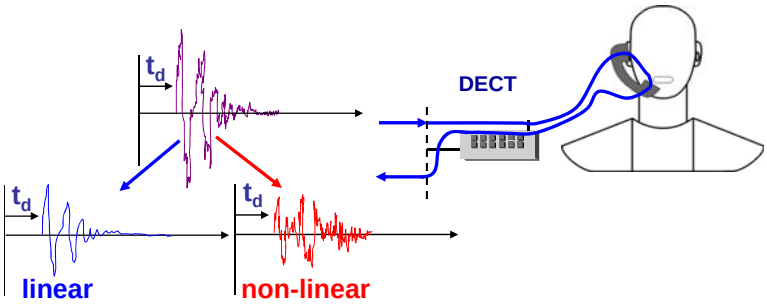
Simulated DECT echo

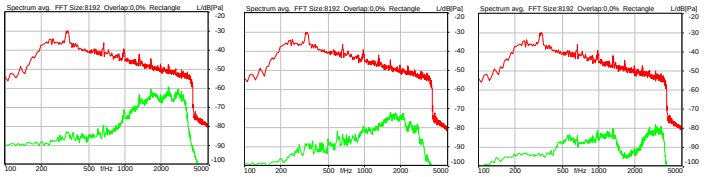
Echo loss 31 dB

t_d 30 ms

+ Network delay 100, 150, … ms !!!

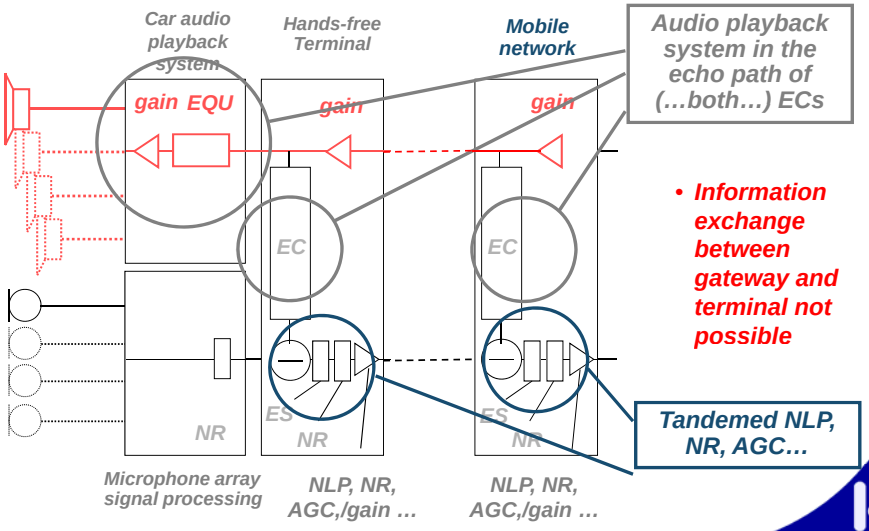
DECT Echoes





Echo loss 17.8 dB | Echo loss 31.0 dB | Echo loss 36.3 dB

Tandeming of signal processing: Car hands-free and network

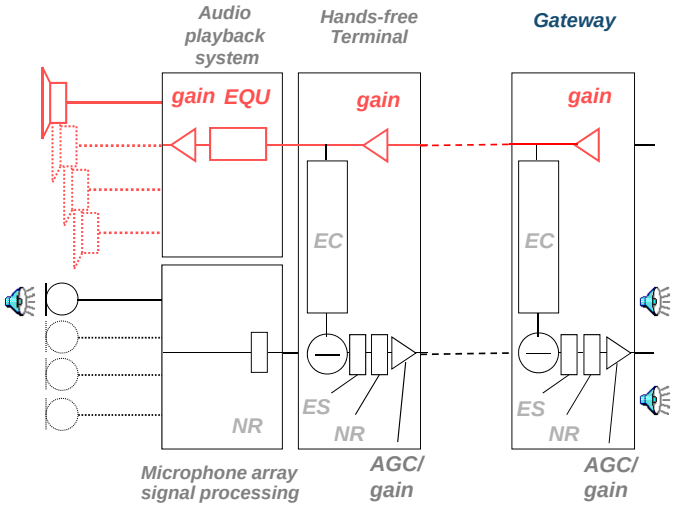


Audio playback system in the echo path of (…both…) ECs

- Information exchange between gateway and terminal not possible

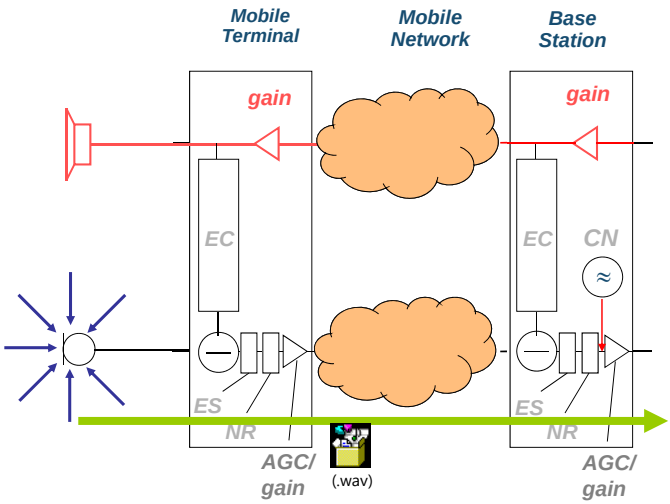
Tandemed NLP, NR, AGC…

Tandeming of noise cancellation



- *Tandeming of noise cancellation:*
- *Typical result*
- *Optimized settings*

Background noise transmission



- *Influence on double talk performance*
- *Both EC & ES active*

Objective Testing based on ETSI EG 202 396-3

MOS for

- **speech quality (S-MOS)**
- **noise transmission quality (N-MOS)**
- **overall quality (G-MOS)**

		Subjective MOS			Objective MOS		
	Noise	S	N	G	S	N	G
	Car 1	4.7	4.7	4.6	4.8	4.8	4.8
	Cafeteria	4.2	2.2	3.1	4.0	2.1	3.1
	Office	4.0	2.8	3.5	4.2	2.6	3.5
	Car 2	2.0	2.3	1.9	2.3	2.3	1.8
	Road	1.4	1.3	1.0	1.3	1.0	1.0

Double Talk – Performance Limits

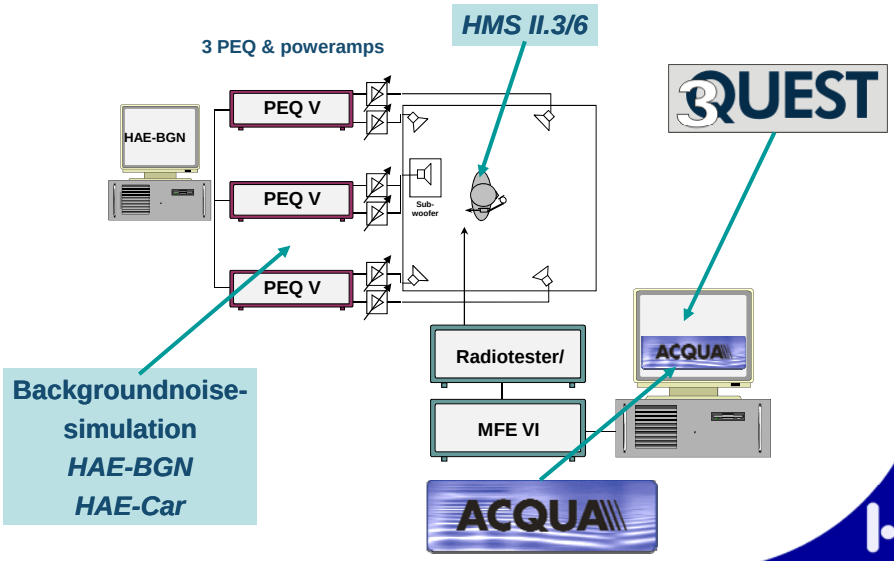
Characterization of HFT acc. to ITU-T P.340

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

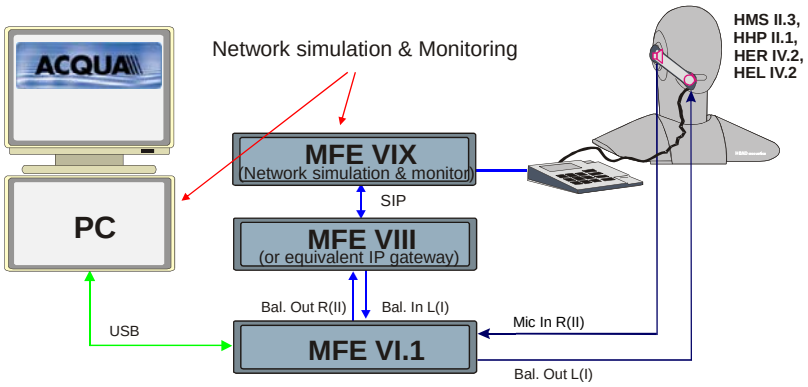
	Type 1	**Type 2**			**Type 3**
		Type 2a	**Type 2b**	**Type 2c**	
$TELR_{DT}$ [dB]	≥ 37	≥ 33	≥ 27	≥ 21	< 21
A_{Hsdt} [dB]	≤ 3	≤ 6	≤ 9	≤ 12	> 12
A_{Hrdt} [dB]	≤ 3	≤ 5	≤ 8	≤ 10	> 10

Terminal test Setup

Background Noise Test



IP-Test Setup



Summary

- Potential problems in a connection may occur due to tandemed signal processing in terminals and networks
- Quality evaluation needs to involve all components of a connection and their interaction!
- Currently no signalling is provided between terminals and networks for information exchange about signal processing active in the different devices
- Related work:
  - G.mdcspne started in SG16 (establish the best arbitration among Network Signal Processing Equipments and terminal signal processing)
  - G.115.x sub series (Logic and protocols for the control of signal processing network elements and functions) may potentially be enhanced to support this capability