

# Speech Quality Assessment for Wideband Communication Scenarios

H. W. Gierlich, S. Völl, F. Kettler  
(HEAD acoustics GmbH)

P. Jax  
(IND, RWTH Aachen)



– Workshop on Wideband Speech Quality in  
Terminals and Networks



supported by:



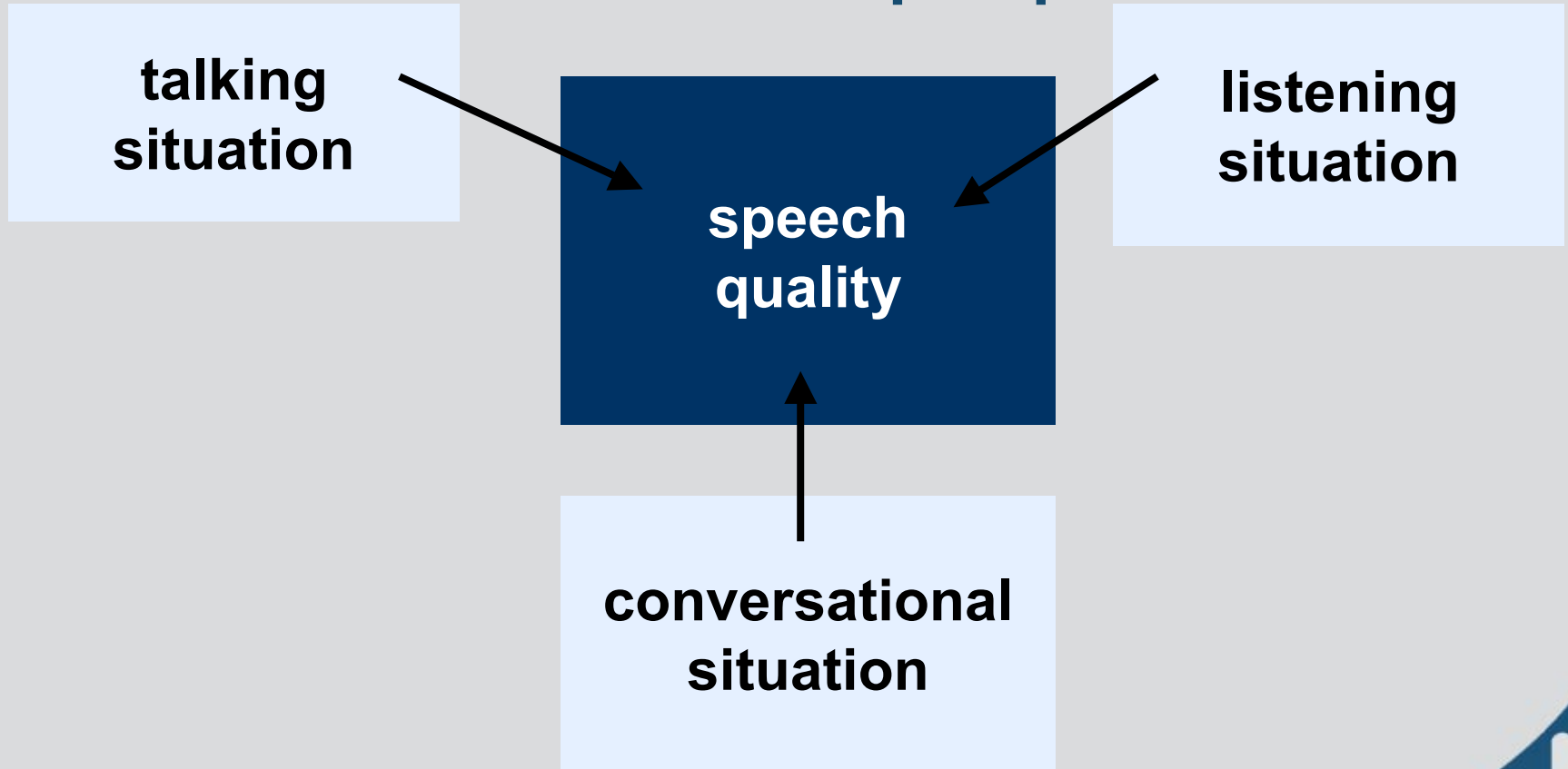
# Outline

- General aspects of speech quality in wideband systems
- Subjective evaluations
  - Conversational tests
  - Speech intelligibility
  - Background noise transmission
  - Echo tests
- Summary



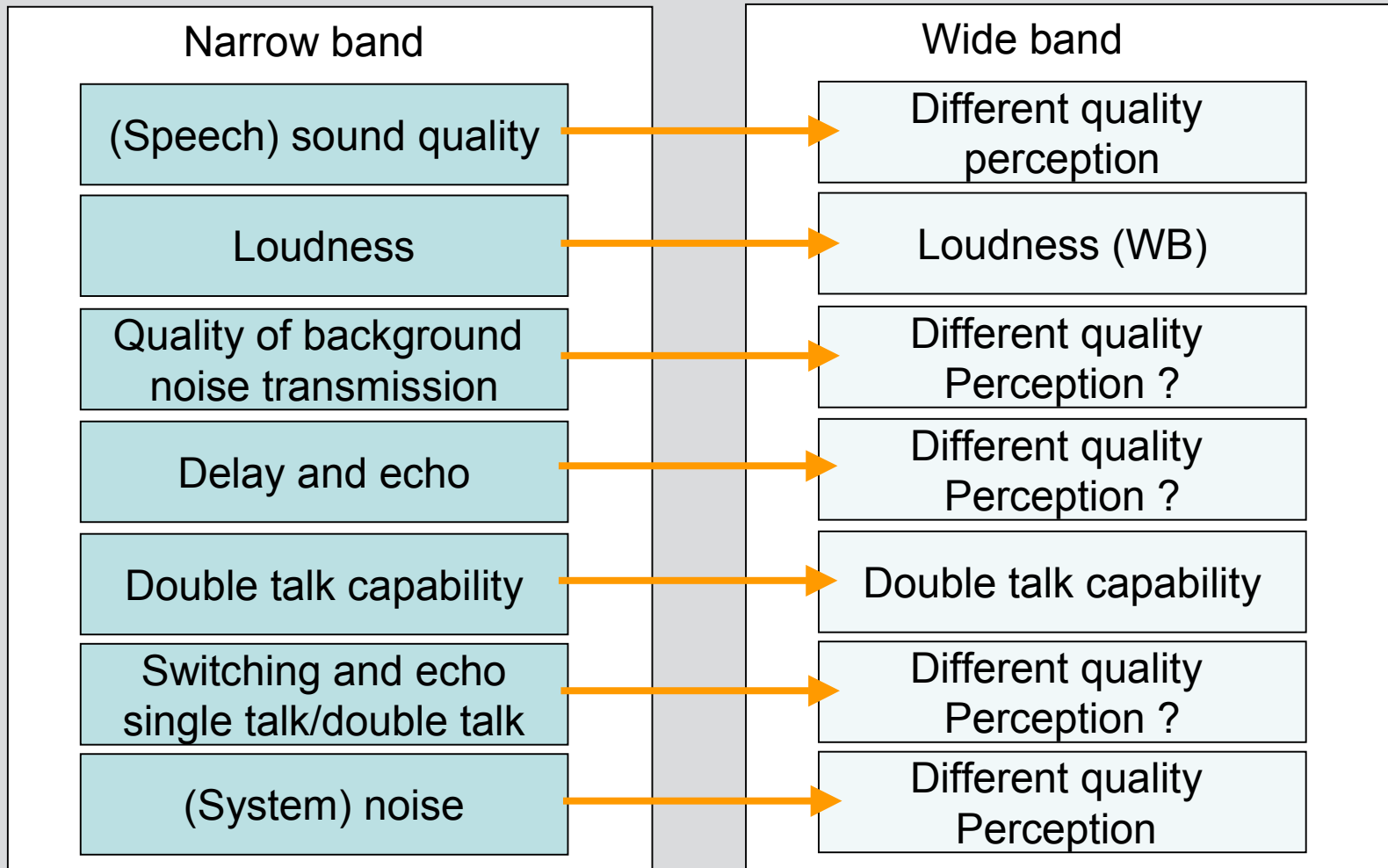
# Speech Quality Parameters

... from the user's perspective



# Auditory Parameters

...contributing to speech quality:



## Roadmap for the development of objective measurements

1. Conversational tests → parameter identification (qualitative)
  2. Listening-only tests → quantitative judgement
  3. Development of objective measurement methods to reproduce the results of the LOT
- Quality evaluation of wideband systems *without* subjective tests



# Outline

- General aspects of speech quality in wideband systems
- **Subjective evaluations**
  - Conversational tests
  - Speech intelligibility
  - Background noise transmission
  - Echo tests
- Summary



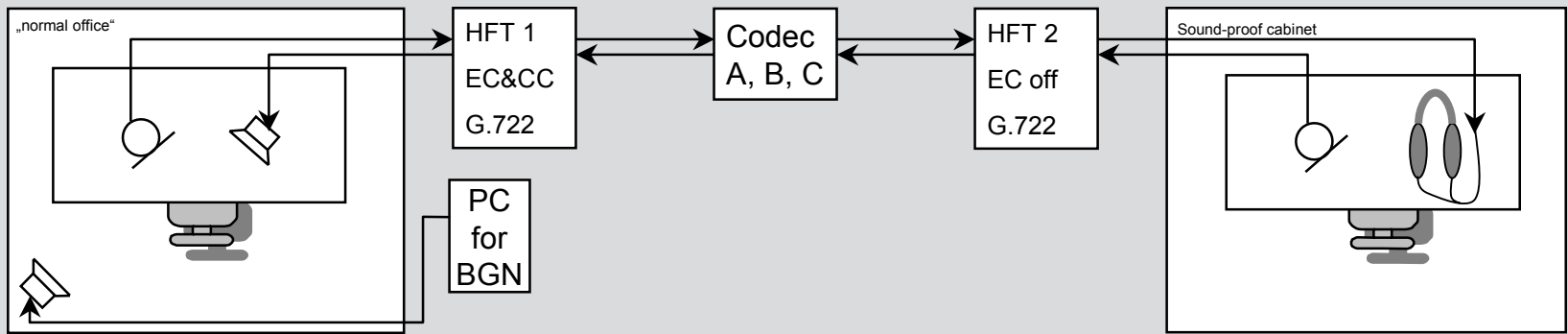
# Conversational Tests

- Purpose: identification of parameters characterizing the communicational quality in wideband systems
- Test conditions:
  - Experts tests
  - “Kandisky”-test
  - 4 wideband codecs under test
  - 3 conditions for each codec: “normal” conversation, with music in office room, with babble in office room
  - “free answering”



# Setup and test procedure

- Setup:



- Test procedure:

- Different codecs included
- One echo-canceller for all tests





# Results

- **Sorting the comments in categories:**
  - Speech sound quality, echo behavior, Quality of background noise, others (e.g. noise, clipping)
- **Example: speech sound quality**

codec	comments
G.722	sounds naturally, high dynamic ...
AMR-WB	sounds rough, naturally, distorted, ...
AMR-WB2	high dynamic, hollow, clank, ...
BWE	rattles, crackles, blunt sound, ...

# Conclusion

- Relevant parameters to be studied further:
  - sound of speech
  - Echo: level, masking, intelligibility
  - quality of background noise transmission
  - Noise
  - Double talk
  - Switching/clipping
- Design of listening-only tests concerning
  - speech intelligibility: narrow band vs. wide band
  - quality of transmitted background noise
  - annoyance of echo



# Outline

- General aspects of speech quality in wideband systems
- Subjective evaluations
  - Conversational tests
  - **Speech intelligibility**
  - Background noise transmission
  - Echo tests
- Summary



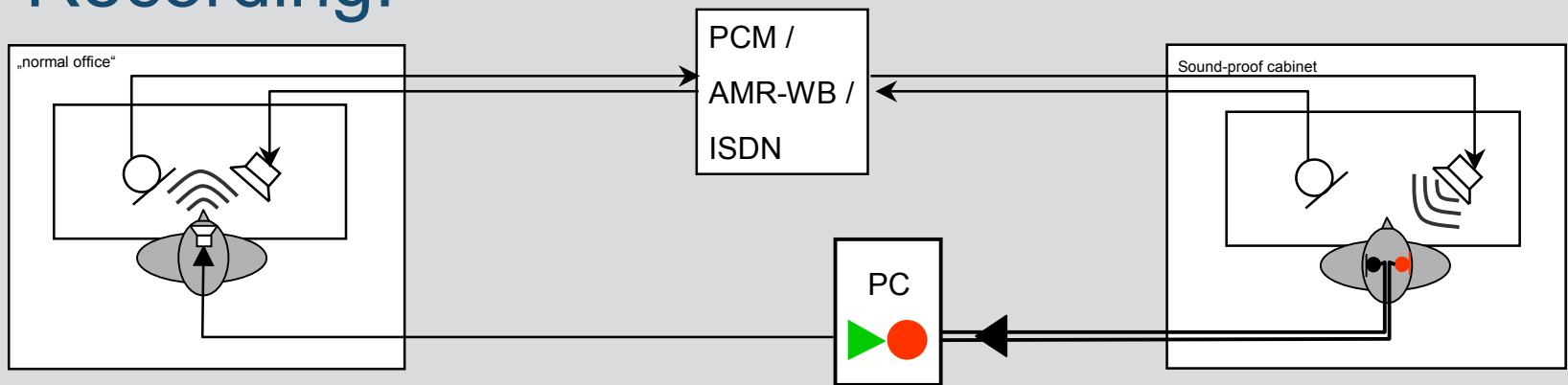
# Speech intelligibility test

- Sensitive test: logatom-test
- Consonant – vowel – consonant
- Informal test:
  - 3x 12 test persons,
  - 29 logatoms
- Test persons note the “word” they understood



# Recording & Listening

- Recording:

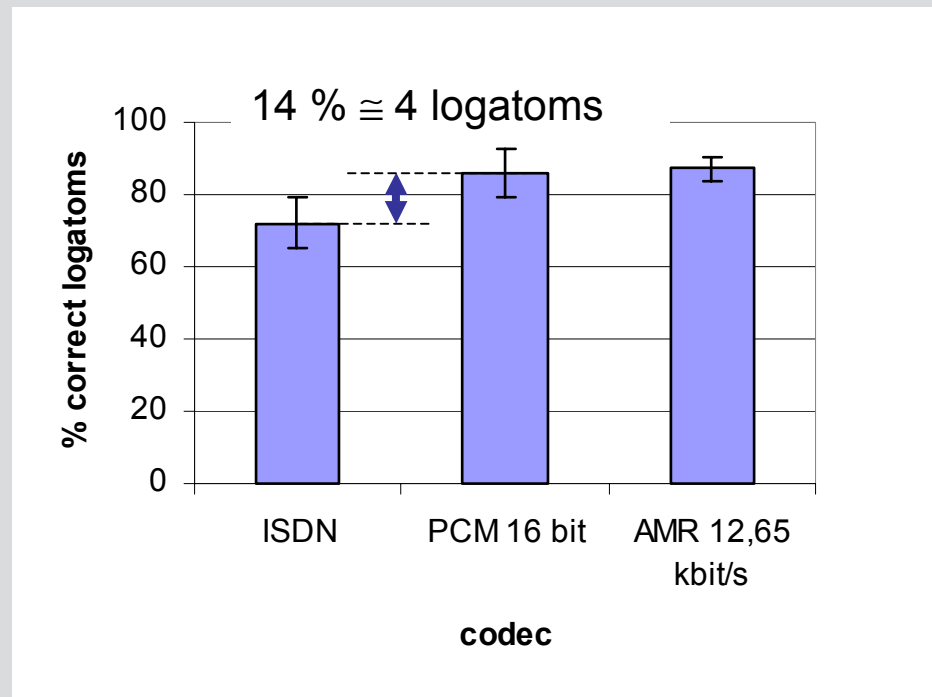


- Listening:  
test persons listen to the artificial head recordings



# Results

- Increased intelligibility for wideband codecs



# Outline

- General aspects of speech quality in wideband systems
- Subjective evaluations
  - Conversational tests
  - Speech intelligibility
  - **Background noise transmission**
  - Echo tests
- Summary



## Background noise assessment: music

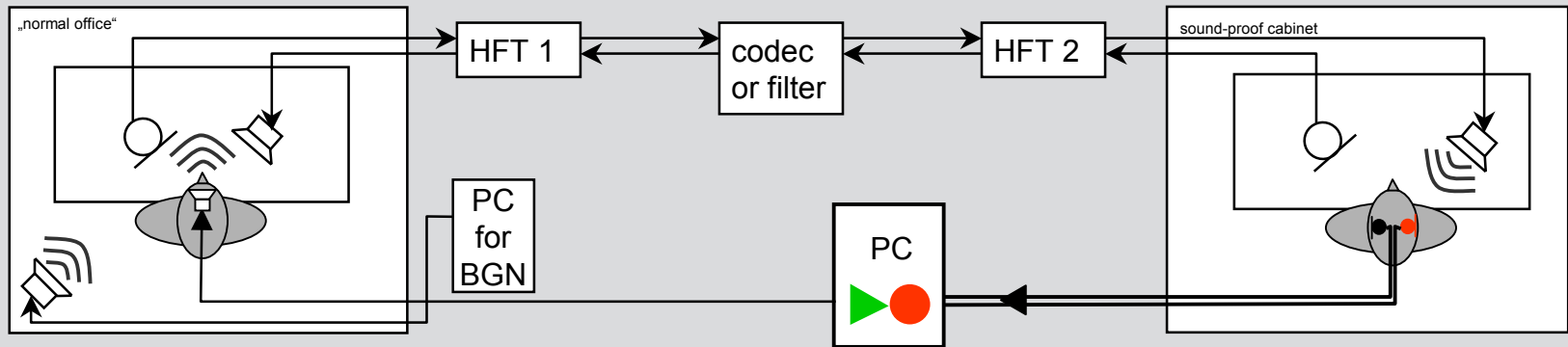
- Background noise: additional information about talkers environment
- Tests with
  - untrained persons: assessment on a 5 point MOS scale
  - experts: assessment on a 5 point MOS scale *and* giving reasons why
- 16 different codecs under test





# Background noise tests

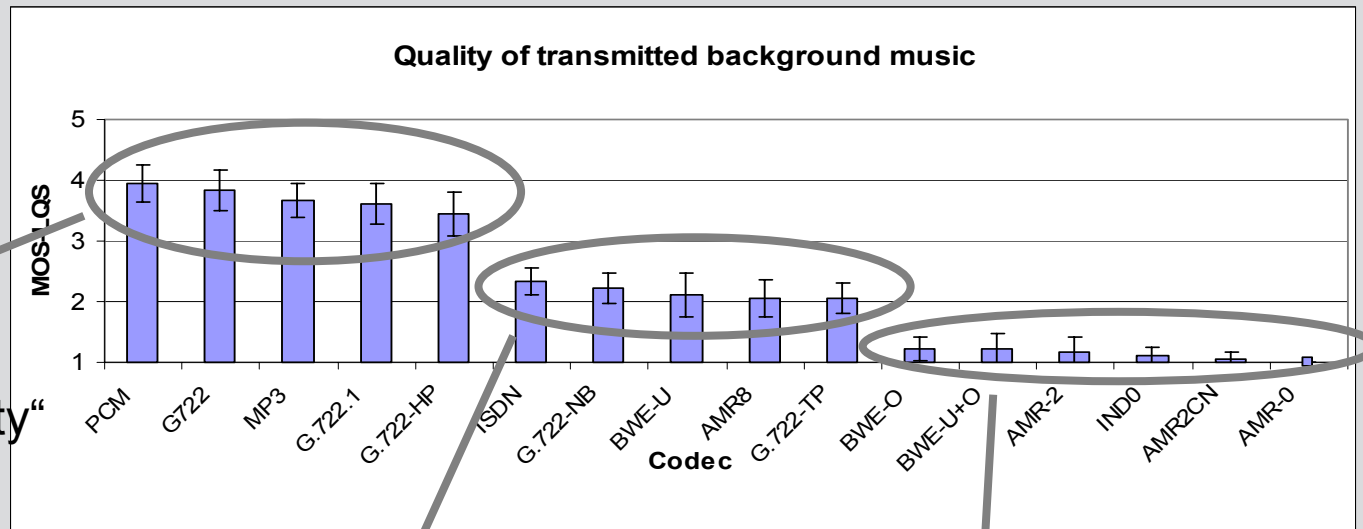
- Recording listening samples:



- Listening:  
test persons listen to artificial head recordings
- ACR scale:  
excellent – good – fair – poor – bad

# Results

- 3 quality levels with significantly different MOS - values



- wideband
- good „intelligibility“ of music



- narrowband
- good „intelligibility“ of music



- mostly wideband
- bad „intelligibility“ of music



# Outline

- General aspects of speech quality in wideband systems
- Subjective evaluations
  - Conversational tests
  - Speech intelligibility
  - Background noise transmission
  - **Echo tests**
- Summary



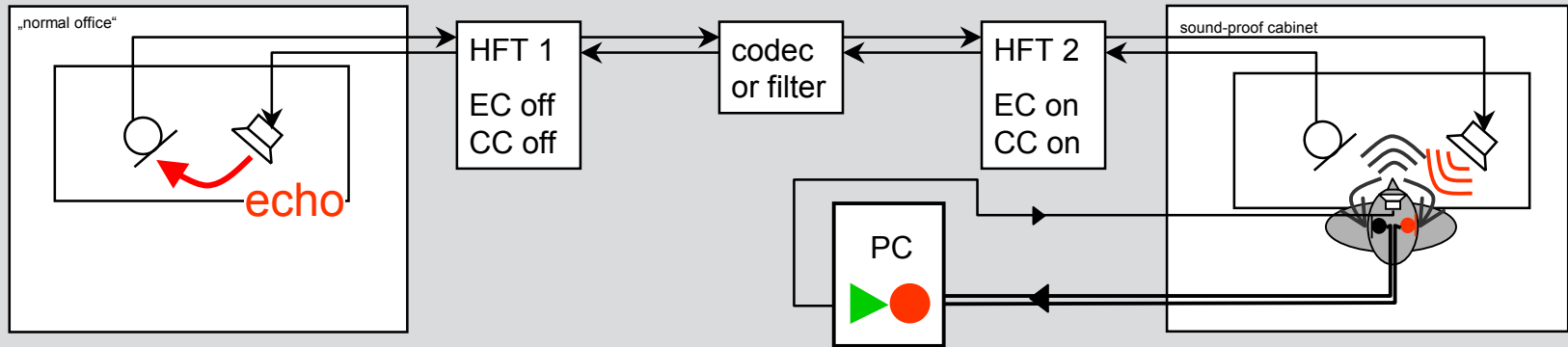
# Echo annoyance test

- Using hands-free telephones → echo disturbances a dominant problem
- Investigation of the annoying aspects of echo using *wide-band* links:
  - influence of echo sound,
  - influence of echo level,
  - influence of codec, ...
- Mean one-way transmission time constant for all listening samples: 170 ms



# Echo annoyance tests

- Recording:



- Listening:

test persons listen to the artificial head recordings:

→ direct speech + echo



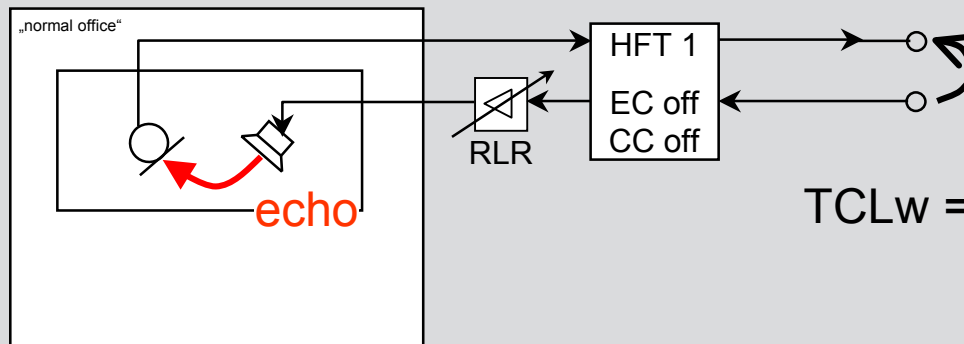
# Tests & assessment

- Tests with
  - untrained persons: assessment on a 5 point MOS scale
  - experts: assessment on a 5 point MOS scale *and* giving reasons why
- DCR scale:
  - 5 – echo is inaudible
  - 4 – echo is audible, but not annoying
  - 3 – echo is slightly annoying
  - 2 – echo is annoying
  - 1 – echo is very annoying



# Echo levels

- $TCL_w$  acc. ITU-T P.79



$TCL_w = \begin{cases} 27 \text{ dB} \rightarrow \text{„low“ echo level} \\ 21 \text{ dB} \rightarrow \text{„medium“ echo level} \\ 13 \text{ dB} \rightarrow \text{„high“ echo level} \end{cases}$



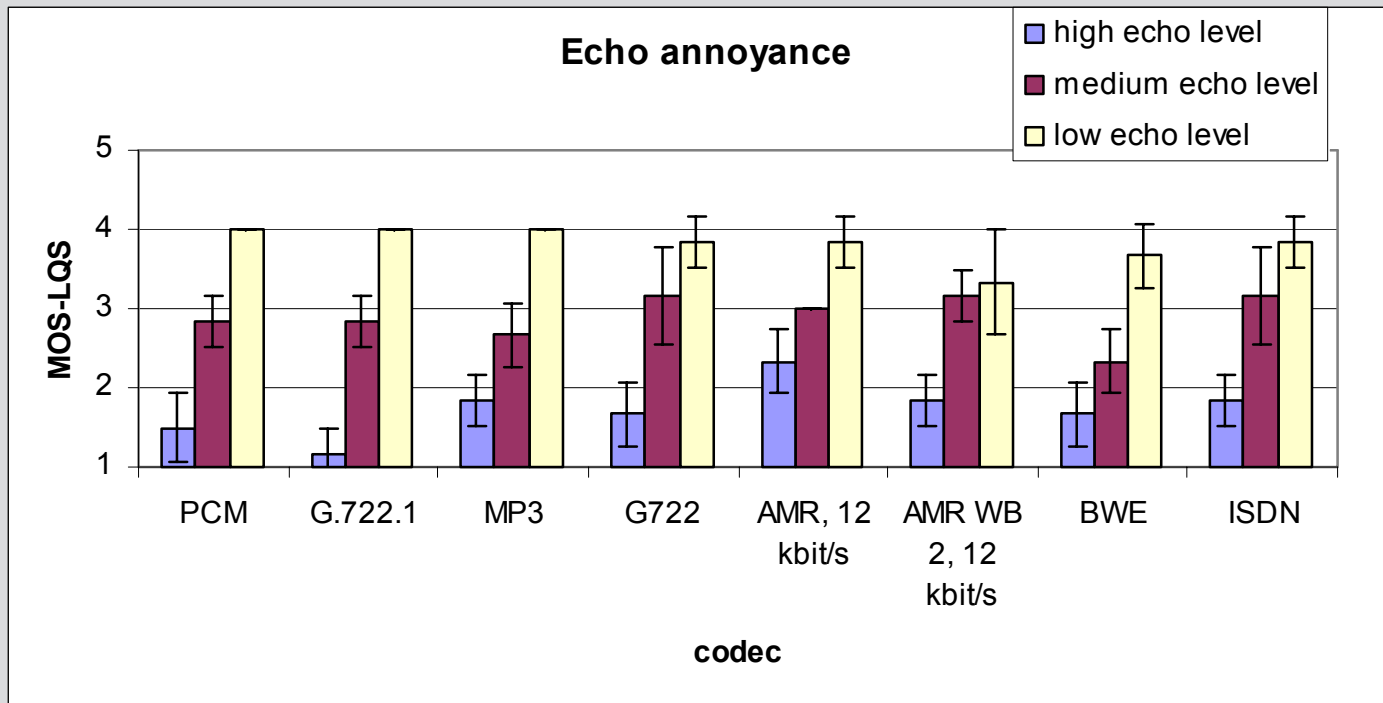
- Note: hands-free on both sides, SLR = 7dB, RLR = 5dB (including HFT correction of 14 dB) => TELR(max) = 39dB

→ Investigation of codec and echo level



# Results

- Comparison of annoyance by echo level (experts)

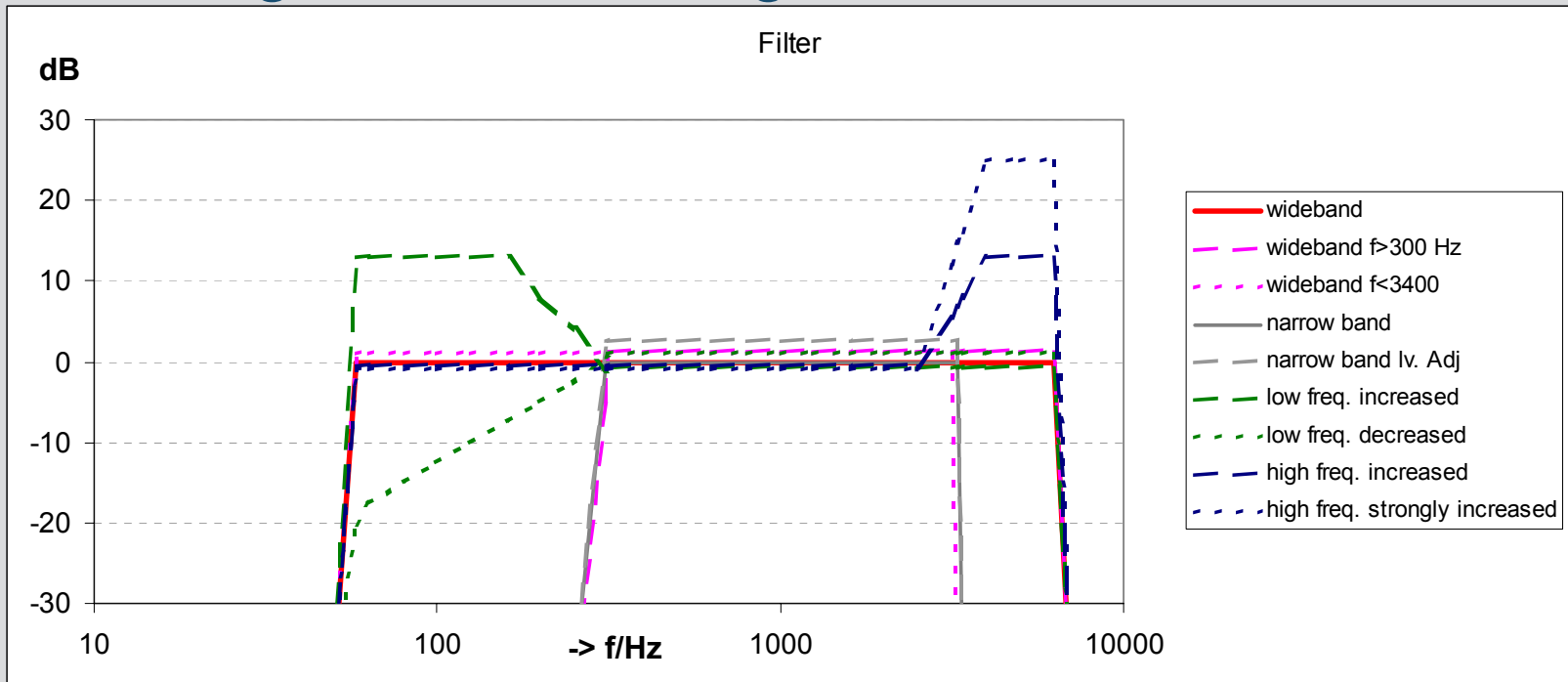


- Differences for echos with the same echo level
- Echo masked by direct speech



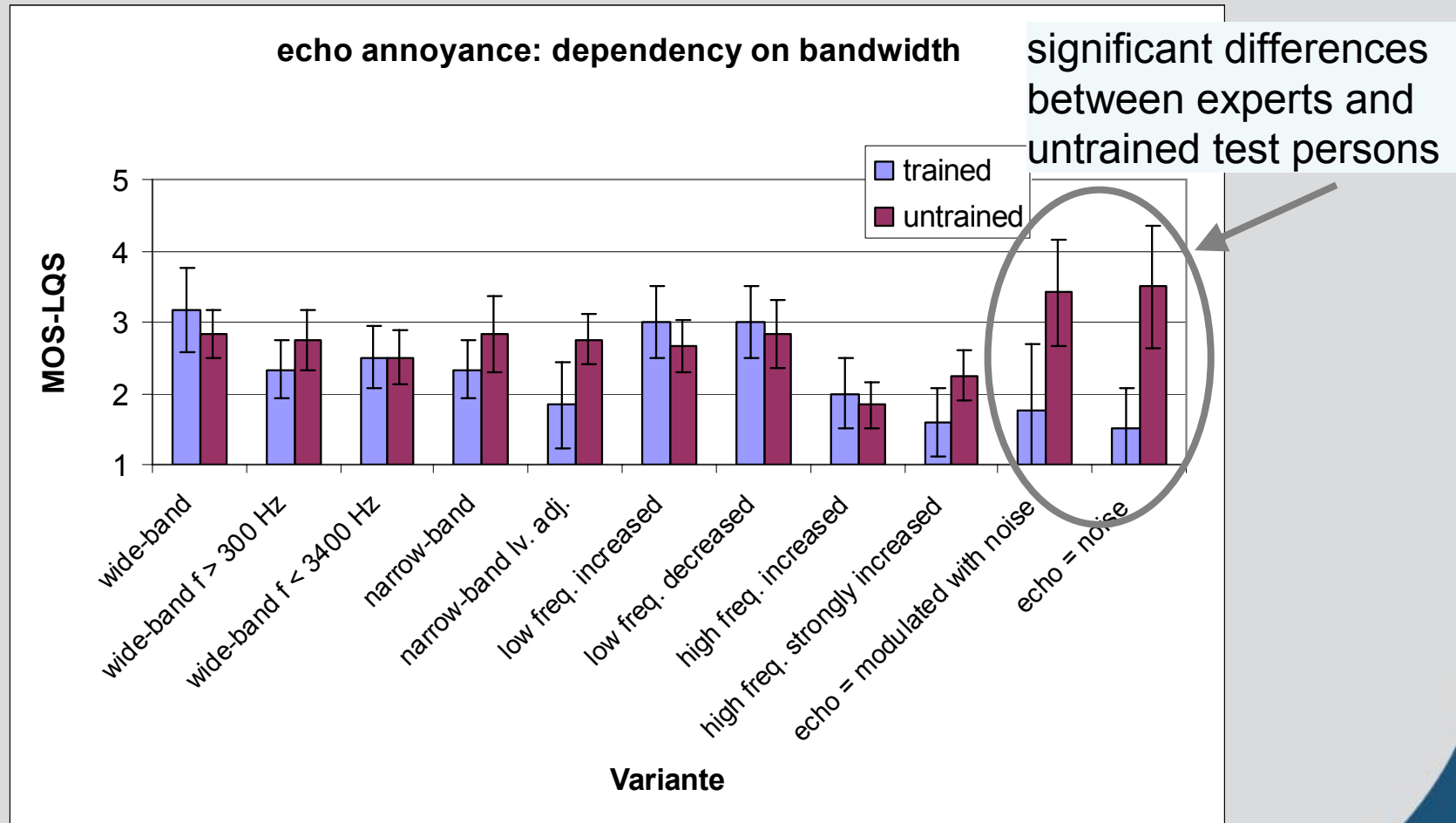
# Influence of the bandwidth

- Filtering of the echo signal



- Speech modulated noise (two examples)
- Level adjustment to  $TCL_w = \rightarrow$  „medium“ echo level

# Results



# Results echo annoyance

- **High frequencies → very annoying**
- **Wide-band and low freq. → slightly annoying**
- **Experts more critical than untrained test persons**
- **Speech modulated noisy echo:**
  - **Experts: very annoying → no advantages**
  - **Untrained: felt insecure**



# Summary (1)

- **Background noise transmission - relevant aspects:**
  - Bandwidth
  - „intelligibility“ / brightness / low distortion (small difference to the original)
- **Echo annoyance - relevant aspects:**
  - Level
  - Masking properties
  - Distortion and frequency characteristics
- **Additional parameters to be investigated subjectively:**
  - Noise
  - Switching/clipping
  - Double talk behavior



## Summary (2)

- **To do:**
  - **Additional subjective testing**
  - **Deriving/Adaptation of methods to measure**
    - Sound quality
    - Echo
    - Background noise transmission
    - Double talk performance
    - Switching/clipping
    - Noise

