Objective measures of speech quality in hearing aids: prediction of listening effort reduction by noise reduction algorithms

Rainer Huber (HörTech),

Michael Schulte, Matthias Vormann (Hörzentrum Oldenburg),

Josef Chalupper (Siemens Audiologische Technik)

Objective

Aim of the study:

Benchmarking of noise reduction programs of four commercial hearing aids (HAs)

- Subjective tests: intelligibility, <u>listening effort</u>, overall preference
- Objective speech/audio quality measures

Stimuli

- Male speech (from Oldenburg sentence test)
  + speech-simulating noise and airplane cabin noise
  at different SNRs [-16...14 dB, 3 dB steps]

- Processing by 4 HAs, with and without NR;
  HAs fitted to average hearing loss of 20 test subjects

- Recording of processed signals (HA output) with artificial
  head (KEMAR) with ear simulator

- Equalization of recordings
  à similar frequency responses across HAs

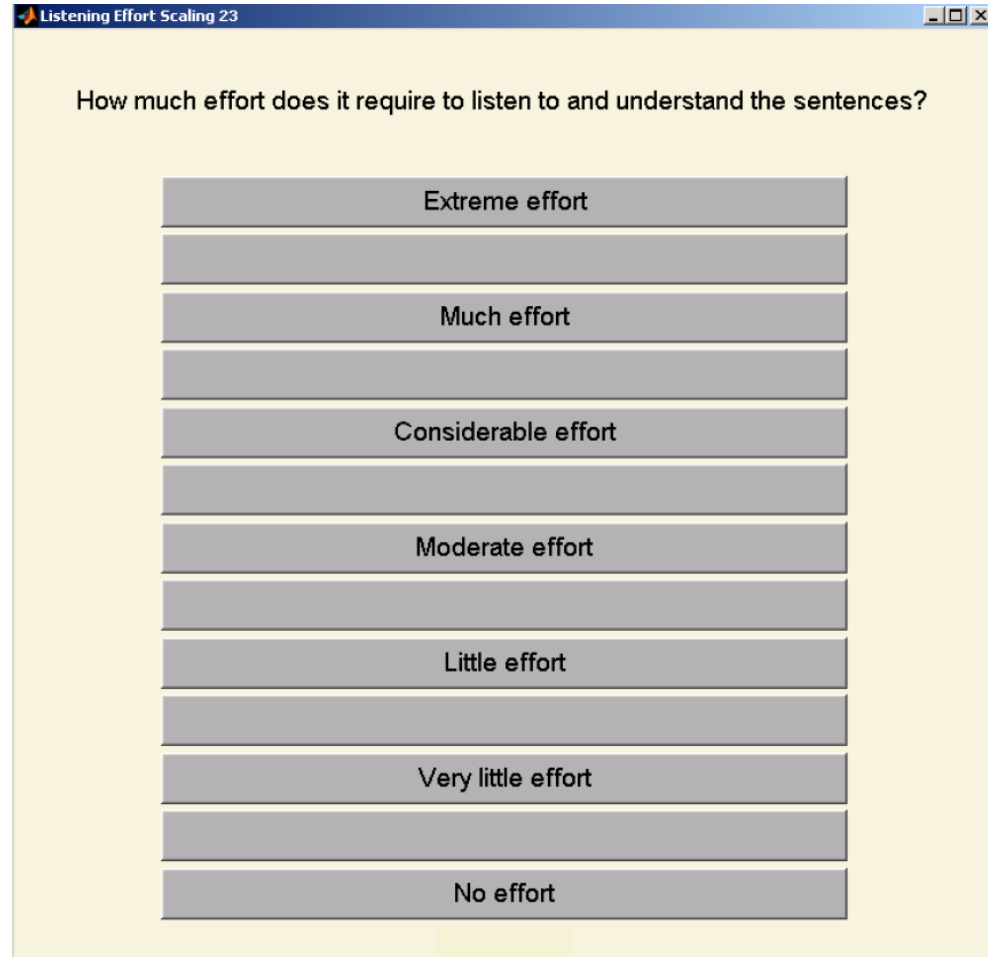- Presentation of test signals over headphones (HDA200)

Subjects

20 moderately hearing impaired subjects (age 26 – 83, median= 71); HA wearers; experienced test subjects

Averaged audiogram
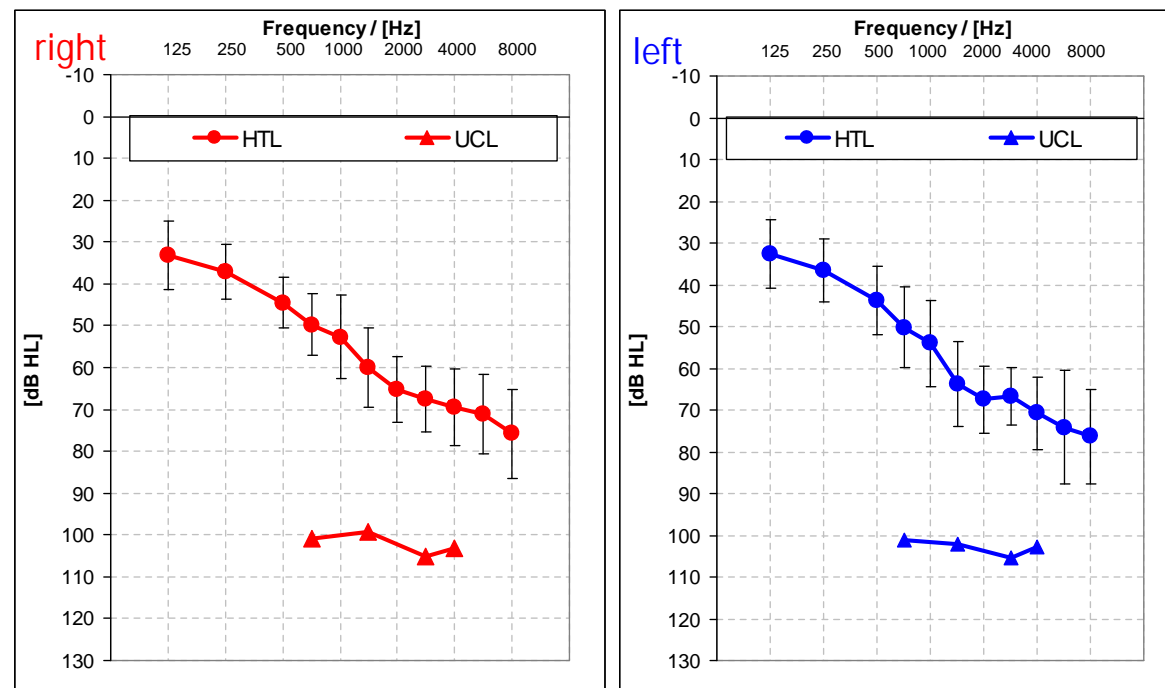
Method:
Listening effort scaling

- 7-category scale (13 steps)
- Subset of SNRs derived from pre-tests, covering range from „extreme effort" to „no effort":
  - ► Airplane cabin noise: SNR = -10, -7, -4, -1, 2, 5dB
  - ► Speech sim. noise: SNR = -1, 2, 5, 8, 11, 14dB

Listening Effort Scaling 23

How much effort does it require to listen to and understand the sentences?

Extreme effort

Much effort

Considerable effort

Moderate effort

Little effort

Very little effort

No effort

Objective quality measures

- ITU-T Rec. P.862 – PESQ (Beerends et al., 2002)

- Speech quality measure q_C (Hansen & Kollmeier, 2000)

- PEMO-Q (Huber & Kollmeier, 2006)

- Loudness Pattern Distortion (LPD) (Chen & Parsa, 2007)

Objective quality measures

- ITU-T Rec. P.862 – PESQ
- Speech quality measure q_C
- PEMO-Q
- Loudness Pattern Distortion (LPD)

Comparison of internal representations

- Weighted Spectral Slope Distance (WSSD) (Klatt, 1982)
- Log-Area Ratio (LAR) (Quackenbush et al., 1988)
- Log-Likelihood Ratio (LLR) (Itakura, 1975)
- Signal-to-Noise Ratio (SNR)

Objective quality measures

- ITU-T Rec. P.862 – PESQ

- Speech quality measure q_C

- PEMO-Q

- Loudness Pattern Distortion (LPD)

Comparison of internal representations

- Weighted Spectral Slope Distance (WSSD)

- Log-Area Ratio (LAR)

- Log-Likelihood Ratio (LLR)

Comparison of linear prediction coefficients

- Signal-to-Noise Ratio (SNR)

All measures are comparison-based, need reference signal;
Reference signal used here: speech + noise at 16 dB SNR

Results - overview

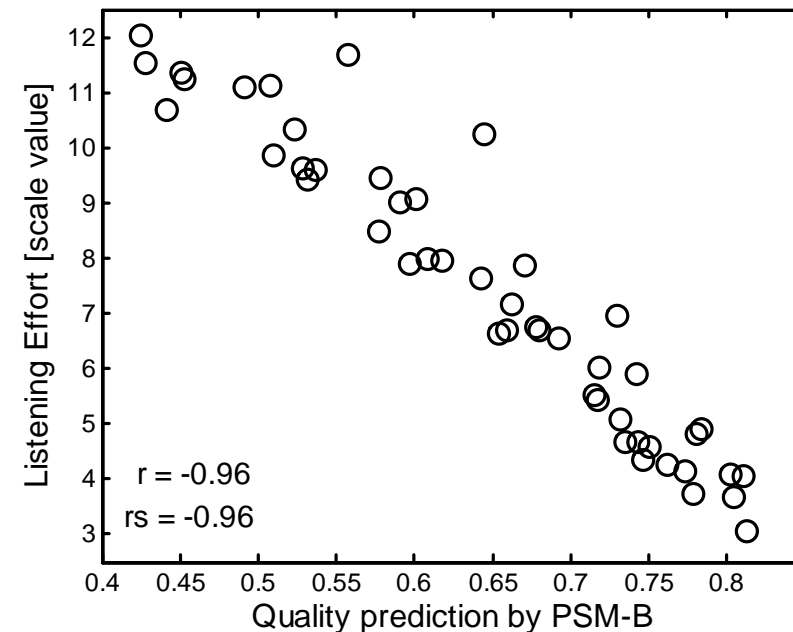## Correlations with subjective ratings

measure	Noise	
	speech sim.	airplane
qc-W	**0.96**	**0.93**
PSM-B	**0.96**	**0.91**
PSM	**0.95**	**0.90**
LPD	**0.95**	0.83
PESQ	0.85	**0.91**
qc	**0.92**	0.77
LAR	**0.91**	0.75
WSSD	0.77	0.86
SNR	**0.90**	0.38
LLR	0.72	0.23

PEMO-Q { PSM-B, PSM

Speech simulating noise

higher effort

better quality

r = -0.96
rs = -0.96
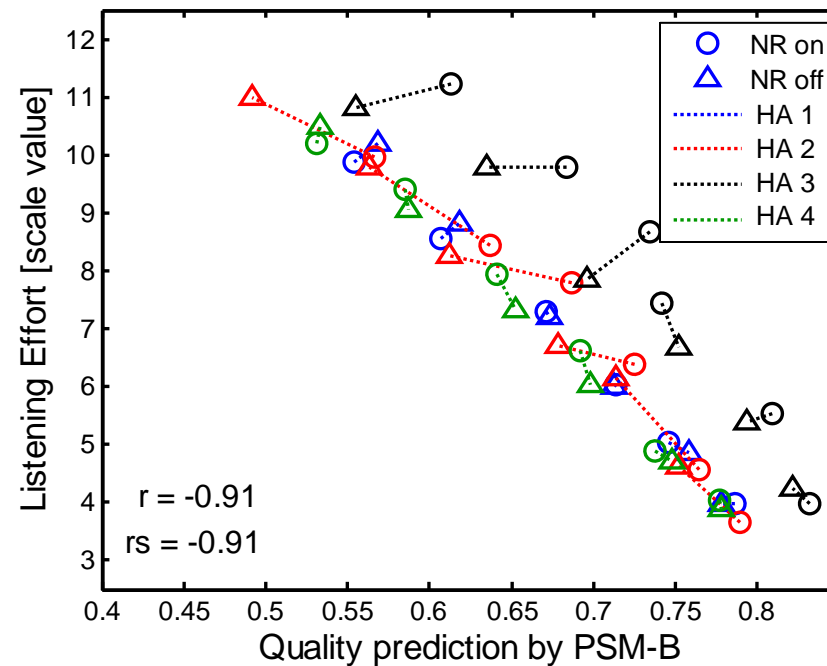
r: linear correlation
rs: rank correlation
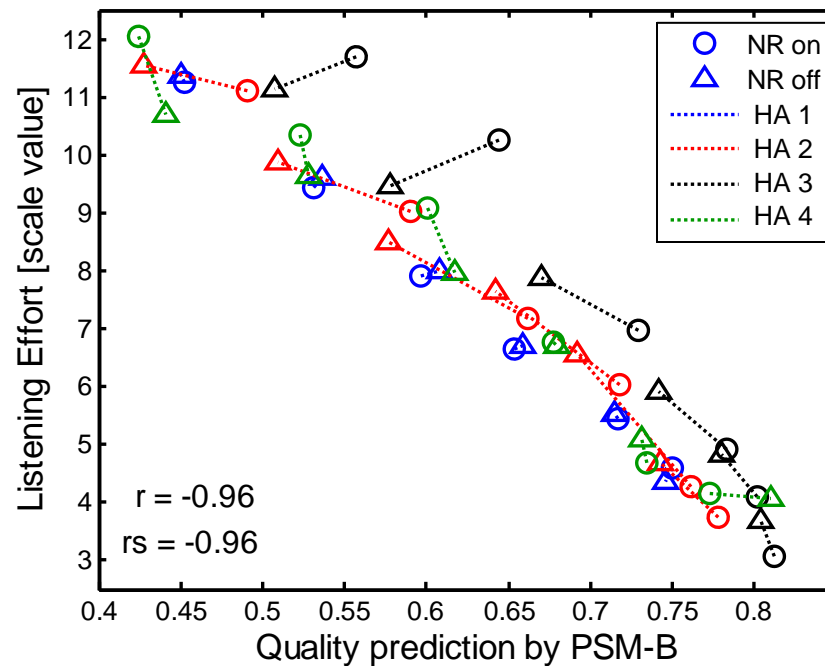
Speech simulating noise, SNR$_{in}$ = -1 ... 14 dB
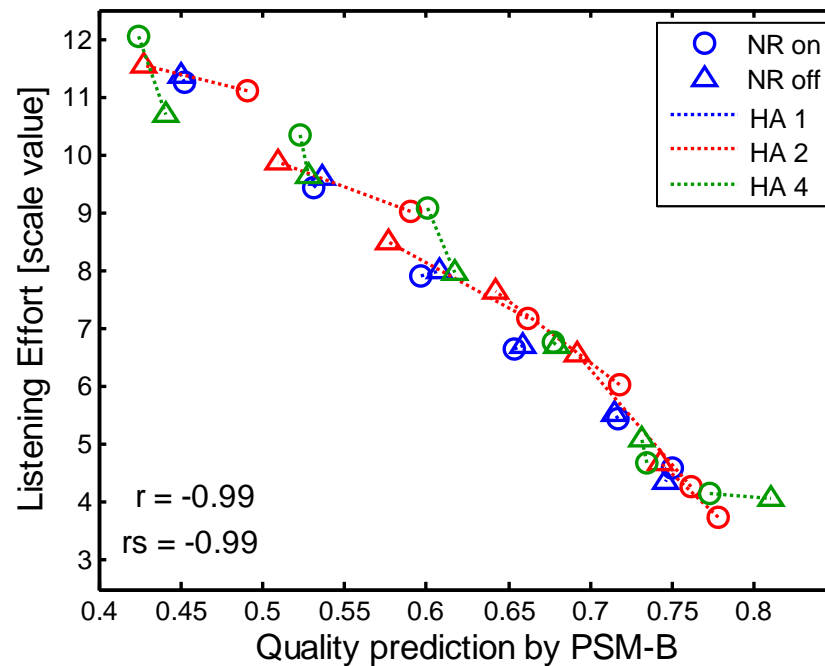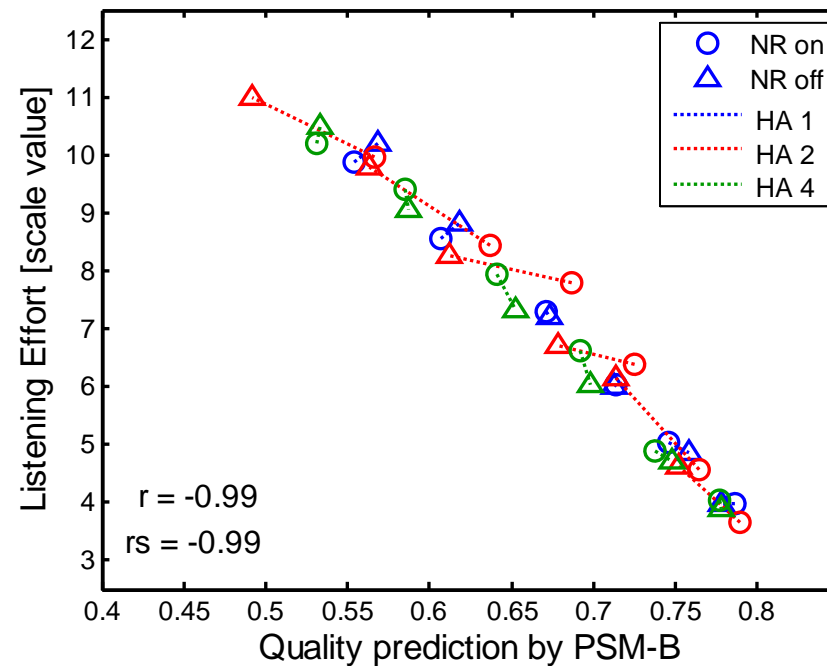
Airplane cabin noise, SNR$_{in}$ = -10 ... 5 dB

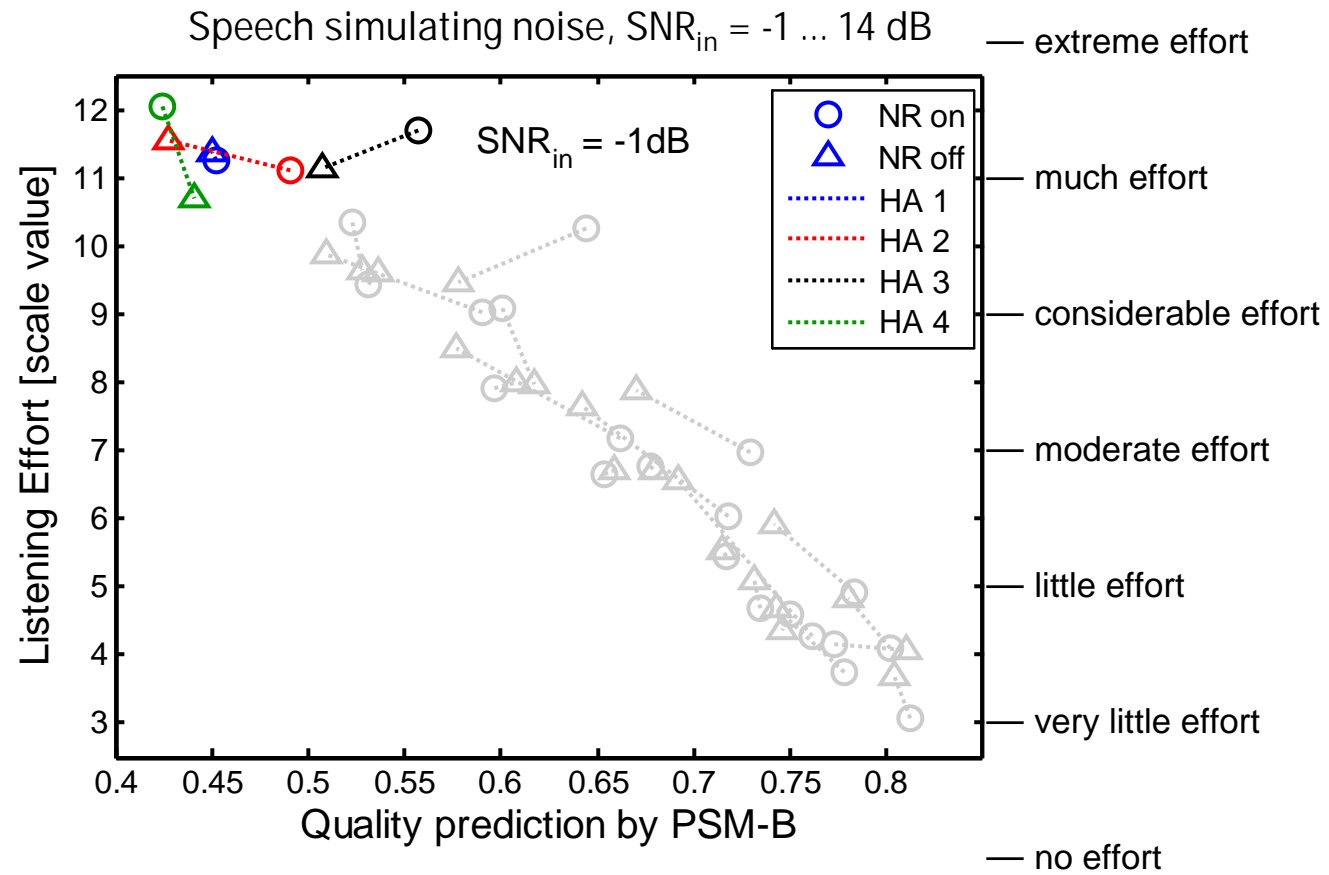Speech simulating noise, SNR$_{in}$ = -1 ... 14 dB

Airplane cabin noise, SNR$_{in}$ = -10 ... 5 dB
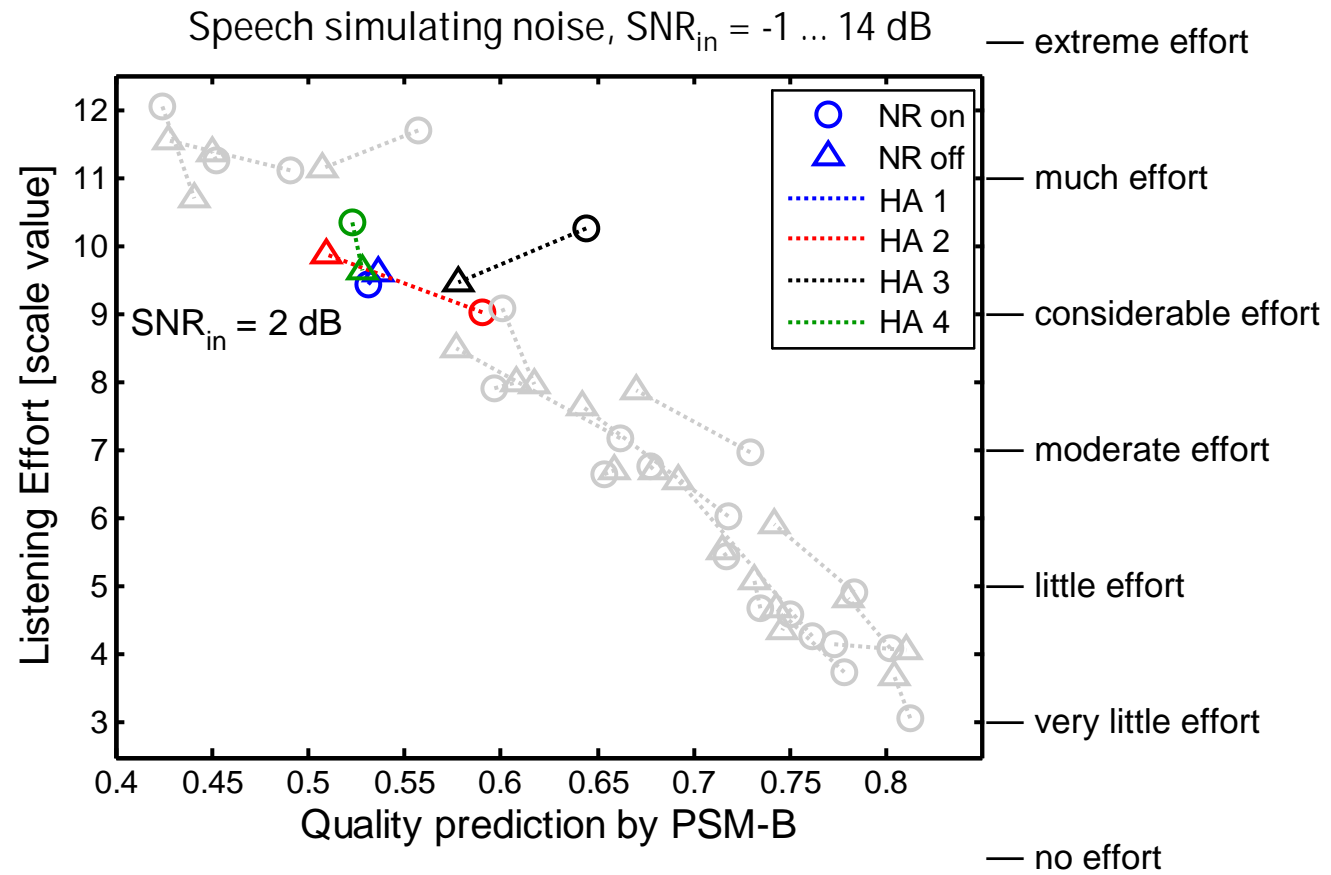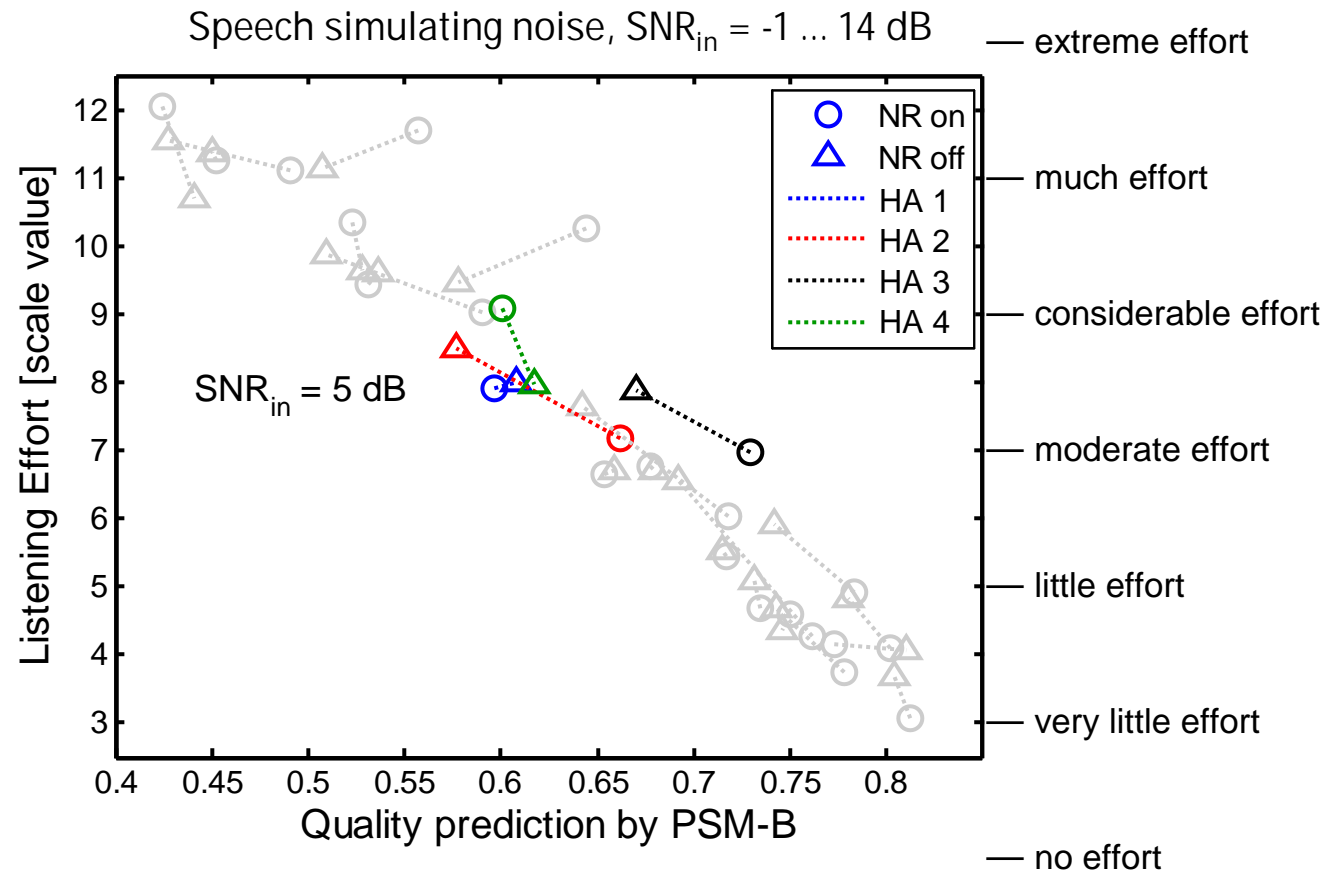
Speech simulating noise, $SNR_{in} = -1 \ldots 14$ dB

Speech simulating noise, SNR_{in} = -1 ... 14 dB

Speech simulating noise, SNR_{in} = -1 ... 14 dB

Influence of input SNR

Speech simulating noise, SNR_{in} = -1 ... 14 dB

— extreme effort

— much effort

— considerable effort

— moderate effort

— little effort

— very little effort

— no effort

Listening Effort [scale value]

Quality prediction by PSM-B

SNR_{in} = 8 dB

Legend:
- ○ NR on
- △ NR off
- .... HA 1
- .... HA 2
- .... HA 3
- .... HA 4

Speech simulating noise, SNR_{in} = -1 ... 14 dB

Speech simulating noise, SNR_{in} = -1 ... 14 dB

Speech simulating noise, SNR_{in} = -1 dB

Correlations per constant SNR$_{in}$

Speech simulating noise, SNR$_{in}$ = 2 dB
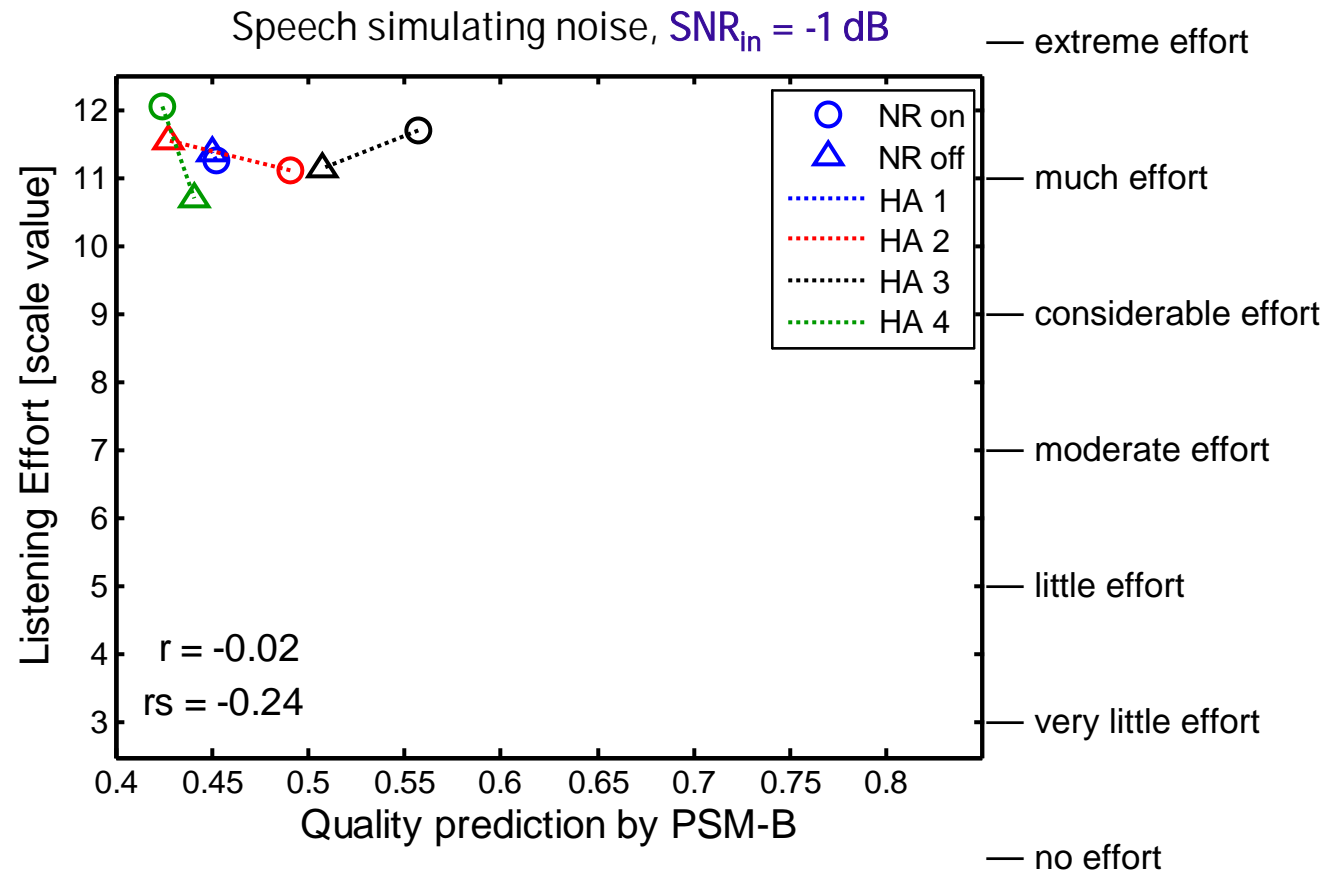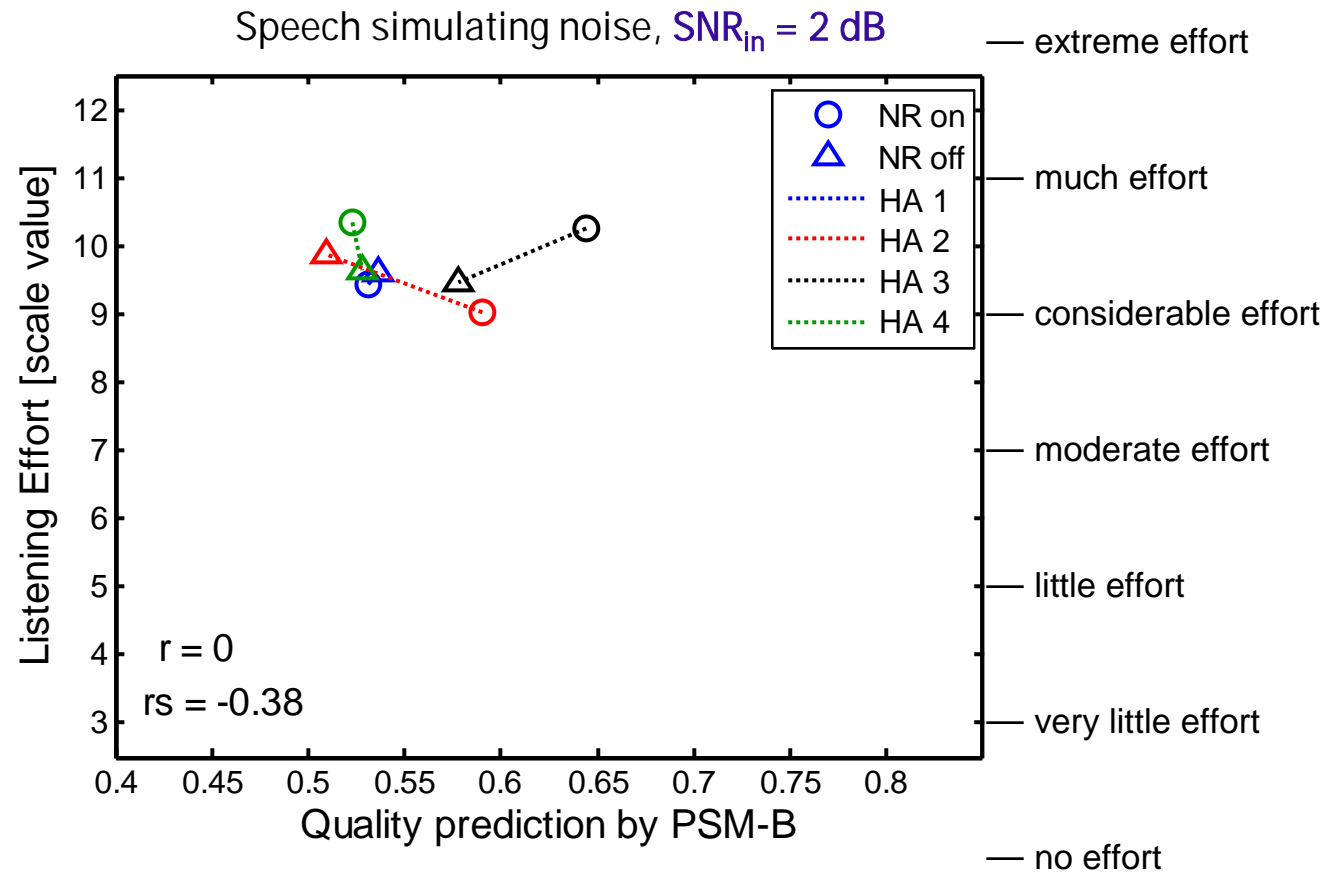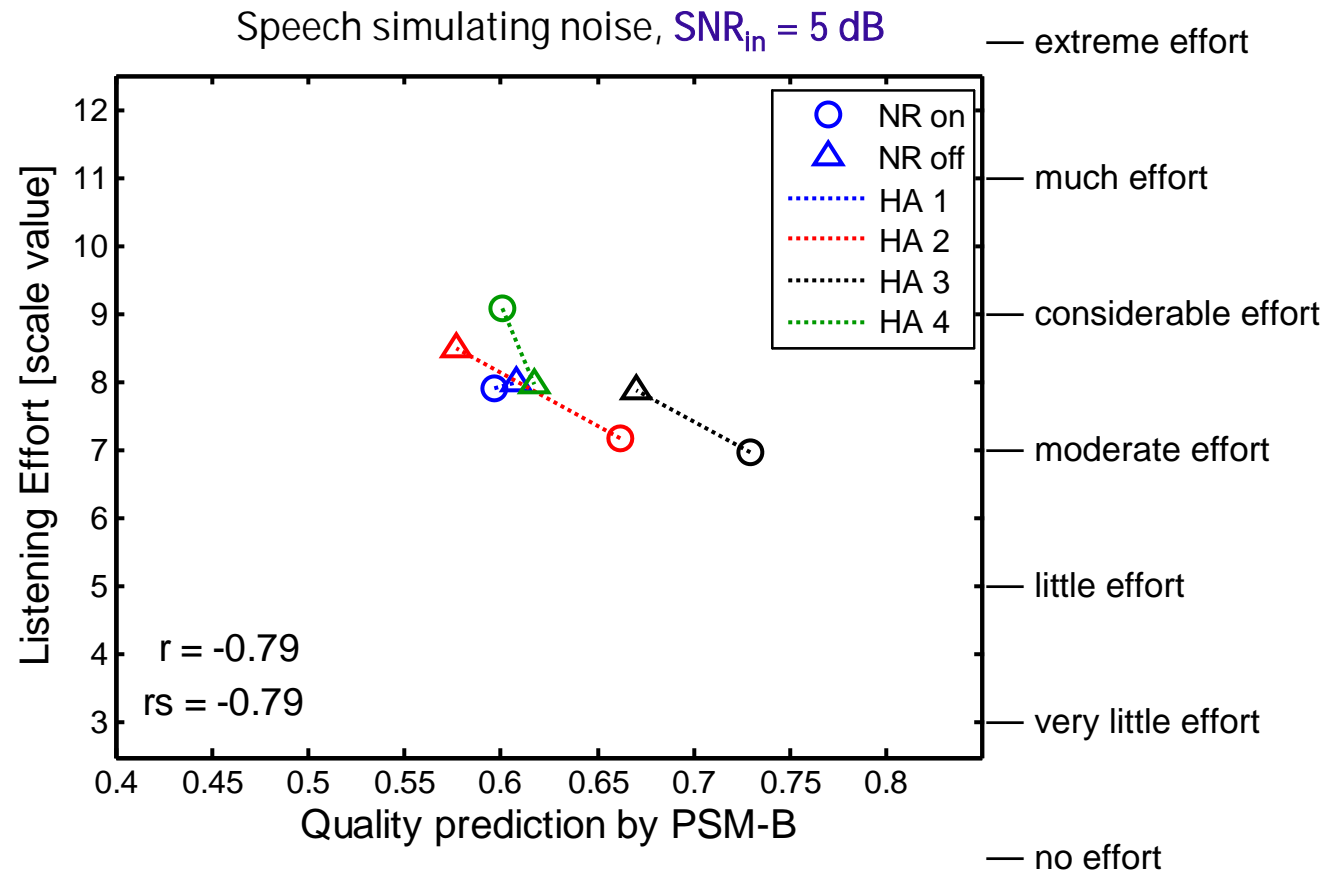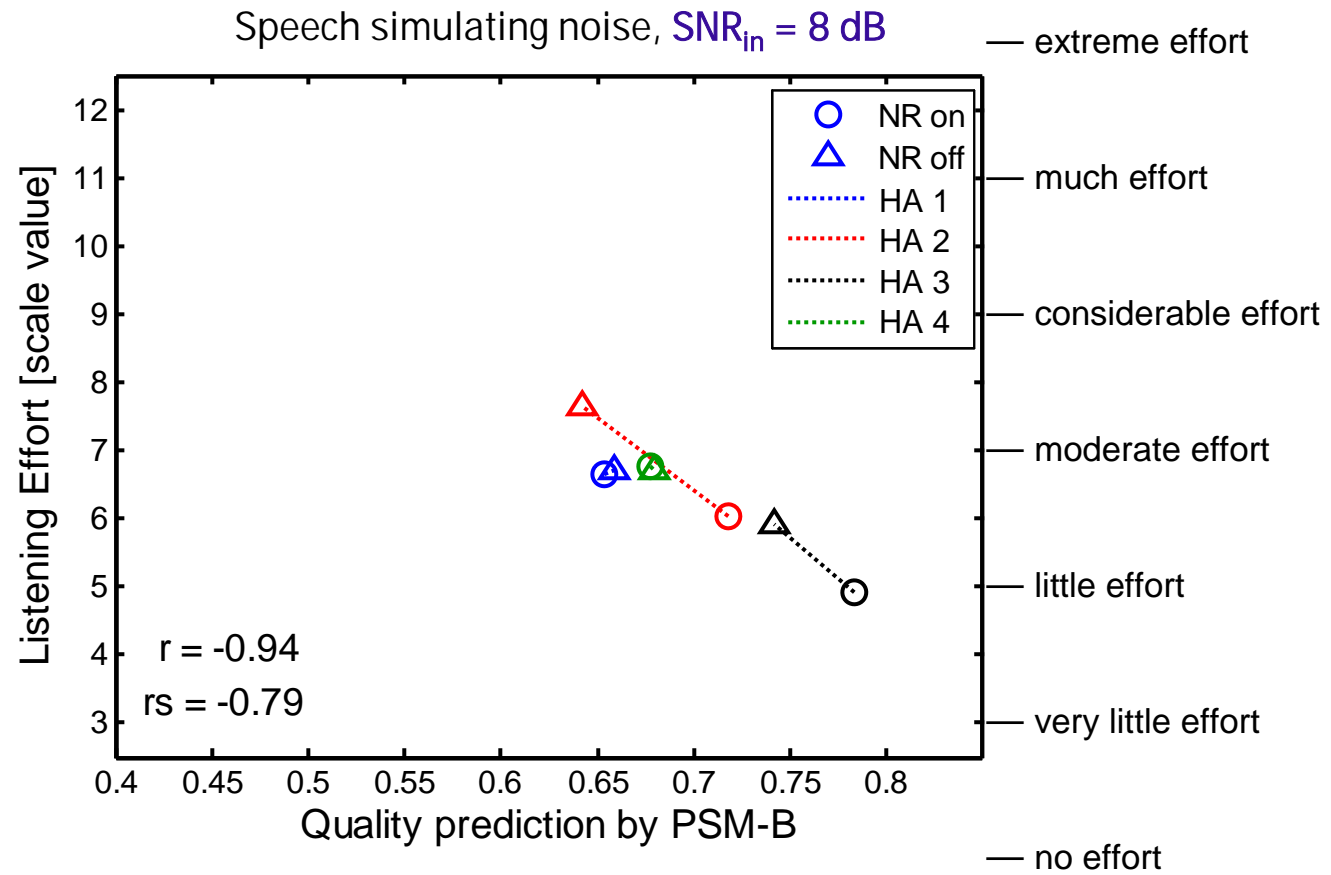
Correlations per constant SNR_{in}

Correlations per constant SNR$_{in}$

Speech simulating noise, SNR$_{in}$ = 8 dB

— extreme effort

Listening Effort [scale value]

Legend	
O	NR on
△	NR off
·······	HA 1
·······	HA 2
·······	HA 3
·······	HA 4

— much effort

— considerable effort

— moderate effort

— little effort

r = -0.94
rs = -0.79

— very little effort

Quality prediction by PSM-B

— no effort

Correlations per constant SNR$_{in}$

Speech simulating noise, SNR$_{in}$ = 11 dB

Listening Effort [scale value]

Quality prediction by PSM-B

— extreme effort
— much effort
— considerable effort
— moderate effort
— little effort
— very little effort
— no effort

- ○ NR on
- △ NR off
- .... HA 1
- .... HA 2
- .... HA 3
- .... HA 4

r = -0.87
rs = -0.93

Correlations per constant SNR_{in}

Speech simulating noise, SNR_{in} = 14 dB

— extreme effort

— much effort

— considerable effort

— moderate effort

— little effort

— very little effort

— no effort

Listening Effort [scale value]

Quality prediction by PSM-B

NR on
NR off
HA 1
HA 2
HA 3
HA 4

r = -0.84
rs = -0.9

Prediction of listening effort reduction

All SNRs

measure	Noise	
	speech sim.	airplane
qc-W	0.65	0.58
qc	0.69	0.54
PSM-B	0.62	0.56
PSM	0.57	0.61
PESQ	0.22	0.79
LPD	0.70	0.53
SNR	0.81	0.17
LAR	0.42	0.57
LLR	0.69	0.16
WSSD	0.06	0.07

PEMO-Q { qc-W, qc, PSM-B, PSM }

Speech simulating noise, SNR_{in} = -1 ... 14 dB

HA 1
HA 2
HA 3
HA 4

r = 0.62
rs = 0.66

Listening Effort reduction

Quality difference prediction ΔPSM-B
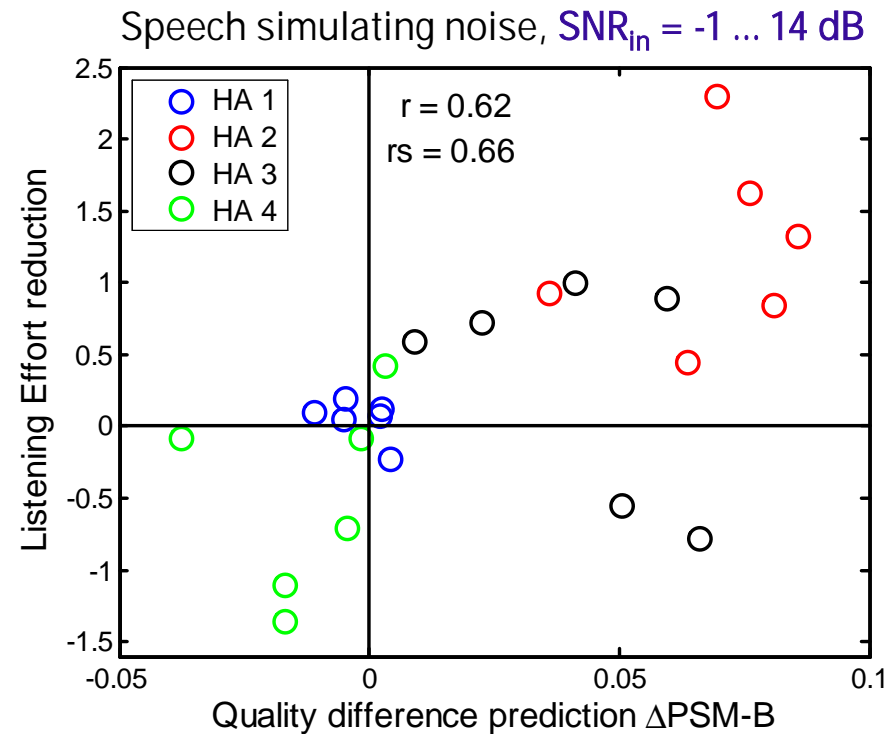
Prediction of listening effort reduction

2 SNRs with highest efforts omitted

measure	Noise	
	speech sim.	airplane
qc-W	0.85	0.67
qc	0.87	0.58
PSM-B	0.86	0.62
PSM	0.80	0.57
PESQ	0.46	0.82
LPD	0.69	0.48
SNR	0.87	0.21
LAR	0.43	0.56
LLR	0.49	0.41
WSSD	0.18	0.16

PEMO-Q { PSM-B, PSM }

Speech simulating noise, SNR_{in} = 5 ... 14 dB

r = 0.86
rs = 0.88

HA 1, HA 2, HA 3, HA 4

Listening Effort reduction

Quality difference prediction ΔPSM-B
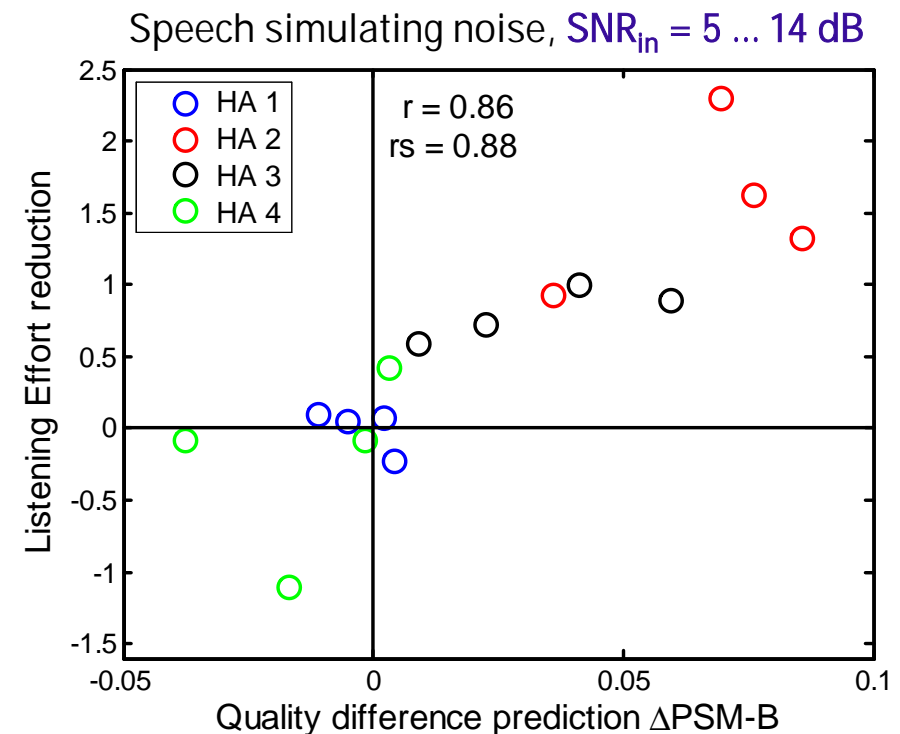
Objective HA benchmarking with speech in noise
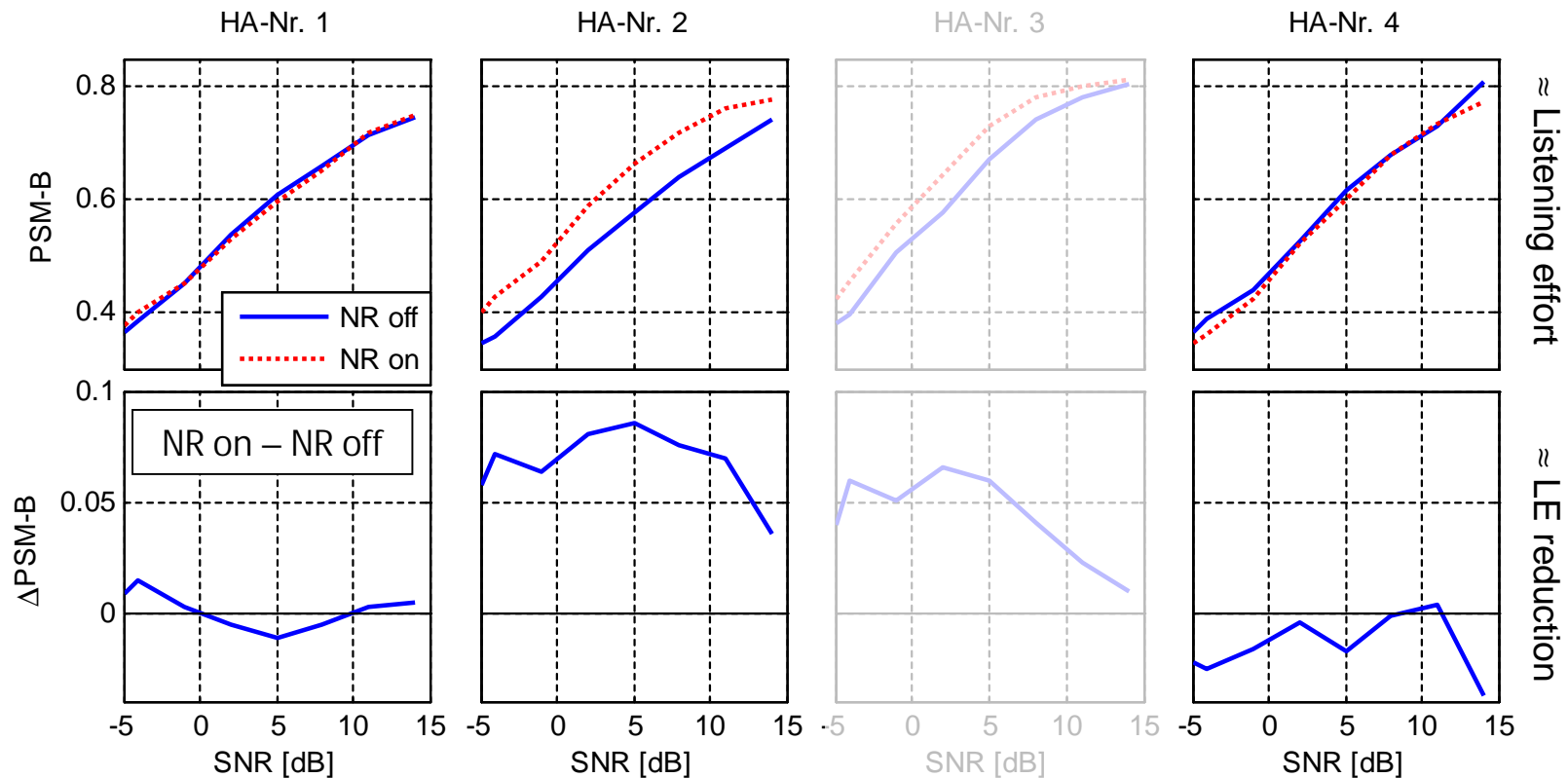
Noise: speech simulating noise

Summary and conclusion

- Benchmark test of noise reduction of commercial hearing aids; criterion: listening effort reduction

- Subjective tests (effort scaling) and application of objective quality measures

- Very high correlations between PEMO-Q quality measures and subjective ratings of (absolute) listening effort

- Prediction of (small) listening effort reductions more difficult; reasonable correlation for not-too-high listening efforts
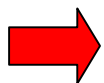
Measurement/prediction of listening effort for noisy speech appears qualified to benchmark hearing aids