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)
Objectives

Aim of the study:
Benchmarking of noise reduction programs of four commercial hearing aids (HAs)

- Subjective tests: intelligibility, listening effort, 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
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)
- 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)

Comparison of internal representations
Objective quality measures

- ITU-T Rec. P.862 – PESQ
- Speech quality measure $q_C$
- PEMO-Q
- Loudness Pattern Distortion (LPD)
- Weighted Spectral Slope Distance (WSSD)
- Log-Area Ratio (LAR)
- Log-Likelihood Ratio (LLR)
- Signal-to-Noise Ratio (SNR)

\[ \text{Comparison of internal representations} \]
\[ \text{Comparison of linear prediction coefficients} \]

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

<table>
<thead>
<tr>
<th>measure</th>
<th>Noise</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>speech sim.</td>
<td>airplane</td>
</tr>
<tr>
<td>qc-W</td>
<td>0.96</td>
<td>0.93</td>
</tr>
<tr>
<td>PSM-B</td>
<td>0.96</td>
<td>0.91</td>
</tr>
<tr>
<td>PSM</td>
<td>0.95</td>
<td>0.90</td>
</tr>
<tr>
<td>LPD</td>
<td>0.95</td>
<td>0.83</td>
</tr>
<tr>
<td>PESQ</td>
<td>0.85</td>
<td>0.91</td>
</tr>
<tr>
<td>qc</td>
<td>0.92</td>
<td>0.77</td>
</tr>
<tr>
<td>LAR</td>
<td>0.91</td>
<td>0.75</td>
</tr>
<tr>
<td>WSSD</td>
<td>0.77</td>
<td>0.86</td>
</tr>
<tr>
<td>SNR</td>
<td>0.90</td>
<td>0.38</td>
</tr>
<tr>
<td>LLR</td>
<td>0.72</td>
<td>0.23</td>
</tr>
</tbody>
</table>

r: linear correlation
rs: rank correlation

better quality

higher effort

r = -0.96
rs = -0.96

Quality prediction by PSM-B

Listening Effort [scale value]

Speech simulating noise
Results – closer look

Speech simulating noise, SNR\textsubscript{in} = -1 ... 14 dB

Airplane cabin noise, SNR\textsubscript{in} = -10 ... 5 dB

Quality prediction by PSM-B

Listening Effort [scale value]

- \( r = -0.96 \)
- \( rs = -0.96 \)

- \( r = -0.91 \)
- \( rs = -0.91 \)
Results – closer look

Speech simulating noise, $SNR_{in} = -1 \ldots 14$ dB

Airplane cabin noise, $SNR_{in} = -10 \ldots 5$ dB

Listening Effort [scale value]

Quality prediction by PSM-B

$\text{Listening Effort [scale value]}$

Quality prediction by PSM-B

$\text{Listening Effort [scale value]}$

$r = -0.99$

$rs = -0.99$
Influence of input SNR

Speech simulating noise, $\text{SNR}_{\text{in}} = -1 \ldots 14 \text{ dB}$

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

Listening Effort [scale value]

Quality prediction by PSM-B
Influence of input SNR

Speech simulating noise, $SNR_{in} = -1 \ldots 14$ dB

Listening Effort [scale value]

- extreme effort
- much effort
- considerable effort
- moderate effort
- little effort
- very little effort
- no effort
Influence of input SNR

Speech simulating noise, $\text{SNR}_{\text{in}} = -1 \ldots 14 \text{ dB}$

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

Listening Effort [scale value]

$\text{SNR}_{\text{in}} = 5 \text{ dB}$

Quality prediction by PSM-B
Influence of input SNR

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

Quality prediction by PSM-B

Listening Effort [scale value]

SNR$_{in}$ = 8 dB

- extreme effort
- much effort
- considerable effort
- moderate effort
- little effort
- very little effort
- no effort
Influence of input SNR

Speech simulating noise, $\text{SNR}_{\text{in}} = -1 \ldots 14 \text{ dB}$

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

Listening Effort [scale value] vs. Quality prediction by PSM-B

$\text{SNR}_{\text{in}} = 11 \text{ dB}$
Influence of input SNR

Speech simulating noise, $\text{SNR}_{\text{in}} = -1 \ldots 14 \text{ dB}$

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

Listening Effort [scale value]

Quality prediction by PSM-B
Correlations per constant $\text{SNR}_{\text{in}}$

Speech simulating noise, $\text{SNR}_{\text{in}} = -1 \text{ dB}$

Listening Effort [scale value]

Quality prediction by PSM-B

$r = -0.02$

$\text{rs} = -0.24$

- extreme effort
- much effort
- considerable effort
- moderate effort
- little effort
- very little effort
- no effort
Correlations per constant $\text{SNR}_{\text{in}}$

Speech simulating noise, $\text{SNR}_{\text{in}} = 2$ dB

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

$\rho = 0$
$rs = -0.38$
Speech simulating noise, $\text{SNR}_{\text{in}} = 5$ dB

- $r = -0.79$
- $rs = -0.79$

Listening Effort [scale value]

Quality prediction by PSM-B

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

Correlations per constant $\text{SNR}_{\text{in}}$
Correlations per constant SNR<sub>in</sub>

Speech simulating noise, SNR<sub>in</sub> = 8 dB

- Correlation: 
  - r = -0.94
  - rs = -0.79

- Listening Effort
  - NR on
  - NR off
  - HA 1
  - HA 2
  - HA 3
  - HA 4

- Quality prediction by PSM-B

- Effort Scale:
  - extreme effort
  - much effort
  - considerable effort
  - moderate effort
  - little effort
  - very little effort
  - no effort
Correlations per constant $SNR_{in}$

Speech simulating noise, $SNR_{in} = 11\, dB$

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

$r = -0.87$
$rs = -0.93$

Listening Effort [scale value]

Quality prediction by PSM-B

2nd Workshop on Speech in Noise - Amsterdam 8.1.2010
Correlations per constant SNR

Speech simulating noise, SNR\textsubscript{in} = 14 dB

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

\begin{align*}
r &= -0.84 \\
r_s &= -0.9
\end{align*}
Prediction of listening effort reduction

All SNRs

<table>
<thead>
<tr>
<th>measure</th>
<th>Noise</th>
<th>Speech sim.</th>
<th>airplane</th>
</tr>
</thead>
<tbody>
<tr>
<td>qc-W</td>
<td>0.65</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>qc</td>
<td>0.69</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>PSM-B</td>
<td>0.62</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td>PSM</td>
<td>0.57</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>PESQ</td>
<td>0.22</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>LPD</td>
<td>0.70</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>SNR</td>
<td>0.81</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>LAR</td>
<td>0.42</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>LLR</td>
<td>0.69</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>WSSD</td>
<td>0.06</td>
<td>0.07</td>
<td></td>
</tr>
</tbody>
</table>

Speech simulating noise, \( SNR_{in} = -1 \ldots 14 \) dB

Quality difference prediction \( \Delta PSM-B \)

PEMO-Q

\( r = 0.62 \)
\( rs = 0.66 \)
Prediction of listening effort reduction

2 SNRs with highest efforts omitted

<table>
<thead>
<tr>
<th>measure</th>
<th>Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>speech sim.</td>
</tr>
<tr>
<td>qc-W</td>
<td>0.85</td>
</tr>
<tr>
<td>qc</td>
<td>0.87</td>
</tr>
<tr>
<td>PSM-B</td>
<td>0.86</td>
</tr>
<tr>
<td>PSM</td>
<td>0.80</td>
</tr>
<tr>
<td>PESQ</td>
<td>0.46</td>
</tr>
<tr>
<td>LPD</td>
<td>0.69</td>
</tr>
<tr>
<td>SNR</td>
<td>0.87</td>
</tr>
<tr>
<td>LAR</td>
<td>0.43</td>
</tr>
<tr>
<td>LLR</td>
<td>0.49</td>
</tr>
<tr>
<td>WSSD</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Speech simulating noise, SNR_{in} = 5 \ldots 14 \text{ dB}

\[ r = 0.86 \quad \text{rs} = 0.88 \]
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