Perceptual evaluation of noise reduction in hearing aids
Brons, I.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Effects of noise reduction on speech intelligibility, listening effort, and personal preference in hearing-impaired listeners

Inge Brons, Rolph Houben, Wouter A. Dreschler

Submitted for publication
4.1 Introduction

Single-microphone noise reduction is a common feature in modern hearing aids that should determine whether the input signal is contaminated with noise and then adjust the hearing aids gain in specific frequency bands to suppress unwanted background noise. In Chapter 3, we directly compared noise reduction from different hearing aids perceptually with each other. For normal-hearing subjects, noise-reduction algorithms appeared to differ perceptually between hearing aids. In this follow-up study, we investigated whether these findings also hold true for hearing-impaired listeners:

Q1. Does hearing aid noise reduction influence speech intelligibility, listening effort, noise annoyance, speech naturalness, and preference for listeners with a moderate sensorineural hearing loss, compared with (a) no noise reduction and (b) noise reduction from other linearly fitted hearing aids?

In this phase, we evaluated noise reduction in isolation without the influence of dynamic-range compression.

4.2 Methods

The methods for hearing aid recording, perceptual measurements, and statistical analyses were identical to those described in Chapter 3.

4.2.1 Hearing aid recordings

We recorded hearing aid output of three linearly-fitted hearing aids from different brands (Phonak Exéïlia M, ReSound Azure AZ80-DVI, and Widex Mind 440) using the method described in Chapter 2. Acoustical analyses of the noise-reduction processing of these hearing aids are given in Chapter 3. Recordings of the three hearing aids with noise reduction activated were randomly coded as conditions NR1, NR2, and NR3. For one hearing aid we also recorded the output when noise reduction was inactive, resulting in an “unprocessed” condition, representing all hearing aids with noise reduction inactivated.

Stimuli consisted of Dutch sentences in babble noise, recorded with an input noise level of 65 dB(A). Stimuli were presented monaurally to the subjects with Sennheiser HDA200 headphones. The noise level was 65 dB(A) for all the stimuli in the unprocessed condition. Additional amplification was applied according to the linear NAL-RP prescription (Byrne et al. 1991) to compensate for listeners individual hearing loss.
4.2.2 Subjects
Twenty hearing-impaired subjects between 48 and 69 years of age (average = 61.3 years) participated in this study. The subjects audiograms were similar (i.e., no more than 10 dB difference at octave frequencies) to audiogram type N3 (moderate hearing loss with moderate slope) in the set of standard audiograms proposed by Bisgaard et al. (2010). All outcomes were measured at both subjects’ individual average SRT\textsubscript{50} and at a fixed SNR of +4 dB.

4.2.3 Intelligibility
Following the adaptive procedure described by Plomp and Mimpen (1979), we measured Speech Reception Thresholds in noise (SRT\textsubscript{50}). At the fixed SNR of +4 dB, we measured the percentage of words correctly repeated, similar to the procedure used in Chapter 3, but at higher SNR. Both measurements started with 13 training sentences followed by one list of 13 sentences per processing condition.

4.2.4 Listening effort
The subjects rated the listening effort on a nine-point rating scale that ranged from “no effort” to “extremely high effort”. The subjects gave ratings for the four processing conditions at the SRT\textsubscript{50} level and at -4, +4, and +10 dB SNR.

4.2.5 Paired-comparison rating
We used paired-comparison rating to measure noise annoyance, speech naturalness, and overall preference. For each combination of processing conditions, subjects indicated which was best on each of the three criteria. All six combinations of conditions were measured three times, both at individual SRT\textsubscript{50} level and at +4 dB SNR.

4.3 Results
4.3.1 Intelligibility
Figure 4.1 shows the group results for speech intelligibility. A repeated-measures ANOVA on the SRT\textsubscript{50} results (left panel) revealed no significant effect of processing condition (F[3,57] = 1.0, p = 0.39). The other outcomes were measured at the individually averaged SRT\textsubscript{50} rounded to whole decibels, ranging from -1 to +4 dB. A repeated measures ANOVA on the rau-transformed percentages (right panel) revealed no significant effect of processing condition (F[3,57] = 2.3, p = 0.085). However, pairwise comparisons with Bonferroni correction for six comparisons showed a significant difference between unprocessed and NR2 (uncorrected p = 0.005).
Noise reduction in linear hearing aids (hearing-impaired listeners)

Figure 4.1: Mean and 95% confidence interval of the SRT$_{50}$ (left panel) and of the percentage of words correctly repeated by the subjects at +4 dB SNR (right panel). “Unpr” is the unprocessed reference condition and NR1, NR2, NR3 are the hearing aid noise reductions. Horizontal bars indicate which processing conditions differ significantly from each other after Bonferroni correction for six comparisons.

4.3.2 Listening effort

Figure 4.2 shows the group-average listening-effort ratings relative to that for unprocessed at SRT$_{50}$ level and averaged over the three fixed levels (right panel) and the average absolute ratings for the three fixed SNRs separately (left panel). A repeated-measures ANOVA on the arcsine-transformed data on SRT$_{50}$ level showed a significant effect of processing condition (F[3,57] = 2.9, p = 0.043), but pairwise comparisons were not significant after Bonferroni correction. The data for the fixed SNRs showed significant effects of SNR (F[2,38] = 124.3, p < 0.001) and processing condition (F[3,57] = 2.8, p = 0.047), but not in Bonferroni-corrected pairwise comparisons of processing conditions.

Figure 4.2: Mean and 95% confidence interval of the listening effort ratings assigned by the 20 subjects relative to unprocessed (Δ listening effort, right panels) at SRT$_{50}$ level (upper right panel) and averaged over the three fixed SNRs (lower right panel), and absolute ratings at -4, +4, and +10 dB SNR (left panel).
4.3.3 Paired-comparison rating

Figure 4.3 shows the average rating scores for each processing condition for the three judgment criteria. Scores from -3 to 3 represent the seven categories in the paired-comparison scale. For statistical analysis, we modeled the data using a log-linear modeling approach for ordinal paired comparisons (Dittrich et al. 2004). A repeated-measures ANOVA on the results showed a significant effect of processing condition for all criteria except for speech naturalness at SRT$_{50}$ level. Horizontal lines in Figure 4.3 indicate which processing conditions differed significantly from each other after Bonferroni correction.

![Figure 4.3: Mean rating scores derived from the paired-comparison data for the three judgement criteria and two SNRs. Scores from -3 to +3 were assigned as 0, indicating no difference; -1 and +1 indicating a minor difference; -2 and +2 indicating a moderate difference; and -3 and +3 indicating a major difference. Error bars show the 95% confidence interval among subjects. Horizontal bars indicate which processing conditions differ significantly from each other (p < 0.05/6 = 0.0083; Bonferroni-corrected threshold for 6 comparisons).](image-url)
4.4 Discussion

4.4.1 Intelligibility
Word scores for NR2 were lower than those for unprocessed. In the results of Chapter 3, NR2 also had the lowest word score, although it was not significantly lower than unprocessed. Most studies found no effect of noise reduction on speech intelligibility (Nordrum et al. 2006). Results of Hu and Loizou (2007a) suggest that noise reduction reduces intelligibility more at lower SNRs. Our results do not support this; we measured no effect of noise reduction at lower SNR (SRT50) and a small negative effect at positive SNR. Hilkhuysen et al. (2012) found no interaction between noise reduction and SNR. However, they did not take measurements at positive SNRs.

4.4.2 Listening effort
We found no significant effect of noise reduction on listening effort. The ranking of conditions was equal for the SRT50 level and fixed SNR (see Figure 4.2). The same ranking was obtained previously for normal-hearing listeners (Chapter 3). Although not statistically significant, this finding suggests that the results for hearing-impaired listeners agree with those of normal-hearing subjects.

4.4.3 Noise annoyance, speech naturalness, and overall preference
Hearing-impaired listeners indicated differences in noise annoyance, speech naturalness, and overall preference between the conditions of noise reduction on and off and between noise-reduction algorithms of different linear hearing aids. Results at +4 dB SNR agreed well with previous results of normal-hearing subjects (Chapter 3), except for speech naturalness. Hearing-impaired subjects rated speech naturalness higher with noise reduction (NR1 and NR2), indicating that they might use other cues to rate naturalness, for instance the absence of noise (Marzinzik 2000).

We repeated the analysis with the subjects divided in two groups based on their SRT50 (12 subjects with SRT50 -1, 0, or 1 dB and eight subjects with SRT50 +2, +3, or +4 dB), and we found that subjects with a low SRT50 rated naturalness at SRT50 level lower after noise reduction than subjects with a high SRT50. This finding confirms that noise reduction affects speech naturalness more at lower SNRs, as was previously found for normal-hearing subjects.

The condition that was most preferred by the subjects (NR2) also caused the lowest intelligibility scores. This trade-off between quality and intelligibility is inherent to noise reduction (Wang 2008; Chapter 6).
4.4.4 Signal-to-noise ratios for evaluation of noise reduction

Noise-reduction processing depends on the input SNR (Hoetink et al. 2009). We therefore measured not only at an individual SNR for each subject (to ensure an equal performance level for all subjects) but also at a fixed SNR to ensure equal noise-reduction processing over all subjects. Group results were similar between the two independent datasets obtained at SRT$_{50}$ level and at fixed SNR. This implies that, for the (small) range of hearing losses included, the approach of a fixed and individual SNR did not influence the results. This conclusion might not hold for a broader range of hearing losses. In that case, the approach of evaluating both from a listeners perspective (individually adjusted SNR) and a processing perspective (fixed SNR) might be considered, although results from fixed SNRs are easier to interpret because the effects of noise reduction and hearing ability are easier to separate.

4.4.5 Limitations

The results of this study were measured in a laboratory setting and cannot be generalized beyond the limited number of conditions included. An important factor not included in this study is the fact that most hearing aids apply dynamic-range compression, which might interact with the effects of noise reduction (Chung 2007; Anderson et al. 2009).

4.5 Conclusions

Noise reduction from all three hearing aids tested was able to reduce the annoyance of babble noise perceived by listeners with moderate sensorineural hearing loss. The noise reduction that reduced noise annoyance the most and that was most preferred caused poorer intelligibility scores, confirming a trade-off between listening comfort and intelligibility. The results of hearing-impaired subjects agree well with those of normal-hearing listeners in a previous study. Subsequent experiments should reveal how dynamic-range compression influences the effects of noise reduction.