Using response times to speech-in-noise to measure the influence of noise reduction on listening effort

ILJA REINTEN[1,*] INGE DE RONDE-BRONS[2], MAJ VAN DEN TILLAART-HAVERKATE[2], ROLPH HOUBEN[2] AND WOUTER DRESCHLER[1]

[1] *Clinical & Experimental Audiology, Amsterdam University Medical Centres location AMC, Meibergdreef 9 1105 AZ Amsterdam, the Netherlands*

[2] *Pento Audiological Centre, Zangvogelweg 150 3815 DP Amersfoort, the Netherlands*

Single microphone noise reduction (NR) can lead to a subjective benefit even when there is no objective improvement in speech intelligibility. A possible explanation lies in a reduction of listening effort. In a previous study, we showed that response times (a proxy for listening effort) to a simple arithmetic task with spoken digits in noise were reduced (i.e., improved) by NR for normal-hearing (NH) listeners. In the current study we complemented the data set with data from twelve hearing-impaired (HI) listeners, the target group for NR. Subjects were asked to add the first and third digit of a digit triplet in noise. Response times to this task were measured, subjective listening effort was rated, and speech intelligibility of the stimuli was tested. Stimuli were presented at three signal-to-noise ratios (SNR; -5, 0, +5 dB) and in quiet. Stimuli were either processed with ideal or non-ideal NR, or unprocessed. In contrast to the previous results with NH listeners, a significant effect of NR on response times was for HI listeners restricted to conditions where speech intelligibility was also affected (-5 dB SNR). We cannot confirm a positive effect on response times to speech-in-noise after applying NR for HI listeners.

INTRODUCTION

It is well known that single microphone noise reduction (NR) in hearing aids can lead to a subjective benefit, in terms of listener preference, even when there is no objective improvement in speech intelligibility (Brons *et al.*, 2014). This suggests that in addition to speech intelligibility there are other factors that determine listener preference for NR, such as a reduction of listening effort. The term listening effort, which is a reflection of the amount of cognitive resources that is required for adequate speech understanding (Hicks and Tharpe, 2002), received increasingly more attention in audiological research over the past few decades. It is closely related to fatigue, and therefore regarded as a fairly subjective measure. In spite of its subjective nature, there has been an ongoing effort to find an objective measure that adequately describes listening effort. Such a measure could be of additional value to describe the non-auditory effects of hearing disabilities and of hearing rehabilitation.

*Corresponding author: i.reinten@amsterdamumc.nl

Proceedings of the International Symposium on Auditory and Audiological Research (Proc. ISAAR), Vol. 7: Auditory Learning in Biological and Artificial Systems, August 2019, Nyborg, Denmark. Edited by A. Kressner, J. Regev, J. C.-Dalsgaard, L. Tranebjærg, S. Santurette, and T. Dau. The Danavox Jubilee Foundation, 2019. ISSN: 2596-5522.

Objective measures for listening effort that are described in literature include physiological values such as the pupil dilation response, heart rate variability, and EEG recordings. Another approach is the use of response times in a dual-task paradigm to measure listening effort. A primary listening task is complemented with a secondary task that requires cognitive processing. It is believed that the additional cognitive processing required for the secondary task is slowed down if the primary listening situation is more effortful (Hicks and Tharpe, 2002). The nature of the secondary task can be non-auditory, for instance response times to a visual cue (Sarampalis *et al.*, 2009; Desjardins and Doherty, 2014). Sarampalis *et al.* (2009) tested response times to a visual cue when listening to speech in noise at different signal-to-noise ratios (SNR). They found that at -6 dB SNR, normal-hearing (NH) listeners responded faster when the stimuli were processed with a NR algorithm based on a minimum mean square estimator (MMSE; Epraihm and Malah, 1984). Desjardins and Doherty (2014) tested performance in a secondary visual tracking task in moderate and difficult listening situations. The authors found that in difficult listening situations, hearing-impaired (HI) listeners performed better when NR from a commercially available HA was applied. The secondary task can also be based on the primary auditory-only task where extra processing is required. An auditory-only set-up requires less equipment and is therefore better suited for clinical applications. Additionally, test-results are not influenced by a possible non-auditory sensory impairment of the listener. In an experiment by Houben *et al.* (2013), such an auditory-only dual-task was performed by presenting digit triplets in noise. Participants had to identify the digits as the primary task and add the first and third digit as the secondary task. The authors showed that for NH listeners, response times to this simple arithmetic task reduced with increasing SNR when speech intelligibility was at its maximum. As a follow-up of this work, van den Tillaart-Haverkate *et al.* (2017) measured response times at different SNRs and for different forms of noise reduction processing. They found for a group of 12 NH listeners that noise reduction also caused a reduction (i.e., improvement) in response time.

Although the study of van den Tillaart-Haverkate *et al.* (2017) shows promising results for untangling the possible advantage of applying NR in hearing aids, the study lacks data of HI listeners. HI listeners are the target users of hearing aids, and they can have significantly different opinions regarding sound quality of hearing aid features. HI listeners should therefore be included in experiments that evaluate features such as noise reduction. Therefore, in this study 12 HI listeners participated in a similar experiment. Presently, we are interested whether response times to speech-in-noise are reduced for HI listeners after the application of NR.

METHODS

This study was approved by the Medical Ethics Committee of the Amsterdam UMC (former AMC) in 2013 (MEC2013_082).

Participants

Twelve HI listeners participated in this experiment. HI listeners had a mean age of 60 ± 5.3 years with a mild- to moderate sensorineural sloping hearing loss. The participants were recruited in the Audiological Centre of the Amsterdam UMC, location AMC. The group averaged audiogram of the included participants is shown in Fig. 1. All participants were native Dutch speakers.

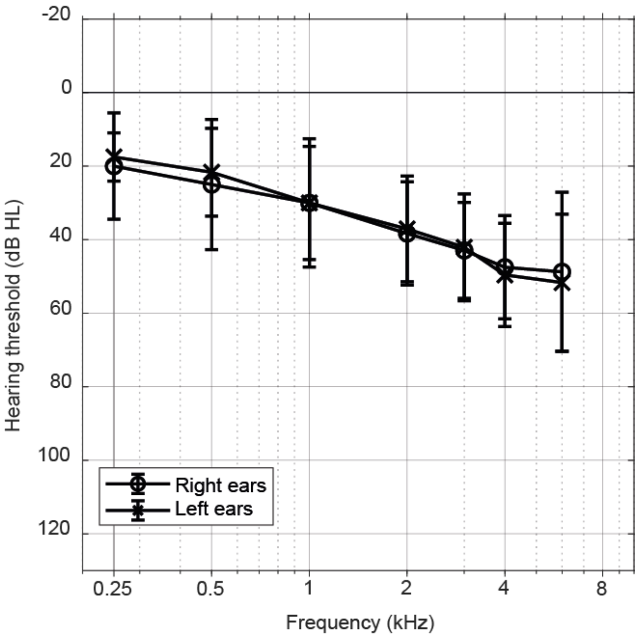

Fig. 1: Group averaged audiograms of the HI subjects with inter-individual standard deviations.

Stimuli and processing

Sixty spoken digit triplets in speech shaped noise were used at four different SNRs: -5, 0 +5 and +∞ (quiet) dB. At each SNR we processed the digit triplets to create three conditions per SNR: one unprocessed condition and two processed conditions with two types of NR algorithms. The two NR algorithms applied in the experiment were the ideal binary mask (IBM; Wang, 2005) and a MMSE (Ephraim and Malah, 1984). The most prominent difference between these two algorithms is that the IBM has a-priori knowledge of the noise and speech material as separate signals and is known to be able to improve speech intelligibility (Wang *et al.*, 2009). Since the IBM requires a-priori knowledge of the noise and speech signals (and the actual SNR), it is not suitable for implementation in real HAs, but it does give insight in the maximum achievable benefit offered by NR. The MMSE on the other hand has to estimate the SNR and is therefore comparable to NR algorithms that are currently implemented in HAs. Implementation of the algorithms was done in MATLAB and is described in detail in van den Tillaart-Haverkate *et al.* (2017), where a detailed description of the stimuli and equipment used can be found as well. All stimuli were presented diotically through headphones. The average level of the speech was 65 dB(A) with an additional linear amplification for each listener according to the NAL-RP rule (Byrne *et al.*, 2001).

Test procedure and data analysis

The primary outcome measure of this experiment was the response time to an arithmetic task (AR-task). For this task, all participants were presented with digit-triplets in noise and were asked to add the first and third digit. Instructions were given to answer as fast as possible on a numerical keypad. Absolute response times were defined as the time between the end of playing the last digit and the subsequent response key-press. Secondary outcome measures were speech intelligibility (SI) and perceived listening effort rating (LEr). These were tested per condition in the following way: first the participant was asked to correctly identify 20 triplets after which they were asked to rate their perceived listening effort. Listening effort rating was scored on a 9-point scale ranging from ‘no effort’ (1) to ‘extremely high effort’ (9) as an answer to the question: “How much effort did it take to understand the last 20 triplets?”

The experiment took place in two visits. The first visit started by measuring hearing thresholds with pure tone audiometry. The AR-task and SI/LEr task were performed in both visits in order to obtain more data points and to allow to investigate the accuracy of the measurement results.

For the AR-task, only correct responses were included in the analysis. For each task, the highest 1.25% of the response times was removed to ensure that unrealistically long response times were not included (Houben *et al.*, 2013). Since absolute response times can have a large inter-individual variation, data analysis was done by using a relative response time. Relative response times were defined by subtracting the response time at $+\infty$ dB SNR from the response time at the other SNRs per processing condition, for each participant.

RESULTS

Fig. 2-A shows the group average absolute response times of the AR-task for all conditions for HI listeners as well as the previously published data (van den Tillaart-Haverkate *et al.*, 2017) of NH listeners. Fig. 2-B shows the mean relative response times of all conditions for HI listeners.

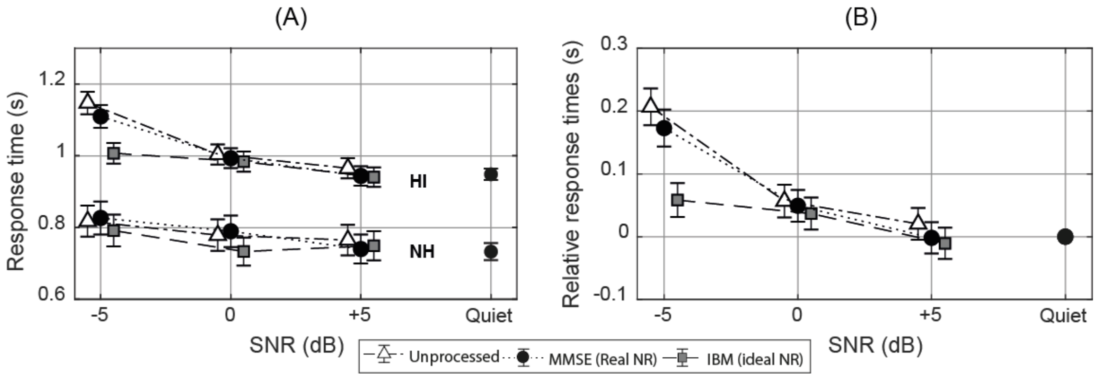

Fig. 2: (**A**) Absolute group-average response times of NH (van den Tillaart-Haverkate *et al.*, 2017) and HI listeners. (**B**) Relative group-average response times of HI listeners. The error bars show 95% confidence intervals.

We analysed the relative response times of the AR-task with a mixed-model ANOVA, with subject and triplet as random effects. Processing condition, SNR and the interaction between processing condition and SNR were considered fixed effects. We found significant effects of processing condition (F = 5.27, p = 0.0184), SNR (F = 107.18, $p < 0.001$), and the interaction between processing condition and SNR (F= 9, $p < 0.001$). Post-hoc pairwise comparisons after Bonferroni corrections (with α = 0.05/27) of the conditions revealed that at -5 dB SNR the IBM condition differed significantly from unprocessed and MMSE. Within the unprocessed and MMSE conditions response times at -5 dB SNR were significantly longer than at all other SNRs and in quiet ($p < 0.001$), and response times at 0 dB SNR were significantly longer than in quiet ($p < 0.001$). Within the IBM condition response times at -5 dB SNR were significantly longer than at +5 dB SNR and in quiet ($p < 0.001$).

Fig. 3-A shows the group average results of the speech intelligibility test in terms of % correct identification of triplets. We analysed the speech intelligibility with a mixed model ANOVA on the rationalized arcsine unit-transformed intelligibility scores, with subject and triplet as random effects. Processing condition, SNR and the interaction between processing condition and SNR were considered fixed effects. We found significant effects of processing condition (F = 12.33, $p < 0.001$), SNR (F = 35.82, $p < 0.001$) and the interaction between processing condition and SNR (F = 5.08, p = 0.001). Post-hoc pairwise comparisons after Bonferroni corrections (with α = 0.05/27) of the conditions revealed that at -5 dB SNR the IBM condition was significantly better intelligible than unprocessed (p = 0.001). Within the unprocessed and MMSE conditions, response times at -5 dB SNR were significantly longer than all other SNRs ($p < 0.001$). Within the IBM condition response times did not differ significantly at all SNRs.

Fig. 3-B shows the group average results of the perceived listening effort rating. We analysed the perceived listening effort rating with an ANOVA with subject as random effect. Processing condition, SNR and the interaction between processing condition

and SNR were considered fixed effects. We found a significant effect of processing condition ($F = 19.3$, $p < 0.001$), SNR ($F = 128.35$, $p < 0.001$), and the interaction between processing condition and SNR ($F = 7.59$, $p < 0.001$). Post-hoc pairwise comparisons after Bonferroni corrections (with $\alpha = 0.05/27$) of the conditions revealed that at -5 dB SNR, the IBM condition differed significantly from the unprocessed and MMSE conditions. Within the unprocessed and MMSE conditions, LEr significantly increased with decreasing SNR ($p < 0.001$), except between +5 dB SNR and 0 dB SNR. Within the IBM condition, LEr at -5 dB SNR was significantly higher than all other SNRs ($p < 0.001$), and LEr at 0 dB SNR was significantly higher than in quiet ($p < 0.001$).

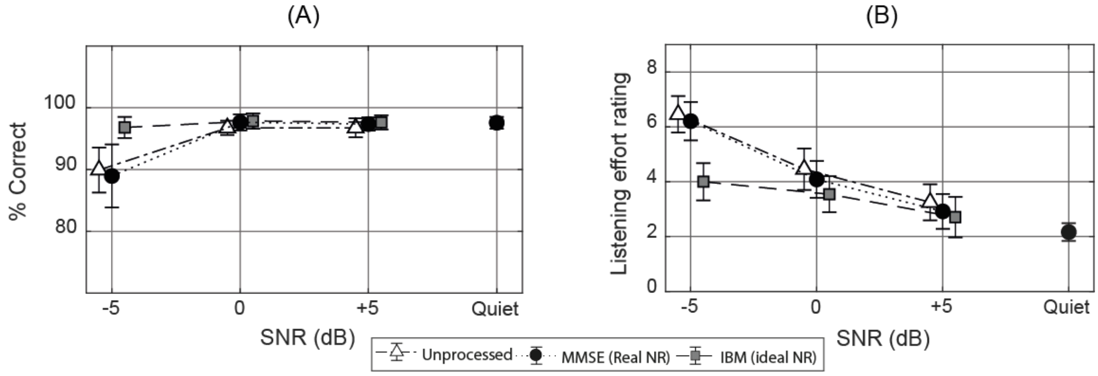

Fig. 3: (**A**) Group averaged speech intelligibility in terms of % correct responses for all SNRs and processing conditions, the error bars show 95% confidence intervals. (**B**) Group averaged perceived listening effort rating for all SNRs and processing conditions, the error bars show 95% confidence intervals.

DISCUSSION

We measured response time of HI listeners on a digit triplet test at different SNRs and for various NR conditions. The results show similarities as well as differences compared with the response time of NH listeners (obtained in a previous study; van den Tillaart-Haverkate *et al.*, 2017). The most obvious difference is that the group of HI listeners needs more time to respond than the previous group of NH listeners. This effect is most likely dominated by the age differences between the two groups. The mean age of the NH listener group (24 ± 4.2 years; van den Tillaart-Haverkate *et al.*, 2017) was on average 36 years younger than the HI listener group of the present study. Response times to tasks are known to increase considerably with age (Verhaeghen and Cerella, 2002; Melzer and Oddsson, 2004). However, we assume that this age effect is negligible in the *relative* response time. Verhaeghen and Cerella (2002) report that reaction times in older adults can commonly be described as a linear transformation to those of younger adults. In the current results we found a significant reduction of relative response times with increasing SNRs. Within the unprocessed and MMSE condition this reduction was significant even when speech intelligibility was maximal. When comparing Fig. 2-B with Fig. 3-B the same trend is found for response times and perceived listening effort: an increase in SNR is accompanied by a decrease in relative response times or perceived listening effort. These observations, confirmed by statistical analyses, are in agreement with our previous studies and support the hypothesis that response times as such might be used as an objective measure for listening effort (Houben *et al.*, 2013; van den Tillaart-Haverkate *et al.*, 2017).

The main purpose of the current experiment was to test whether response times in a dual-task paradigm are reduced by applying NR for HI listeners, the target group for using NR. Fig. 2-B shows that at all SNRs, the relative response times for the processed conditions are consequently shorter than unprocessed signals. This suggests a positive effect of NR on response times. However, given the significant interaction found between processing condition and SNR, we cannot directly interpret the overall effect of NR on relative response times but instead we have to analyze each SNR separately. This analysis has only revealed a significant reduction of response time between IBM and the other two conditions at -5 dB SNR. This reduction is most likely caused by the large improvement of speech intelligibility for IBM at -5 dB SNR. The decrease of response times with an increase of speech intelligibility is an effect that has been observed before (Gatehouse and Gordon, 1990; Baer *et al.*, 1993), but we are most interested in an effect of NR on response times at SNRs where speech intelligibility is maximal. In this so-called area of interest the reductions in response times by applying NR were not significant, which is in contrast with our previous results in NH listeners (van den Tillaart-Haverkate *et al.*, 2017). A possible explanation for these contrasting results might lie in the prominent reduction in relative response times for IBM at -5 dB SNR, causing the statistical analysis to lose power. The effect of hearing loss also needs to be considered in the test performance. It is well-known that sound perception is different for HI listeners, but little is known how this can affect cognitive processes as measured in dual-task paradigms. Our results are consistent with findings from Sarampalis *et al.* (2009) who also report a reduction in listening effort by applying a similar NR algorithm at a difficult listening situation (-6 dB SNR). Their study did not include HI listeners. Desjardins and Doherty (2014), who did include HI listeners, also measured a reduction of listening effort by NR at a more complex listening condition. Both studies used a visual dual-task paradigm, whereas we used an audiological-only dual task. Our auditory-only dual-task gave similar results to the mentioned visual dual-task paradigms. This finding suggests that an auditory secondary task is suitable for evaluating listening effort. However, the issue remains that the beneficial effects of NR in scenarios where speech intelligibility is maximal may be hard to interpret.

In conclusion, the current dataset of response times to a dual-task paradigm for HI listeners shows a significant and positive effect of increasing SNRs on response times. These results concur with the subjective results of perceived listening effort rating. Nevertheless, in spite of the observed overall effect of NR on response times we cannot statistically confirm a positive effect on response times to speech-in-noise after applying realistic NR for HI listeners.

REFERENCES

Baer, T., Moore, B. C., and Gatehouse, S. (**1993**). "Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: Effects on intelligibility, quality, and response times," J. Rehabil. Res. Dev., **30**, 49-49.

Brons, I., Houben, R., and Dreschler, W. A. (**2014**). "Effects of noise reduction on speech intelligibility, perceived listening effort, and personal preference in hearing-impaired listeners," Trends Hear., **18**, 1-10.

Byrne, D., Dillon, H., Ching, T., Katsch, R., and Keidser, G. (**2001**). "NAL-NL1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures," J. Am. Acad. Audiol., **12**(1).

Desjardins, J. L., and Doherty, K. A. (**2014**). "The effect of hearing aid noise reduction on listening effort in hearing-impaired adults," Ear Hearing, **35**(6), 600-610.

Gatehouse, S., and Gordon, J. (**1990**). "Response times to speech stimuli as measures of benefit from amplification," Br. J. Audiol., **24**(1), 63-68.

Hicks, C. B., and Tharpe, A. M. (**2002**). "Listening effort and fatigue in school-age children with and without hearing loss," J. Speech Lang. Hear R., **45**(3), 573-584.

Houben, R., van Doorn-Bierman, M., and Dreschler, W. A. (**2013**). "Using response time to speech as a measure for listening effort," Int. J. Audiol., **52**(11), 753-761.

Melzer, I., and Oddsson, L. I. (**2004**). "The effect of a cognitive task on voluntary step execution in healthy elderly and young individuals," J. Am. Geriatr. Soc., **52**(8), 1255-1262.

Sarampalis, A., Kalluri, S., Edwards, B., and Hafter, E. (**2009**). "Objective measures of listening effort: Effects of background noise and noise reduction," J. Speech Lang. Hear R., **52**(5), 1230-1240.

van den Tillaart-Haverkate, M., de Ronde-Brons, I., Dreschler, W. A., and Houben, R. (**2017**). "The influence of noise reduction on speech intelligibility, response times to speech, and perceived listening effort in normal-hearing listeners," Trends Hear., **21**, 1-13.

Verhaeghen, P., and Cerella, J. (**2002**). "Aging, executive control, and attention: A review of meta-analyses," Neurosci. Biobehav. Rev., **26**(7), 849-857.

Wang, D. (**2005**). "On ideal binary mask as the computational goal of auditory scene analysis," in *Speech Separation by Humans and Machines*, Springer, Boston MA, pp. 181-197.

Wang, D., Kjems, U., Pedersen, M. S., Boldt, J. B., and Lunner, T. (**2009**). "Speech intelligibility in background noise with ideal binary time-frequency masking," J. Acoust. Soc. Am., **125**(4), 2336-2347.