Effects of attention on the speech reception threshold and pupil response of people with impaired and normal hearing

Koelewijn, Thomas; Versfeld, Niek J; Kramer, Sophia E

Published in:
Hearing Research

DOI:
10.1016/j.heares.2017.08.006

Hearing Research 354 (2017) 56–63

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research Paper

Effects of attention on the speech reception threshold and pupil response of people with impaired and normal hearing

Thomas Koelewijn*, Niek J. Versfeld, Sophia E. Kramer

Section Ear & Hearing, Department of Otolaryngology-Head and Neck Surgery and Amsterdam Public Health Research Institute, VU University Medical Center, Amsterdam, The Netherlands

A R T I C L E I N F O

Article history:
Received 23 February 2017
Received in revised form
21 August 2017
Accepted 25 August 2017
Available online 30 August 2017

Keywords:
Hearing impairment
Attention
Pupil response
Speech reception
Location uncertainty

A B S T R A C T

For people with hearing difficulties, following a conversation in a noisy environment requires substantial cognitive processing, which is often perceived as effortful. Recent studies with normal hearing (NH) listeners showed that the pupil dilation response, a measure of cognitive processing load, is affected by 'attention related' processes. How these processes affect the pupil dilation response for hearing impaired (HI) listeners remains unknown. Therefore, the current study investigated the effect of auditory attention on various pupil response parameters for 15 NH adults (median age 51 yrs.) and 15 adults with mild to moderate sensorineural hearing loss (median age 52 yrs.). Both groups listened to two different sentences presented simultaneously, one to each ear and partially masked by stationary noise. Participants had to repeat either both sentences or only one, for which they had to divide or focus attention, respectively. When repeating one sentence, the target sentence location (left or right) was either randomized or blocked across trials, which in the latter case allowed for a better spatial focus of attention. The speech-to-noise ratio was adjusted to yield about 50% sentences correct for each task and condition. NH participants had lower ('better') speech reception thresholds (SRT) than HI participants. The pupil measures showed no between-group effects, with the exception of a shorter peak latency for HI participants, which indicated a shorter processing time. Both groups showed higher SRTs and a larger pupil dilation response when two sentences were processed instead of one. Additionally, SRTs were higher and dilation responses were larger for both groups when the target location was randomized instead of fixed. We conclude that although HI participants could cope with less noise than the NH group, their ability to focus attention on a single talker, thereby improving SRTs and lowering cognitive processing load, was preserved. Shorter peak latencies could indicate that HI listeners adapt their listening strategy by not processing some information, which reduces processing time and thereby listening effort.

1. Introduction

Hearing loss can result in a degraded representation of the auditory scene, which makes it harder to differentiate target speech from competing sounds (Shinn-Cunningham and Best, 2008). By making more auditory information available through acoustic amplification (e.g. by using hearing aids), listening in complex situations may become easier. For instance, benefits of bilateral over unilateral hearing aid fittings have been shown for listening conditions that require spatial auditory attention (Noble and Gatehouse, 2009). Conversely, more auditory input leads to the necessity to process more information, which results in higher levels of listening effort, especially when there is uncertainty about the location of the speaker (Koelewijn et al., 2015, 2014a). One question is: how do audible binaural spatial cues affect listening effort during speech processing by people with sensorineural hearing loss?

Listening effort has recently been defined as 'the deliberate allocation of mental resources to overcome obstacles in goal pursuit when listening' (Pichora-Fuller et al., 2016). Based on the attention model of Kahneman (1973) and a recent modified version of it called the Framework for Understanding Effortful Listening (FUEL, Pichora-Fuller et al., 2016), one can argue that attention, manipulated for instance by means of task instructions, can affect the allocation of cognitive resources and thereby performance. The

* Corresponding author. VUmc – Ear & Hearing, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands.
E-mail address: t.koelewijn@vumc.nl (T. Koelewijn).

http://dx.doi.org/10.1016/j.heares.2017.08.006

availability of these resources is linked to levels of arousal of an individual. These levels of arousal can be measured as autonomic responses by means of pupillometry.

Recent studies (Koelewijn et al., 2015, 2014a) showed an effect of divided attention on the pupil dilation response for young normal hearing adults during processing of speech in noise. Consistent with FUEL it was shown that when participants were instructed to repeat two streams of masked speech instead of one, their performance dropped and their evoked pupil dilation response became larger (Koelewijn et al., 2014a). This is consistent with the idea that allocation of more resources (higher load) leads to larger pupil dilation (Just et al., 2003; Kahneman and Beatty, 1966). These pupillometry studies (Koelewijn et al., 2015, 2014a) were based on a design of Best et al. (2010), who showed that when normal hearing (NH) and hearing impaired (HI) participants were presented with two sentences in noise, one to each ear, performance dropped when both sentences had to be repeated instead of one. Apart from requiring a more favorable signal to noise ratio (SNR) than for the NH participants, the HI listeners' performance, when repeating one or two sentence over a range of fixed SNRs, was strikingly similar to that of the NH group.

Previous research (Kidd et al., 2005; Kitterick et al., 2010) showed that knowing *where* to listen has a positive effect on speech perception performance. Knowing where to listen also seems to reduce listening effort. When the location of the target speech was known, NH participant's pupil dilation response was significantly smaller than when the location was uncertain (Koelewijn et al., 2015). However, sensorineural hearing loss is known to affect binaural hearing (Moore, 1996) by affecting the ability to detect interaural time differences (ITD) and interaural level differences (ILD), both strong cues in spatial hearing in the horizontal plane. Additionally, binaural hearing is more strongly affected in people with an asymmetrical than with symmetrical loss, as is shown in studies using the Speech hearing, Spatial hearing and Qualities of hearing (SSQ) questionnaire (Gatehouse and Akeroyd, 2006; Noble and Gatehouse, 2004). Considering that sensorineural hearing loss has been shown to affect spatial hearing, it might also affect listening effort in spatially uncertain listening conditions.

The current study uses the pupil response to speech-in-noise processing as an objective measure of listening effort. During processing of an auditory event, both the mean pupil dilation (MPD) and peak pupil dilation (PPD) are known to be sensitive indices of cognitive processing load (listening effort). Peak latency, the time from stimulus onset to PPD (Zekveld et al., 2011), is an indicator of the speed of cognitive processing (e.g., Hyönä et al., 1995). Hence, a shorter latency may indicate faster cognitive processing. Peak latency is also affected by the amount of processed information (Koelewijn et al., 2015), with less information leading to shorter latencies. Additionally, the baseline pupil size prior to the pupil response provides information about an individual's anticipation of resource allocation for the task at hand (e.g., Aston-Jones and Cohen, 2005).

Effects of divided attention on the pupil dilation response and thereby listening effort have been found for young normally hearing adults during processing of speech in noise (Koelewijn et al., 2014a). It is not known, however, how attentional processes affect the pupil response of HI listeners. The aim of the current study was to explore how spatial manipulations of auditory attention would affect listening effort for adults with hearing loss. The question addressed was whether people with symmetrical mild to moderate hearing loss are able to effectively use spatial auditory cues to enhance speech perception and to lower their listening effort. More specifically, does dividing attention over two talkers instead of focusing on one and knowing the location of the target speech have an effect on performance and the pupil responses for HI participants? We additionally aimed to compare these findings with those obtained for NH listeners.

The PPD previously observed for NH listeners (Koelewijn et al., 2015, 2014a) was closely tied to the amount of attentional resources required and how effectively these could be deployed during speech processing in adverse listening conditions. Based on previous research we hypothesized that the HI group would require an overall increase in SNRs compared to the NH group (e.g., Festen and Plomp, 1990) to reach the same level of intelligibility in all listening conditions. Consistent with Best et al. (2010), it was hypothesized that both groups would require an increase in the SNR on dual-target task compared to the single-target task. Between tasks, consistent with previous results for NH participants (Koelewijn et al., 2015, 2014a), we expected both the NH and HI participants to show a larger PPD in the dual-target task than in the single-target task because of increased processing demands (Pichora-Fuller et al., 2016). Additionally, it was hypothesized that focusing attention on a location would enable listeners to filter out irrelevant information, which in turn would reduce processing load. This should lead to a smaller PPD and a decrease in SNR, as shown previously (Koelewijn et al., 2015). Finally, given that spatial hearing is affected by mild to moderate sensorineural hearing loss, it was hypothesized that, a difference between the HI and the NH group in the effect of location uncertainty (on SNR and PPD) would be observed.

2. Methods

2.1. Participants

Fifteen NH adults (2 males, 13 females, age between 33 and 66 yrs., median age 51 yrs.) and fifteen HI adults (4 males, 11 females, age between 34 and 72 yrs., median age 52 yrs.), recruited at the VU University Medical Centre, participated in the study. The sample size of this study was based on the outcomes of two previous studies (Koelewijn et al., 2015, 2014a). NH was defined as pure-tone thresholds less than or equal to 20 dB HL over the octave frequencies 0.25–4 kHz. A single 25 dB HL dip at one of these frequencies in one ear was allowed. NH participant's pure-tone hearing thresholds averaged over both ears and over the octave frequencies 1–4 kHz (three-frequency pure-tone average), ranged from 1.7 to 13.3 dB (dB) hearing level (HL) (mean = 8.1 dB HL, standard deviation (SD) = 3.3 dB). HI participants had a three-frequency pure-tone average, averaged over the two ears, ranging from 30.8 to 62.5 dB HL (mean = 47.2 dB HL, SD = 9.6 dB). The differences between the three-frequency pure-tone average of the better and poorer ears for the HI participants ranged from 0 to 5 dB, so all had symmetrical hearing loss. Mean audiograms for the better and poorer ears for both groups are shown in Fig. 1. All participants in the HI group had an air-bone gap less than or equal to 10 dB, in one (the better) ear at 1 and 2 kHz, indicating sensorineural hearing loss. Participants in both groups had no history of neurological diseases and reported normal or corrected-to-normal vision. They were native Dutch speakers and provided written informed consent in accordance with the Ethics Committee of the VU University Medical Center, Amsterdam.

2.2. Tasks and materials

Participants were presented with two different everyday Dutch sentences (Versfeld et al., 2000), one to each ear, simultaneously via headphones. An example sentence is ‘Hij maakte de brief snel open’, which means ‘He quickly opened the letter’. One sentence was spoken by a female talker (S1) and the other by a male talker (S2). Each sentence was masked by stationary noise (see below),

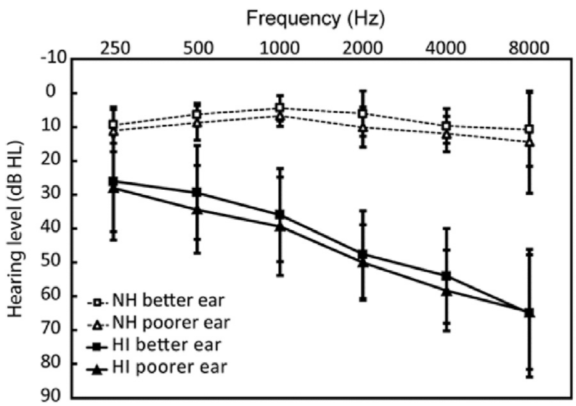

Fig. 1. Mean audiograms (dB HL) (ISO 389, 1991) for the better and poorer ears based on the three-frequency pure-tone average (1–4 kHz) of the participants in the NH and HI groups. Error bars show the standard deviations within each group for each ear and frequency.

which was independent at the two ears. In total, there were five conditions (see Table 1), subdivided into three 'single-target' conditions, one 'dual-target' condition, and one 'control' condition. In the single-target conditions, participants were asked to repeat S1 and ignore S2. In the first single-target task condition S1 was always presented to the left ear and in the second single-target condition, S1 was always presented to the right ear. In the third condition S1 was presented randomly to the left or right ear. In these conditions participants were told whether S1 would be presented to the same ear or at random to the left or right ear. The location fixed (left and right ear) and random conditions in the single-target task were included to assess the effect of location uncertainty. In the 'dual-target' condition, participants were asked to repeat S1 first and then S2, and the presentation side of S1 and S2 was randomized between trials. Finally, in the 'control' condition only S1 was presented, randomly to the left or right ear in the presence of speech shaped stationary noise that was uncorrelated at the two ears. Participants had to repeat S1. The dual-target and single-target conditions were included to compare the effects of divided and focused attention. The control condition was included to contrast the single-target condition and in order to investigate the effect of S2 on S1 processing. The five conditions, containing 40 trials each, were presented block-wise and the order was balanced between participants in a Latin square design. Participants were informed about what condition was going to be presented before the start of each block. Whilst performing the listening tasks, listeners did not receive any feedback.

After each condition, participants were asked to give subjective effort, performance, and motivation ratings. The questions were similar to the ones used in previous studies (Koelewijn et al., 2012a; Zekveld et al., 2010). For the effort rating, participants had to indicate how much effort it took on average to perceive the speech during each condition. This was rated on a visual analog scale from 0 ('no effort') to 10 ('very effortful'). To obtain an indication of how the participants perceived their own performance on the task, they were asked to estimate the percentage of sentences they had perceived correctly. The range was from 0 ('none of the sentences were intelligible') to 10 ('all sentences were intelligible'). Finally, to assess motivation during the course of the test, participants were requested to indicate how often during each condition they had abandoned the listening task because the task was too difficult. The range was from 0 ('this happened for none of the sentences') to 10 ('this happened for all of the sentences'). Prior to analysis the motivation rating was reversed, so that higher ratings reflected higher motivation.

Table 1
Overview of the tasks presented in each block, showing the stimuli presented, the side of presentation, and what target to repeat.

Condition	Task	Presentation side	Stimuli presented	Repeat
1	Single-target	fixed (S1 left ear)	S1, S2	only S1
2		fixed (S1 right ear)	S1, S2	only S1
3	Single-target	random	S1, S2	only S1
4	Dual-target	random	S1, S2	S1, then S2
5	Control	random	S1	S1

Shade shows the fixed location conditions in the Single target task.

Participants did not wear hearing aids during the listening tasks. Instead, sounds were spectrally shaped, for each participant and each ear individually, based on their pure tone thresholds. Before shaping, the speech level was adjusted to 55 dBA after which sound files were band pass filtered between 0.25 and 4 kHz with slopes of 48 dB/octave and shaped according to the NAL-R amplification rule (Byrne and Dillon, 1986). Note that this procedure was performed for both HI and NH participants. The long-term average spectra of the sentences spoken by the male talker and the associated stationary noise were matched to that of the female talker. Matching was performed to allow the same masker to be used on both ears while providing the same level of energetic masking. This had as a benefit that participants had no indication of what ear the male or female talker would be presented to based on the 3 s of noise presented prior to the onset of the sentence. Bandpass filtering was performed in Adobe Audition and all other sound manipulations were performed using Matlab. The speech level was always fixed and the masker level was adaptively varied by means of a staircase procedure in order to obtain a speech reception threshold (SRT) corresponding to 50% intelligibility for S1, also in the dual-target task. This allowed comparison of the SRT and the pupil dilation response for the processing of S1 at 50% intelligibility when the task was to either ignore or to process S2. Note that the stimuli were identical in the single-target and dual-target tasks, but different sets of sentences were used and the task instruction differed. After correctly repeating the whole S1 sentence, for which the scoring was strict, the masker level for the following trial was increased by 2 dB *at both ears*. After an incorrect response, the masker level decreased by 2 dB at both ears. For each condition, the SNR for the first trial started below threshold at −8 dB for the NH group and at −4 dB for the HI group. The first sentence of each condition was repeated, while the masker level was decreased in steps of 2 dB, until the participant correctly repeated S1. This was also the case for the dual-target task, for which performance on S2 did not influence the staircase procedure. S2 performance was scored separately in percentage words correct. Performance and the pupil dilation response for the first five trials were excluded from analysis. The onset of the noise masker was 3 s prior to the onset of both sentences and continued for 3 s after the end of the longer of the two sentences. The mean duration of the sentences was 1.9 s for the female talker (range = 1.3–2.7 s, SD = 0.26 s) and 2.0 s for the male talker (range = 1.3 = 2.9 s, SD = 0.31 s). At the end of each trial, a 0.5-s 1000-Hz prompt tone was presented, after which participants were allowed to respond. Participants responded verbally and their response was scored in real time by the experimenter. Participants were instructed to repeat the whole sentence or as many words as they could recall. SRT scores were used as the performance measure for S1 and proportion of words correct per sentence was used as the performance measure for S2. Prior to the experiment, participants were familiarized with the task by listening and responding to 8 practice trials for each task.

To test for linguistic ability independent of the auditory modality, participants also performed the text reception threshold

(TRT) task (Besser et al., 2012; Zekveld et al., 2007), which is a visual analog of the speech reception threshold task (Plomp and Mimpen, 1979). Similar TRT scores for the NH and HI group would indicate that differences in SRT scores were not based on differences in linguistic ability. In the TRT task sentences were visually presented on a computer screen in a red font (lower case Arial, vertical visual angle of 0.48°) on a white background partially masked by black vertical bars. These bars were evenly distributed across the screen and the width of the bars was varied by means of an adaptive staircase procedure, targeting the percentage of unmasked text required to read 50% of the sentences without any error. Sentences were presented on a screen word-by-word with word-onset timings similar to those for the corresponding recorded SRT sentences. After the onset of the last word the full sentence remained on the screen for 500 ms. Three lists of 13 sentences were presented, of which the first list was for practice purposes and was not included in the analysis. The TRT score was defined by the average percentage of unmasked text in the two remaining tests with the first four sentences of each list excluded. Lower TRT scores indicated better performance.

2.3. Apparatus and procedure

During the whole session participants were seated in a sound-treated room. Participants performed the TRT task followed by the speech perception tasks. During the TRT task participants were seated at about 60 cm from the computer screen. During the speech perception tasks participants were seated at approximately 3.5-m distance from a white wall. While listening to the sentences they had to gaze at a dot (diameter 0.47°) that was located at the participant's eye-height on the horizontal middle of the wall. An overhead light source illuminating the wall was placed at 3.5-m distance from the wall, outside the participants' field of view. The light intensity (lx) was adjusted such that, for each participant, the pupil diameter was around the middle of its dynamic range as measured by examination of the pupil size at 0 lx and 250 lx.

During the speech perception tasks, the pupil diameter of the left eye was measured by an infrared eye-tracker (SMI, 2D Video-Oculography, version 4) with a spatial resolution of 33 pixels per centimeter and a 50-Hz sampling rate. Separate audio files (44.1 Hz, 16 bit) for target sentences and maskers were mixed and presented binaurally from a PC by an external soundcard (asus Xonar Essence One) through headphones (Sennheiser, HD 280, 64 U) by a dedicated program (written in Matlab 2012a). After each condition, participants were asked to rate their perceived effort, performance and motivation. The whole test session, including measurement of pure-tone hearing thresholds, near vision acuity, practicing and performing the TRT task, fitting the eye-tracker, and practicing and performing the speech perception tasks with a 15-min break halfway through took 2–2.5 h.

2.4. Pupil data analysis

For each participant, the mean and SD of the pupil diameter were calculated for each pupil trace, recorded during each trial. These calculations were performed over a time period of 5.3 s including the baseline period that started 1 s before speech onset. Zero values and diameter values more than 3 SDs smaller than the mean diameter were coded as blinks. Traces in which more than 15% of their duration consisted of blinks were excluded from further analysis (3.1% of all trials). For the remaining traces, blinks were removed by linear interpolation between the fifth sample before and eighth sample after the blinks. A five-point moving average smoothing filter was passed over the de-blinked pupil traces to remove any high-frequency artifacts. A spike detection algorithm was used to detect eye movements on both the x- and y-traces. This algorithm used a 5-sample time window sliding in 1-sample steps, in which the maximum amplitude differences were calculated between all possible time point combinations within the window. The SD was calculated for each x- and y-trace between the start of the baseline and the response prompt. All trials for which the maximum x- or y-amplitude difference exceeded 2 SDs were excluded from analysis (13.8%). All remaining traces were baseline corrected by subtracting the trial's baseline value from the value for each time point within that trace. This baseline value was the mean pupil size within the 1-s period prior to the onset of the sentence (when listening to noise alone). The baseline period is shown by the left and middle dotted vertical lines in both plots in Fig. 3. Average traces in each condition were calculated separately for each participant. Within the average trace, mean pupil dilation (MPD, mm) was defined as the average pupil dilation relative to baseline within a time window ranging from the start of the sentence to the start of the response prompt, shown by the middle and right dotted vertical lines in both plots in Fig. 3. Within this same time window, the peak pupil dilation (PPD, mm) was defined as the largest value relative to the baseline. The latency of the PPD (ms) was defined relative to the sentence onset. Finally, for each participant and each condition the average pupil diameter at baseline was calculated.

2.5. Statistical analysis

For all dependent variables (SRT, MPD, and PPD), we performed a mixed 2×3 analysis of variance (ANOVA) (task analysis) with 'group' (NH and HI) as between-subject variable and 'task' (control, single-target, and dual-target; all random presentation) as repeated measure within-subject variable. For SRT, all pupil measures, and subjective motivation ratings, Mauchly's test of sphericity was significant. Therefore, for these measures the degrees of freedom were corrected by the Greenhouse-Geisser method. The fixed location condition was included only in the single-target task. Therefore, we performed a separate mixed 2×3 ANOVA (location analysis) with 'group' (NH and HI) as between-subject variable and 'location uncertainty' as repeated measure including the single-target random, fixed left, and fixed right conditions. For each dependent variable that showed a main effect of location, a planned comparison was performed in the form of two-sided paired-samples t-tests between the single-target conditions 1 and 2 (see Table 1) where the location of S1 was fixed on either the left or right ear. This is of importance because lateralized language processing in the brain is known to affect speech perception differently for speech presented to the left or right ear (Kimura, 1967).

3. Results

For the TRT task, participants in the NH group showed an average score of 58% (SD = 4.3), while the average score for the HI group was 59% (SD = 3.8), a Levene's Test for Equality of Variances showed no significant difference in variance ($p = 0.405$) and a two-sided independent samples *t*-test for Equality of Means showed no significant difference in means ($p = 0.334$) between groups. Average SRTs, pupil measures, and subjective ratings for each group as a function of condition are presented in Table 2, average SRTs are separately plotted in Fig. 2, and average pupil traces over participants, for each group and each condition, are plotted in Fig. 3.

3.1. Task effects

The outcomes of the *task analyses* (conditions 3, 4, and 5) on the SRTs, pupil measures, and subjective ratings are shown in Table 3. Analysis of the S1 SRTs showed main effects of group and task. The

Table 2
Average SRT, pupil measures, and subjective ratings for each group as a function of task/condition.

	Condition				
	Single-target			Dual-target	Control
Location	Fixed left	Fixed right	Random	Random	Random
Behavioral	SRT (SD), dB SNR				
NH	0.21 (1.21)	0.28 (1.64)	0.66 (1.47)	4.96 (3.45)	−0.41 (1.30)
HI	1.96 (2.24)	2.06 (2.00)	3.67 (4.90)	9.07 (4.44)	0.65 (1.66)
Pupil	Mean pupil dilation (SD), mm				
NH	0.13 (0.10)	0.12 (0.11)	0.15 (0.11)	0.23 (0.15)	0.11 (0.14)
HI	0.10 (0.08)	0.11 (0.08)	0.15 (0.05)	0.20 (0.10)	0.06 (0.06)
	Peak pupil dilation (SD), mm				
NH	0.26 (0.16)	0.25 (0.16)	0.28 (0.17)	0.38 (0.23)	0.23 (0.19)
HI	0.21 (0.09)	0.21 (0.11)	0.28 (0.07)	0.32 (0.14)	0.15 (0.09)
	Peak latency (SD), s				
NH	2.46 (0.84)	2.50 (0.90)	2.52 (0.72)	3.00 (0.46)	2.69 (0.99)
HI	2.25 (0.89)	2.30 (0.67)	2.06 (0.37)	2.67 (0.49)	2.00 (0.92)
	Baseline (SD), mm				
NH	3.89 (0.60)	3.91 (0.61)	3.99 (0.50)	4.10 (0.60)	3.87 (0.55)
HI	3.93 (0.56)	3.90 (0.63)	3.87 (0.55)	4.15 (0.67)	3.85 (0.60)
Subjective	Self-rated Effort (SD) (0 = low, 10 = high)				
NH	6.95 (1.23)	6.43 (1.33)	6.84 (1.11)	8.49 (0.83)	6.09 (1.57)
HI	6.79 (1.42)	7.46 (1.02)	7.51 (1.05)	7.91 (1.24)	6.82 (0.91)
	Self-rated Performance (SD) (0 = low, 10 = high)				
NH	5.31 (1.46)	5.04 (1.73)	5.37 (1.27)	3.51 (1.18)	5.56 (1.34)
HI	5.54 (1.60)	4.88 (1.39)	4.96 (1.51)	3.93 (1.58)	5.24 (1.45)
	Self-rated Motivation (SD) (0 = low, 10 = high)				
NH	2.25 (1.96)	1.96 (1.75)	8.17 (1.52)	7.61 (2.22)	8.04 (1.81)
HI	1.83 (1.30)	2.05 (1.26)	8.28 (0.88)	7.55 (1.55)	8.15 (0.96)

Shade shows the fixed location conditions in the Single target task.

NH participants showed overall lower (better) SRTs (average = 1.7 dB SNR) than the HI participants (average = 4.5 dB SNR) and for both groups the SRTs increased with the complexity of the listening task. Analysis of S2 performance (words correct) in the dual-target condition in the form of a Levene's Test for Equality of Variances showed no significant difference in variance ($p = 0.261$) and a two-sided independent samples t-test for Equality of Means showed no significant difference in means ($p = 0.833$) between the NH group (50%, SD = 14) and HI group (49%, SD = 21).

For the pupil measures, both MPD and PPD showed a main effect of task; the MPD and PPD were larger in the dual-target task than in the single-target task. The smallest MPD and PPD values occurred for the control condition where only one sentence was presented. No group effects or interactions were found for either the MPD or PPD. Peak latency showed a between-group effect and a main effect of task. The peak latency was shorter for the HI than for the NH group and the within-group peak latency was shorter for the single-target task than for the dual-target task. Finally, analyses of the pupil baseline showed a main effect of task; the baseline was largest for the dual-target task.

Analyses of self-rated effort showed an effect of task and a significant interaction between group and task. This interaction might be explained by a higher effort rating in the dual-target task for the NH group than for the HI group, while the HI group had relatively higher effort ratings than the NH group in the control and single-target tasks (see Table 2). However, post-hoc analysis using three two-sided heteroscedastic t-tests showed no significant differences between groups for the control ($p = 0.14$), single-target ($p = 0.1$), and dual-target ($p = 0.15$) tasks. For self-rated performance, there was a main effect of task, which was consistent with the SRTs. Analysis of self-rated motivation showed a violation of sphericity ($\chi^2(2) = 7.785$, $p = 0.02$) and therefore a Greenhouse-Geisser correction was performed. The data showed a main effect of task; motivation for both groups was smaller for the dual-target task than for the single-target task.

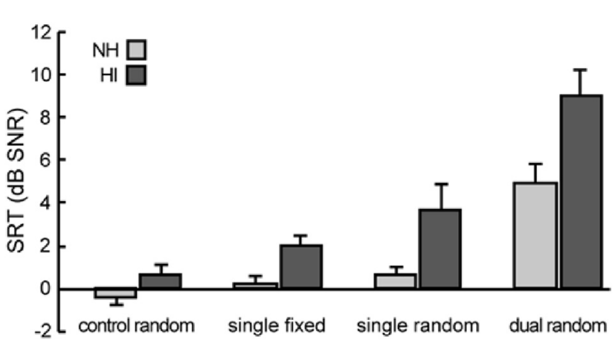

Fig. 2. The SRT for each task/condition, averaged over participants for each group. Error bars indicate the standard error of the mean.

3.2. Location uncertainty effects

Outcomes of the *location analyses* (conditions 1, 2, and 3) on the SRTs, pupil measures, and subjective ratings are shown in Table 4. Analysis of the SRTs showed main effects of group and location uncertainty. NH participants showed better SRTs than the HI participants. Both groups showed lower SRTs when the location was fixed than when it was random. Planned comparison between conditions 1 and 2 showed no effect of laterality ($t < 1$) between the two single-target fixed conditions. Hence, to give a clear depiction of the location effect on the SRT, conditions 1 and 2 in Fig. 2 are presented as one averaged value.

For the pupil measures, both MPD and PPD showed a main effect of location uncertainty; both MPD and PPD were smaller when the location was fixed than when it was random. No group effects or interactions were found for either the MPD or PPD. Planned comparisons showed no effect of laterality for either the MPD ($t < 1$) or PPD ($t < 1$). For a clear depiction of the location effect on the pupil responses, the mean responses of conditions 1 and 2 in Fig. 3 are presented as one average response. Location analysis showed no significant effects for peak latency or pupil baseline.

Analysis of self-rated effort showed a significant interaction between group and location uncertainty. Post-hoc analysis using three two-sided heteroscedastic t-tests showed no significant differences between groups for the fixed left condition ($p = 0.75$) and the random condition ($p = 0.1$). However, there was a significant difference between groups for the fixed right condition ($p = 0.02$) with HI listeners showing higher self-rated effort than NH listeners in the fixed right condition. Analysis of self-rated performance showed no significant effect of location uncertainty. Analysis of self-rated motivation showed a violation of sphericity ($\chi^2(2) = 9.146$, $p = 0.01$) and therefore a Greenhouse-Geisser correction was performed. After correction, the data showed no effect of location uncertainty on motivation.

3.3. Summary

There was an overall effect of group on the SRTs (worse mean SRT for the HI group) and a similar pattern of SRT scores within groups for the different conditions. Both task (control, single-target, and dual-target) and location uncertainty (fixed vs. random) affected peak pupil dilation. Listening to a single- rather than to a dual-target and to a fixed rather than to a random location, both reduced the MPD and PPD for both the NH and HI groups. Interestingly, in the task analysis (Table 3), HI participants showed a

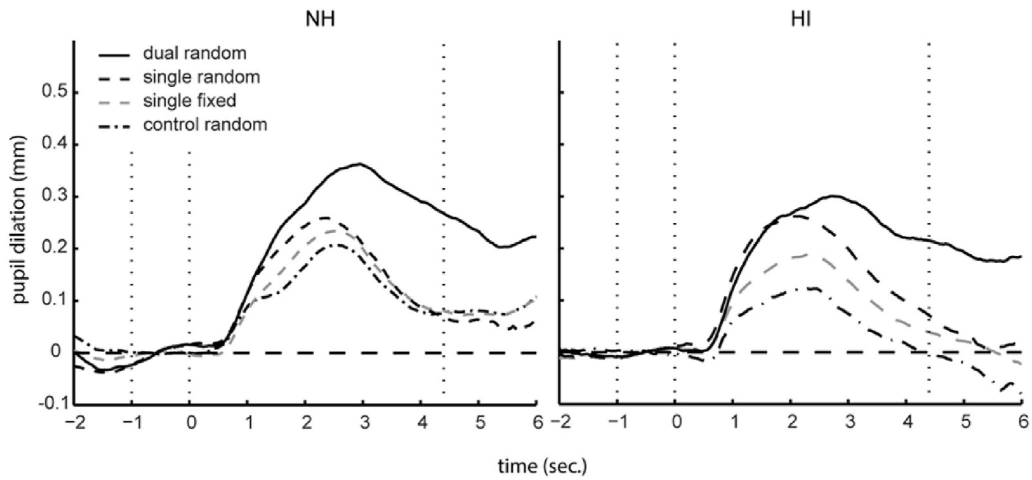

Fig. 3. Pupil responses for each task/condition, averaged over participants within each group. The onset of the sentences was at 0 s. The baseline, calculated as the average pupil diameter over 1 s preceding the start of the sentence, is shown by the dashed horizontal line. The time window over which the mean pupil dilation was computed corresponds to the range between the second and third dotted vertical lines.

Table 3
Outcomes of SRT, pupil measures, and subjective ratings using a mixed 2 × 3 ANOVA with between-subject variable 'group' (NH and HI) and repeated measure 'task' (control, single-target, and dual-target; all random presentation).

	Task analysis					
	Group		Task		Group*Task	
	$F_{[1,28]}$	*p*	$F_{[2,56]}$	*p*	$F_{[2,56]}$	*p*
SRT[a]	9.12	**0.005**	58.89	**<0.001**	2.82	0.080
Mean pupil dilation[a]	0.62	0.438	32.78	**<0.001**	2.02	0.160
Peak pupil dilation[a]	0.86	0.362	24.08	**<0.001**	1.53	0.230
Peak latency[a]	7.29	**0.012**	7.47	**0.004**	0.71	0.453
Baseline[a]	0.03	0.865	16.38	**<0.001**	1.67	0.203
Self-rated Effort	0.73	0.401	28.62	**<0.001**	5.08	**0.009**
Self-rated Performance	0.10	0.749	25.82	**<0.001**	1.72	0.188
Self-rated Motivation[a]	0.01	0.918	3.19	**0.036**	0.08	0.886

Significant p-values are presented in bold.
[a] Degrees of freedom for the averaged tests of significance were corrected by means of the Greenhouse-Geisser model.

significantly shorter peak latency than the NH participants. Additionally, an interaction was found between group and task for the subjective effort ratings, where the HI listeners seemed to experience less increase of effort with increasing task complexity than the NH group. Finally, an interaction was found between group and location uncertainty for the subjective effort ratings, where HI participants perceived more effort in the fixed-right condition than NH listeners.

4. Discussion

This study aimed to investigate the effect of attention on SRT and pupil measures for NH and HI participants. This was done by instructing participants to either focus attention on one target sentence or to divide attention over two target sentences, and by manipulating location uncertainty in the single-target task.

4.1. Task effects

Consistent with previous research (Best et al., 2010), both NH and HI participants showed higher (worse) SRTs for dual-target than for single-target performance. The impact of attention on the pupil dilation response did not differ between groups. As predicted by the FUEL model (Pichora-Fuller et al., 2016), both groups showed larger MPDs and PPDs with increasing task demand, i.e. when processing two sentences simultaneously instead of one. Processing two sentences instead of one also affected participants'

Table 4
Outcomes of SRT, pupil measures, and subjective ratings using a mixed 2 × 3 ANOVA with between-subject variable 'group' (NH and HI) and repeated measure 'location' (fixed left, fixed right, and random for the single-target task).

	Location analysis					
	Group		Location		Location*Group	
	$F_{[1,28]}$	p	$F_{[2,56]}$	p	$F_{[2,56]}$	p
SRT[a]	7.49	**0.011**	4.00	**0.038**	1.43	0.248
MPD	0.20	0.662	4.54	**0.015**	1.03	0.364
PPD	0.47	0.499	5.71	**0.006**	0.91	0.407
Peak latency	1.51	0.230	0.43	0.653	0.68	0.510
Baseline	0.02	0.895	0.13	0.880	1.68	0.195
Self-rated Effort	2.14	0.155	1.02	0.369	3.67	**0.032**
Self-rated Performance	0.06	0.816	1.85	0.166	0.89	0.415
Self-rated Motivation[a]	0.09	0.766	0.87	0.403	0.66	0.486

Significant p-values are presented in bold.
[a] Degrees of freedom for the averaged tests of significance were corrected by means of the Greenhouse-Geisser model.

'anticipation' of resource allocation as reflected by the larger pupil baseline in that condition (Aston-Jones and Cohen, 2005). Peak latency was longest in the dual-target task, indicating that processing two sentences instead of one took more time (Hyönä et al., 1995). This suggests that the two target sentences could not be processed fully in parallel or that processing was slower when resources were divided over two streams of incoming information. Consistent with previous studies (Koelewijn et al., 2015, 2014a) the dual-target task was also perceived as more effortful.

SRT, MPD, and PPD were all affected by processing interfering information in the contralateral ear as observed when contrasting the single-target task to the control task. This effect of interfering speech on the PPD was observed in previous studies for both NH and HI participants (Koelewijn et al., 2012a, 2014b). One of these studies (Koelewijn et al., 2012b) showed that the PPD being larger in the condition with an interfering talker masker than in the condition with an energetic masker correlated with working memory capacity. This may suggest that information from the distractor sentence in this study was (partly) processed, using working memory capacity, which may have resulted in additional processing load.

For both groups, task had an effect on subjective performance and motivation ratings. Participants seemed to have noticed their performance drop in the dual-target task. Although the speech-to-noise ratio was adjusted to yield about 50% sentence correct on S1 for each task and condition, the average performance in the dual-target task was rated based on the performance on both S1 and S2. Because S2 performance was 50% words correct, which translated in this study to around 25% sentences correct, their average performance over S1 and S2 in the dual-target task was indeed less than their performance on S1 in the single-target or control task. This might have affected the participant's motivation in performing the dual-target task, as reflected by the lower subjective ratings.

4.2. Location uncertainty effects

An effect of location uncertainty on the SRTs was shown for both groups and is consistent with previous results for NH participants (Kidd et al., 2005; Kitterick et al., 2010). Moreover, for both groups larger MPDs and PPDs were observed in the single-target random than in the fixed location condition, as previously shown for NH participants (Koelewijn et al., 2015). These results suggest that knowledge about the location of the target gave both the NH and HI groups an advantage. Note that the actual amount of information needed to be processed in the fixed and random conditions was the same. This might explain why there were no effects on peak latency and baseline. Remarkably, the effect of location uncertainty was not subjectively experienced, something that was also shown in our previous study of NH participants (Koelewijn et al., 2015).

There was no significant interaction between location uncertainty and group. This suggests that mild to moderate symmetrical hearing loss had no effects on spatial attention other than those found for the NH group. This disagrees with what was hypothesized based on previous research (e.g., Gatehouse and Akeroyd, 2006; Moore, 1996). It must be noted however that stimuli were presented dichotically over headphones, which provides optimal conditions for spatial separation. Stimuli presented through loudspeakers with closely spaced azimuths might have shown different results. Interestingly, a study by Zekveld et al. (2014) showed that although spatial separation influenced speech recognition performance, it did not influence the pupil dilation response of NH participants, indicating that spatial cues may reflect perceptual processing more than cognitive processing. Importantly, the current results suggest that HI participant's ability to direct spatial attention is not necessarily affected by sensorineural hearing loss in optimal circumstances.

4.3. Group effects

Consistent with previous research (e.g., Festen and Plomp, 1990), SRTs were higher for the HI than for than for the NH participants, despite the use of NAL-R amplification for both groups. For the pupil measures, the task analysis showed a significant group effect of peak latency. Shorter peak latency was observed for the HI than for the NH group, which means that maximum processing was reached faster, a surprising result that has not been reported before. A shorter peak latency may suggest that less information was being processed at a cognitive level (Koelewijn et al., 2015). It is common to miss or mishear parts of speech during a conversation, and to rely on redundancies in speech to extract the right message. People with hearing loss are especially well accustomed to this process, and might be more comfortable with missing information than listeners with normal hearing. As a benefit, shorter peak latency might reflect that HI participants preserved energy at a cognitive level. Thus, this newly observed latency effect may indicate different strategies for speech processing by HI listeners than by NH listeners. Since in the current study intelligibility levels were fixed, the reduced processing of redundant information did not appear in the results. However, previous research (Ng et al., 2013) suggests that for HI participants processing speech in noise can affect memory processing resulting in decreased recall of sentences that at first instance were recognized fully and correctly. Using such recall tasks in future studies might shed light on the use of such different strategies of listeners and provide valuable information.

No group effects were observed for the MPD or PPD. This might be due to the observed shift in SRTs for the HI compared the NH group. Note that HI participants showed similar pupil responses but in order to reach the required intelligibility levels this occurs at much more favorable SNRs. This tradeoff might explain why the results don't show an increase in the MPD, PPD or self-rated listening effort in HI participants compared to NH participants. A recent study by Ohlenforst et al. (2017) shows that when presenting speech in noise over a wide range of SNRs, for NH and HI participants the maximum PPD occurs at different SNRs but both at around 50% speech intelligibility. These results show that at a fixed SNR, PPDs can differ between NH and HI individuals. Note that these effects are not restricted to the pupil response as an index of listening effort (Wu et al., 2016).

4.4. Group similarities

Observing similar patterns of effects for NH and HI participants when manipulating task difficulty and when other measures of listening effort are used is not uncommon (Wu et al., 2016). One should keep in mind that in all conditions the SNR was adjusted to yield about 50% correct performance. The results showed that, in order to reach 50% correct, for both NH and HI participants the SNR needed to be increased as the task became more demanding (e.g., dual-target vs. single-target task), and in addition greater listening effort was required, as indicated by increased MPD and PPD. The current results extend our previous observations for young NH adults (Koelewijn et al., 2015, 2014a) to NH and HI participants with a broader age range.

Finally, in previous studies (Koelewijn et al., 2015, 2014a) using a similar design, performance data were in the form of proportion words correct and the SNR was fixed at −9, −3, and 3 dB. To compare the pupil response of the NH and HI groups at the same performance level, for this study the target intelligibility was fixed at 50% and the SRT (in dB SNR) was the dependent measure. That this did not affect the ability to show effects of divided versus focused attention and location uncertainty.

4.5. Summary and conclusions

The results show that the amount of information processed and the uncertainty of the target talker's location affect SRT scores and listening effort for NH and HI individuals in a similar manner. The HI group showed overall higher SRTs than the NH group in our complex listening situations, while there was the absence of a between group effect for the PPDs. This demonstrates that the HI participants had the same levels of listening effort, but at more favorable SNRs than the NH group. Importantly, the results show that listening effort is only partly related to hearing status. Notably, sensorineural hearing loss had no influence on HI participant's ability to focus their attention. One thus can argue that hearing loss causes an overall decrease in performance, but has not necessarily an effect on higher level processes such as attention. Finally, the results for peak latency suggest that HI listeners process less speech information, which reduces their total listening effort.

Acknowledgments

This work was financed by The Netherlands Organization for Scientific Research (NWO) (Veni grant 451-12-039). The authors thank the reviewers and the editor Brian Moore for their helpful comments during the revision of this manuscript.

References

Aston-Jones, G., Cohen, J.D., 2005. An integrative theory of locus coeruleus-norepinephrine function: adaptive Gain and Optimal Performance. Annu. Rev. Neurosci. 28, 403–450.
Besser, J., Zekveld, A., Kramer, S., Ronnberg, J., Festen, J., 2012. New measures of masked text recognition in relation to speech-in-noise perception and their associations with age and cognitive abilities. J. Speech Lang. Hear. Res. 55, 194–209.
Best, V., Gallun, F.J., Mason, C.R., Kidd, G., Shinn-Cunningham, B.G., 2010. The impact of noise and hearing loss on the processing of simultaneous sentences. Ear Hear 31, 213–220.
Byrne, D., Dillon, H., 1986. The National Acoustic Laboratories'(NAL) new procedure for selecting the gain and frequency response of a hearing aid. Ear Hear 7, 257–265.
Festen, J.M., Plomp, R., 1990. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. J. Acoust. Soc. Am. 88, 1725–1736.
Gatehouse, S., Akeroyd, M., 2006. Two-eared listening in dynamic situations. Int. J. Audiol. 45, 120–124.
Hyönä, J., Tommola, J., Alaja, A., 1995. Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. Quart. J. Exp. Psychol. A Hum. Exp. Psychol. 48, 598–612.
Just, M., Carpenter, P., Miyake, A., 2003. Neuroindices of cognitive workload: neuroimaging, pupillometric and event-related potential studies of brain work. Theor. Issues Ergon. Sci. 4, 56–88.
Kahneman, D., 1973. Attention and Effort. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.
Kahneman, D., Beatty, J., 1966. Pupil diameter and load on memory. Science 154, 1583–1585.
Kidd, G., Arbogast, T.L., Mason, C.R., Gallun, F.J., 2005. The advantage of knowing where to listen. J. Acoust. Soc. Am. 118, 3804–3815.
Kimura, D., 1967. Functional asymmetry of the brain in dichotic listening. Cortex 3, 163–178.
Kitterick, P.T., Bailey, P.J., Summerfield, A.Q., 2010. Benefits of knowing who, where, and when in multi-talker listening. J. Acoust. Soc. Am. 127, 2498–2508.
Koelewijn, T., de Kluiver, H., Shinn-Cunningham, B.G., Zekveld, A.A., Kramer, S.E., 2015. Hear. Res. 323, 81–90.
Koelewijn, T., Shinn-Cunningham, B.G., Zekveld, A.A., Kramer, S.E., 2014a. The pupil response is sensitive to divided attention during speech processing. Hear. Res. 312, 114–120.
Koelewijn, T., Zekveld, A.A., Festen, J.M., Kramer, S.E., 2014b. The influence of informational masking on speech perception and pupil response in adults with hearing impairment. J. Acoust. Soc. Am. 135, 1596–1606.
Koelewijn, T., Zekveld, A.A., Festen, J.M., Kramer, S.E., 2012a. Pupil dilation uncovers extra listening effort in the presence of a single-talker masker. Ear Hear 33, 291.
Koelewijn, T., Zekveld, A.A., Festen, J.M., Rönnberg, J., Kramer, S.E., 2012b. Processing load induced by informational masking is related to linguistic abilities. Int. J. Otolaryngol. 1–11.
Moore, B.C.J., 1996. Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids. Ear Hear. 17, 133–161.
Ng, E.H.N., Rudner, M., Lunner, T., Pedersen, M.S., Rönnberg, J., 2013. Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. Int. J. Audiol. 52, 433–441.
Noble, W., Gatehouse, S., 2009. Effects of bilateral versus unilateral hearing aid fitting on abilities measured by the Speech, Spatial, and Qualities of Hearing scale (SSQ). Int. J. Audiol. 45, 172–181.
Noble, W., Gatehouse, S., 2004. Interaural asymmetry of hearing loss, speech, spatial and Qualities of hearing scale (SSQ) disabilities, and handicap. Int. J. Audiol. 43, 100–114.
Ohlenforst, B., Zekveld, A.A., Lunner, T., Wendt, D., Naylor, G., Wang, Y., Versfeld, N.J., Kramer, S.E., 2017. Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation. Hear. Res. 351, 68–79.
Pichora-Fuller, K.M., Kramer, S.E., Eckert, M.A., Edwards, B., Hornsby, B.W.Y., Humes, L.E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C.L., Naylor, G., Philips, N.A., Richter, M., Rudner, M., Sommers, M.S., Tremblay, K.L., Wingfield, A., 2016. Hearing impairment and cognitive energy: the Framework for understanding effortful listening (FUEL). Ear Hear. 37 (Suppl. 1), 5S–27S.
Plomp, R., Mimpen, A., 1979. Improving the reliability of testing the speech reception threshold for sentences. Int. J. Audiol. 18, 43–52.
Shinn-Cunningham, B.G., Best, V., 2008. Selective attention in normal and impaired hearing. Trends Amplif. 12, 283–299.
Versfeld, N., Daalder, L., Festen, J., Houtgast, T., 2000. Method for the selection of sentence materials for efficient measurement of the speech reception threshold. J. Acoust. Soc. Am. 107, 1671–1684.
Wu, Y.-H., Stangl, E., Zhang, X., Perkins, J., Eilers, E., 2016. Psychometric functions of dual-task paradigms for measuring listening effort. Ear Hear. 37, 660–670.
Zekveld, A., George, E., Kramer, S., Goverts, S., Houtgast, T., 2007. The development of the text reception threshold test: a visual analogue of the speech reception threshold test. J. Speech Lang. Hear. Res. 50, 576–584.
Zekveld, A.A., Kramer, S.E., Festen, J., 2010. Pupil response as an indication of effortful listening: the influence of sentence intelligibility. Ear Hear. 31, 480–490.
Zekveld, A.A., Kramer, S.E., Festen, J., 2011. Cognitive load during speech perception in noise: the influence of age, hearing loss, and cognition on the pupil response. Ear Hear. 32, 498–510.
Zekveld, A.A., Rudner, M., Kramer, S.E., Lyzenga, J., Rönnberg, J., 2014. Cognitive processing load during listening is reduced more by decreasing voice similarity than by increasing spatial separation between target and masker speech. Front. Neurosci. 8, 88.