Pupil Responses of Adults With Traumatic Brain Injury During Processing of Speech in Noise

Thomas Koelewijn[1], José A. P. van Haastrecht[1], and
Sophia E. Kramer[1]

Abstract

Previous research has shown the effects of task demands on pupil responses in both normal hearing (NH) and hearing impaired (HI) adults. One consistent finding is that HI listeners have smaller pupil dilations at low levels of speech recognition performance ($\leqslant 50\%$). This study aimed to examine the pupil dilation in adults with a normal pure-tone audiogram who experience serious difficulties when processing speech-in-noise. Hence, 20 adults, aged 26 to 62 years, with traumatic brain injury (TBI) or cerebrovascular accident (CVA) but with a normal audiogram participated. Their pupil size was recorded while they listened to sentences masked by fluctuating noise or interfering speech at 50% and 84% intelligibility. In each condition, participants rated their perceived performance, effort, and task persistence. In addition, participants performed the text reception threshold task—a visual sentence completion task—that measured language-related processing. Data were compared with those of age-matched NH and HI participants with no neurological problems obtained in earlier studies using the same setup and design. The TBI group had the same pure-tone audiogram and text reception threshold scores as the NH listeners, yet their speech reception thresholds were significantly worse. Although the pupil dilation responses on average did not differ between groups, self-rated effort scores were highest in the TBI group. Results of a correlation analyses showed that TBI participants with worse speech reception thresholds had a smaller pupil response. We speculate that increased distractibility or fatigue affected the ability of TBI participants to allocate effort during speech perception in noise.

Keywords

listening effort, traumatic brain injury, pupillometry, speech reception in noise

Date received: 17 January 2018; revised: 21 September 2018; accepted: 14 October 2018

Introduction

The pupil dilation response to speech-in-noise processing is considered to be an objective measure of listening effort (Pichora-Fuller et al., 2016). There is a large range of studies showing that the pupil response is sensitive to properties of the speech perception tasks and stimuli used (for a review see, Zekveld, Koelewijn, & Kramer, 2018). In our previous studies, we demonstrated that pupil response is sensitive to the level of speech intelligibility (e.g., Kramer, Kapteyn, Festen, & Kuik, 1997; Zekveld, Kramer, & Festen, 2010) and type of background sound (Koelewijn, Zekveld, Festen, & Kramer, 2012; Wendt, Koelewijn, Książek, Kramer, & Lunner, 2018). We found this effect in both normal hearing (NH) and hearing impaired (HI) listeners.

According to the framework for understanding effortful listening (FUEL; Pichora-Fuller et al., 2016), listening effort depends on more than just task demands. Listener factors may also control the allocation of resources needed to overcome obstacles in goal pursuit. Research has shown that hearing status is one type of listener factor, which affects the pupil dilation response.

[1]Amsterdam UMC, Vrije Universiteit Amsterdam, Otolaryngology – Head and Neck surgery, Ear & Hearing, Amsterdam Public Health research institute, Amsterdam, The Netherlands

Corresponding Author:
Thomas Koelewijn, Dept. of Otolaryngology – Head and Neck Surgery, section Ear – Hearing, Location VUmc, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands.
Email: t.koelewijn@vumc.nl

To illustrate, while replicating the effects of intelligibility level and masker type on the pupil response, Zekveld, Kramer, and Festen (2011) observed a *smaller* pupil dilation response for HI compared with NH listeners when listening to speech in noise at intelligibility levels close to 50% correct performance. The same finding was observed by Kuchinsky et al. (2014) and Kramer, Teunissen, and Zekveld (2016) as well as Ohlenforst et al. (2017). Remarkably, other studies reported a *larger* pupil response in HI listeners compared with NH listeners. An example is a study by Ayasse, Lash, and Wingfield (2017) who used sentences that were audible to all participants. Ohlenforst et al. (2017) observed an interaction effect with *smaller* pupil dilations in HI listeners compared with NH listeners for low speech recognition performance in noise ($\leqslant$50%) but *larger* pupil dilations in HI listeners compared with NH listeners for high speech recognition performance in noise ($\geqslant$60%). Thus, whereas hearing loss does seem to be a listener factor influencing the pupil response, the association between hearing loss and the pupil dilation response is still complex (for a review see, Zekveld et al., 2018).

Apart from hearing loss, there are other listener factors that have shown to be associated with the pupil response. Examples are cognitive and linguistic ability (for a review see, Zekveld et al., 2018). An individual's linguistic processing ability can be assessed using a visual sentence completion test, known as the text reception threshold (TRT) task (Zekveld, George, Kramer, Goverts, & Houtgast, 2007). First, the TRT scores are known to be correlated with effects of interfering speech on both the speech reception threshold (SRT) and the pupil dilation response (e.g., Koelewijn, Zekveld, Festen, Rönnberg, & Kramer, 2012). Second, TRT scores correlate with working memory capacity measures (e.g., Koelewijn, Zekveld, Festen, Rönnberg, et al., 2012). This is of interest since a person's cognitive capacity could limits one's ability to exert effort (Pichora-Fuller et al., 2016). In addition, Wang et al. (2017) showed that during speech processing in noise, a higher level of fatigue was related to a smaller pupil response. These findings together indicate that changes in the pupil dilation response evoked during speech recognition in noise not only depend on task-related factors, but the magnitude of the pupil response seems to also vary across different listener groups. Importantly, previous research does not allow for fully dissociating the impact of hearing loss as diagnosed by a pure-tone audiogram (PTA) or by a speech-in-noise task on the pupil dilation response. Although FUEL suggests the pupil response to reflect central processing, to our knowledge, there is no study yet that has investigated the pupil response in people who report difficulties in speech processing in noise but who have a normal PTA. Hence, it is unclear whether the observed differences in the pupil response between NH and HI listeners are related to differences in the PTAs alone or whether differences in central (functional) speech processing may underlie these differences.

A significant proportion of people with traumatic brain injuries (TBIs) report difficulties and high levels of effort when listening to speech in a noisy background even when these individuals have normal thresholds in the audiogram (Zhao & Stephens, 2007). A recent review by Werff (2016) encompasses an extensive overview of the impact mild TBI may have on auditory processing. Often both frontal and temporal regions of the brain are damaged in people with TBI caused by external force. These regions are typically vulnerable for trauma (French, Taber, Helmick, Hurley, & Warden, 2011), caused by fast acceleration or deceleration in movement, for example, a car accident or a fall. Given that temporal regions are often damaged, it is generally known that TBI may affect functions related to primary auditory processing needed for speech segregation. In this study, we included participants with TBI caused by external force or a CVA. All participants had a normal PTA but they reported serious difficulties understanding speech in a noisy background. These characteristics make the TBI group a unique group of listeners. The aim of this study was to examine their pupil dilation responses evoked by an SRT task (Plomp & Mimpen, 1979) and compare their responses with those of NH and HI listeners. Testing this group of TBI participants allowed the effects of functional hearing difficulties on the pupil response to be investigated, despite normal thresholds in the audiogram.

For this study, we used the same design and tasks as used in our previous studies including NH and HI participants without any reported neurological problems (Koelewijn, Zekveld, Festen, & Kramer, 2014; Koelewijn, Zekveld, Festen, Rönnberg, et al., 2012). We used the SRT task to measure an individual's ability to perceive sentences presented in noise. Participants listened to speech in fluctuating noise and to speech masked by a single talker at two different levels of intelligibility. The SRT task was used in combination with pupillometry, allowing us to investigate the effort individuals exploited in these different listening situations. Both peak pupil dilation (PPD) and mean pupil dilation (MPD) can be used as parameters to measure listening effort (Zekveld et al., 2011). PPD latency is an indicator of the time needed for cognitive processing (e.g., Hyönä, Tommola, & Alaja, 1995), and the baseline pupil size prior to the pupil response is considered to be an index of arousal or cognitive resources exerted in preparation for the task (e.g., Aston-Jones & Cohen, 2005). An individual's linguistic processing ability was also assessed by administering the TRT task (Zekveld et al., 2007). We used the same setup as in our previous studies because our primary goal was to examine whether we

would be able to replicate the effects of intelligibility level and masker type on both the SRT and pupil parameters in this unique group of participants with TBI.

In summary, the objective of this study was twofold. The first goal was to examine whether we would be able to replicate the effects of intelligibility and masker type on the pupil response in a group of listeners with TBI. Based on previous research showing that the pupil response is sensitive to intelligibility level and masker type independent of hearing status, we hypothesized (H1) that TBI participants would show similar effects. The second goal was to compare the results of the TBI participants with those of NH and HI listeners. We hypothesized (H2a) that the participants with TBI would have an increased SRT, indicating worse speech processing in noise, compared with what has been previously shown for NH participants without any reported neurological problems (Koelewijn, Zekveld, Festen, Rönnberg, et al., 2012). The pupil dilation response was recorded while participants performed the SRT tasks. Based on the previously shown difference in PPD between NH and HI participants, we hypothesized (H2b) that a smaller PPD for the TBI compared with NH participants would be observed. Finally, we tested whether any speech processing deficit, when observed, would be restricted to the auditory domain. This was verified by having the TBI participants performing a TRT task. We hypothesized (H2c) that if the effect of TBI is restricted to auditory processing, TRT scores for the TBI group should be similar to those of NH and HI participants without any neurological problems (Koelewijn et al., 2014; Koelewijn, Zekveld, Festen, Rönnberg, et al., 2012). This would indicate that cognitive capacity needed for language processing was not affected in the included group of TBI participants.

Methods

Participants

Twenty adults (aged between 26 and 62 years, mean age 47 years, four males) with a history of acute TBI participated in this study. Of them, seven participants reported TBI, nine reported a CVA, and four participants had another unspecified cause of TBI. All participants were recruited via a Dutch TBI information website (hersenletsel-uitleg.nl) and at meeting points (brain injury-cafés) where Dutch people with traumatic or non-TBI gather. All participants reported difficulties with listening to speech in background sound and for at least 1 year did not receive TBI-related medical treatment or rehabilitation therapy. All participants had a normal PTA, which averaged over both ears and across octave frequencies 1 to 4 kHz. was 9.1 dB Hearing Level (HL, $SD = 4.8$ dB). All had normal or corrected-to-normal vision and were screened for near-vision acuity (Bailey & Lovie, 1980).

They were native Dutch speakers and provided written informed consent in accordance with the Ethics Committee of the VU University Medical Center.

SRT—Task

Speech perception in noise was measured using SRT tasks (Plomp & Mimpen, 1979) with speech presented in fluctuating noise or with a single-talker masker (Festen & Plomp, 1990; Koelewijn et al., 2014). Target sentences were everyday Dutch sentences (Versfeld, Daalder, Festen, & Houtgast, 2000), spoken by a female voice. An example of an everyday sentence is "Hij maakte de brief snel open," which directly translates to "He quickly opened the letter." The single-talker masker contained concatenated sentences from another set uttered by a male talker. Fluctuating noise was created by multiplying noise by the speech amplitude envelope of the single-talker masker for two separate frequency bands below and above 1 kHz (Festen & Plomp, 1990). Both the single-talker masker and the fluctuating noise had a long-term average frequency spectrum identical to the spectrum of the target speech signal. The value of the SRT (dB signal-to-noise ratio [SNR]) was estimated for speech presented at 50% and at 84% intelligibility (sentences entirely correct) using a 1-up-1-down (Plomp & Mimpen, 1979) and a 4-up-1-down staircase procedure (Levitt, 1971), respectively. For each condition, the target speech level was fixed at 55 dB SPL.

The onset of the masker was 3 s prior to the onset of the target sentence and continued for 3 s after the offset of the target sentence. The length of each trial varied with the length of the presented sentence, which had a mean duration of 1.9 s (range 1.4–2.5 s). At the end of the trial, a 1000 Hz prompt tone was presented for 1 s after which participants were instructed to respond. Manipulation of both masker type and intelligibility level resulted in a total of four different conditions that were presented in a blocked fashion (i.e., 50% fluctuating noise, 50% single talker, 84% fluctuating noise, and 84% single talker). Each block contained 39 trials and the order of the blocks was counterbalanced (Latin square) over participants. The SRT was calculated for each condition (block) by averaging the SNRs at which the sentences were presented, with the exception of the first four sentences that were excluded. Prior to the experiment, participants were familiarized to the task at an intermediate sentence intelligibility level of 71% (2-up-1-down procedure) by listening and responding to 13 practice sentences for both masker types.

SRT—Self-Ratings

While performing the SRT tasks, listeners did not receive any feedback. After each block, participants were asked to rate their effort, performance, and quitting

(task persistence; Koelewijn, Zekveld, Festen, & Kramer, 2012; Zekveld et al., 2010). For the effort rating, participants indicated how much effort it took on average to perceive the speech during that last block. This was rated on a visual analog scale from 0 (*no effort*) to 10 (*very effortful*). In addition, they were asked to estimate the percentage of sentences they had perceived correctly. This was rated from 0 (*none of the sentences were intelligible*) to 10 (*all sentences were intelligible*). Finally, to assess quitting rate, participants were requested to indicate how often during the last block they had abandoned the listening task because the task was too difficult. This was rated from 0 (*this happened for none of the sentences*) to 10 (*this happened for all of the sentences*).

TRT—Task

As in previous studies (Koelewijn et al., 2014; Koelewijn, Zekveld, Festen, Rönnberg, et al., 2012), participants performed a TRT task (Zekveld et al., 2007). During this 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 computer screen and the width of the bars was varied depending on the required percentage of unmasked text. Sentences were presented word-by-word with onset timings similar to the corresponding recorded SRT sentences. After the onset of the last word, the full sentence remained on screen for 500 ms (Besser, Koelewijn, Zekveld, Kramer, & Festen, 2012). Before this experiment, an adaptive 1-up-1-down procedure was applied, targeting the percentage of unmasked text required to read 50% of the sentences without any error. Four lists of 13 sentences each were presented. The first list was for practice purposes only and was excluded from the analysis. The TRT score was defined by the average percentage of unmasked text in the three remaining tests under exclusion of the first four sentences of each list. Lower scores indicate better performance.

Apparatus and Procedure

All testing was performed in a sound-treated room. After recording the participant's audiogram, the test session started with the SRT tasks. During the SRT tasks, participants were asked to fixate their gaze at a dot (diameter 0.47°) located at eye level on a white wall at 3.5-m distance. During the SRT test, the pupil diameter of the left eye was measured using an infrared eye-tracker (SMI, 2D Video-Oculography, version 4). The light intensity 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 01× and 2501× (see also, Zekveld et al., 2010).

For the SRT tasks, audio in the form of stereo wave files (44.1 Hz, 16 bit) was presented binaurally by an external soundcard (Creative SoundBlaster, 24 bit) through headphones (Sennheisser, HD 280, 64 Ω). The TRT test was administered following the SRT tasks. During the TRT task, participants were seated in front of a computer screen (Dell, 17 in.) at 60 cm viewing distance. All tests were presented by a PC running Windows XP (Dell, Optiplex GX745, 2.66 GHz 2 Core). The whole procedure, including measurement of pure-tone hearing thresholds, near vision acuity, fitting the eye-tracker, practicing and performing the SRT tasks, practicing and performing the TRT task, and a 15-min break halfway through plus additional breaks requested by participants took 2 to 3 h. Note that seven participants requested an additional break, one participant requested a longer break halfway through without an additional break, and one participant took a nap during the break halfway through without requesting an additional break.

SRT—Pupillometry

During the SRT tasks, for each participant and trial, the pupil diameter and pupil x- and y-coordinate traces were recorded at a 50 Hz-sampling rate. Pupil diameter and x- and y- traces of the first four sentences were omitted from further analysis. For all remaining diameter traces, the mean pupil diameter and *SD* were calculated in the interval from 1 s before sentence onset till the start of the response prompt. Zero values and diameter values more than 3 *SD*s smaller than the mean were coded as blinks. Traces containing more than 15% blinks were excluded from further analysis (on average over participants 4.0% of the traces). The remaining diameter traces were deblinked by means of linear interpolation that started four samples before an ended eight samples after the blinks. The x- and y-coordinate traces of the pupil center (reflecting eye-movements) were also *deblinked* by application of the same procedure. After deblinking on both the x- and y-coordinate traces, a spike detection algorithm was used to detect eye movements. This algorithm uses a 100 ms time window that slides with 20 ms steps in which maximum amplitude differences are calculated between all the time points within the window for each step. Traces containing an amplitude difference larger than 2 *SD*s were excluded (on average 15.9 % of the traces). The remaining traces were each separately low-pass filtered by means of 5-point moving averaging. The baseline value was calculated for each trace by computing the mean pupil size within the 1-s period prior to speech onset after which each individual trace was baseline corrected by subtracting a baseline value from the trace. Finally, within each subject and for each of the four conditions, the remaining traces were all processed as described and averaged.

Within these averaged traces, the MPD, PPD relative to baseline, and PPD latency were calculated. PPD was defined as the highest value (in mm) in the interval between sentence onset and the response prompt and relative to baseline, PPD latency as the timing of the PPD relative to the onset of the sentence, and MPD as the mean value (in mm) within the interval between speech onset and the response prompt.

Statistical Analysis

To examine whether we would be able to replicate the effects observed in previous studies (H1), a repeated measures analysis of variance (ANOVA) testing the effects of intelligibility (50% and 84%) and masker type (fluctuating noise and single-talker masker) was performed on the SRT score, the pupil measures, and the subjective ratings of the TBI participants. Statistically significant ($p < .05$) interactions were further analyzed by means of Bonferroni corrected ($\alpha = .05/m$) two-tailed paired samples t tests. All statistical analyses were performed using SPSS version 20.

For testing the Hypotheses 2a to 2c, the current results (PTA, SRT, PPD, self-rated effort, and TRT) were compared with data previously recorded for NH and HI participants without any neurological problems (Koelewijn et al., 2014; Koelewijn, Zekveld, Festen, Rönnberg, et al., 2012). The NH and HI groups used in our previous studies each included 32 participants. From each of these groups, we selected a subgroup of 20 participants that was age matched to the current group of 20 TBI participants. This allowed us to compare the scores of the current TBI group comprising 20 participants with those of the NH and HI groups. The creation of age-matched samples was accomplished by removing the 12 oldest participants from the NH and HI groups. After matching, the average age of the NH group was 46.5 years ($SD = 3.7$) and that of the HI group was 54.2 years ($SD = 7.9$). The average age of the TBI group was 46.6 years ($SD = 10.3$) and independent samples t tests showed a significant difference with the HI group ($p = .015$) but not with the NH group ($t < 1$). For testing H2a and H2b, we compared the PTA, the SRTs and PPDs both averaged over all four conditions of TBI participants to those of NH participants, by means of two-tailed independent samples t tests. For testing H2c, we used one-way ANOVA to compare the current TRT scores to those of NH and HI participants.

Results

Speech Reception Threshold

Analyses on the SRTs (see Table 1 and Figure 1) revealed a significant main effect of intelligibility,

Table 1. Average SRT, MPD, PPD, PPD Latency, Pupil Baseline, and Self-Rated Effort, Performance, and Quitting Rate Scores, as a Function of Intelligibility Level and Masker Type in the TBI Group.

	Intelligibility	Fluctuating	Single-talker
SRT		SNR (SD), dB	
	50%	−10.2 (1.3)	−11.3 (1.7)
	84%	−4.5 (1.9)	−3.8 (2.7)
Pupil		MPD (SD), mm	
	50%	0.07 (0.08)	0.09 (0.09)
	84%	0.03 (0.09)	0.05 (0.08)
		PPD (SD), mm	
	50%	0.18 (0.11)	0.20 (0.14)
	84%	0.12 (0.11)	0.15 (0.13)
		PPD latency (SD), s	
	50%	2.52 (0.90)	2.32 (0.42)
	84%	2.34 (0.93)	1.94 (0.73)
		Baseline (SD), mm	
	50%	4.30 (0.75)	4.30 (0.73)
	84%	4.14 (0.72)	4.23 (0.76)
Self-rated		Effort (Low = 0; high = 10)	
	50%	7.6 (1.3)	7.6 (1.6)
	84%	6.8 (1.5)	7.0 (1.3)
		Performance (Low = 0; high = 10)	
	50%	5.9 (1.3)	5.6 (1.1)
	84%	7.2 (1.3)	7.2 (0.8)
		Quitting rate (Low = 0; high = 10)	
	50%	1.5 (1.1)	2.0 (1.2)
	84%	1.0 (0.6)	1.3 (0.9)

Note. SRT = speech reception threshold; MPD = mean pupil dilation; PPD = peak pupil dilation; TBI = traumatic brain injury.

$F(1, 19) = 500.31$, $p < .001$, as indicated by lower SRTs for the 50% compared with the 84% intelligibility. No significant main effect of masker type ($F < 1$) was shown, but there was a significant interaction, $F(1, 19) = 8.15$, $p = .010$. Post hoc analysis ($\alpha = .0125$) revealed significant differences between the 50% and 84% intelligibility conditions for the fluctuating noise masker ($p < .001$) and the single-talker masker ($p < .001$), confirming the main effect of intelligibility level. However, the difference between the single-talker masker and fluctuating noise masker conditions was only statistically significant at 50% intelligibility ($p = .002$), as reflected by a lower SRT in the single-talker masker condition compared with the fluctuating noise masker condition, and not at 84% intelligibility ($p = .336$), which explained the interaction and the absence for a main effect of masker type.

Pupil Measures

Analysis on the MPDs (see Table 1 and Figure 2) revealed a significant main effect of intelligibility,
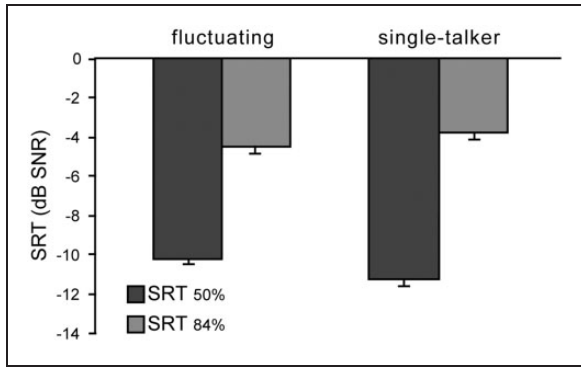
Figure 1. SRTs (dB SNR) at two intelligibility levels for both masker types, averaged over subjects. The error bars show the standard errors for each condition. SRT = speech reception threshold; SNR = signal-to-noise ratio.
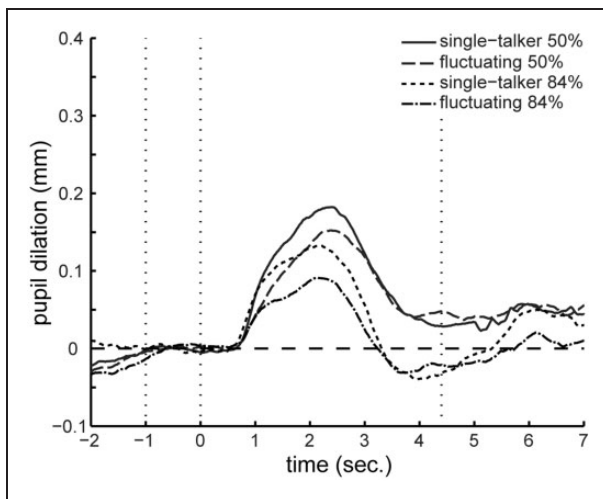

Figure 2. Pupil responses per condition averaged over participants. The onset of the sentences is at 0 s. The baseline is indicated as the average pupil diameter over 1 s preceding the start of the sentence. The area between the second and third dotted lines indicates the time window used for calculating the mean pupil dilation.

$F(1, 19) = 19.25$, $p = .000$, as reflected by larger MPDs at 50% compared with the 84% intelligibility. No significant main effect of masker type, $F(1, 19) = 2.09$, $p = .164$, and no interaction ($F < 1$) were shown. For the PPDs, there was a significant main effect of intelligibility, $F(1, 19) = 24.25$, $p = .000$, as reflected by larger PPDs at 50% compared with the 84% intelligibility condition. No significant main effect of masker type, $F(1, 19) = 2.95$, $p = .102$, or interaction ($F < 1$) was shown. Analyses on the PPD latency revealed no effect of intelligibility, $F(1, 19) = 3.72$, $p = .069$, no effect of masker type, $F(1, 19) = 2.84$, $p = .108$, and no interaction ($F < 1$). Finally, for the pupil baseline, an effect of intelligibility

was observed, $F(1, 19) = 9.99$, $p = .005$, with the baseline in the 50% condition being somewhat higher than in 84% intelligibility condition. The analysis of the pupil baseline did not reveal an effect of masker type ($F < 1$) and no significant interaction, $F(1, 19) = 1.69$, $p = .210$.

Self-Rated Scores

Analysis of the self-rated effort ratings (see Table 1) revealed a significant main effect of intelligibility, $F(1, 19) = 9.89$, $p = .005$, no effect for masker type ($F < 1$), and no interaction ($F < 1$). Analysis on self-rated performance data showed a significant main effect of intelligibility, $F(1, 19) = 63.49$, $p = .000$, no significant effect of masker type, $F(1, 19) = 1.130$, $p = .301$, and no interaction ($F < 1$). Self-rated effort was higher and self-rated performance was lower at 50% compared with 84% intelligibility. Finally, analysis on quitting rate showed an effect of intelligibility, $F(1, 19) = 22.91$, $p = .000$, an effect of masker type, $F(1, 19) = 7.116$, $p = .015$, and no interaction ($F < 1$). Quitting rate was higher at 50% compared with 84% intelligibility and higher for the single-talker compared with the fluctuating noise conditions.

Correlation Analyses

To make sure that interindividual differences in pupil dilation could not be explained by the absolute noise level of the masker, Pearson correlations were calculated between the pupil baseline and the SRTs for each of the four conditions ($\alpha = .0125$). None of these correlations were significant (single talker 50%, $r = -.165$, $p = .487$; fluctuating 50%, $r = -.050$, $p = .835$; single talker 84%, $r = .160$, $p = .500$; and fluctuating 84%, $r = -0.202$, $p = .393$). Three Pearson correlations ($\alpha = .0167$) between the PPD, SRT (averaged over conditions), and TRT were calculated. A significant correlation between the SRT and PPD ($r = -.629$, $p = .003$) was observed, with higher (worse) SRTs related to smaller PPDs (for a scatterplot see Figure 4).

Comparison of TBI, NH, and HI Groups (Testing H2a, H2b, and H2c)

The PTA between the TBI group, PTA (1–4 KHz) = 9.1 dB HL, and NH group, PTA (1–4 KHz) = 9.0 dB HL, did not differ significantly ($t = .087$, $p = .931$). For the NH and TBI group, the average SRT scores were calculated over conditions, which showed a significantly higher (poorer) SRT ($t = 4.133$, $p < .001$) for the TBI (SRT = −7.4 dB SNR) compared with the NH group (SRT = −9.3 dB SNR). This confirms H2a indicating that the TBI participants had worse speech processing in noise than the NH group, despite the absence of a

difference in PTA. For testing H2b, PPD scores between the TBI group (PPD = .16 mm) and the NH group (PPD = .22 mm) were compared. No significant difference was found ($t = -1.349$, $p = .185$). Self-rated effort scores between groups were compared with a one-way ANOVA. The results showed a significant group effect, $F(2, 57) = 8.782$, $p < .001$. Post hoc tests showed a significantly higher effort rating for the TBI group (mean = 7.2) compared with the NH ($p < .01$, mean = 5.8) and HI ($p < .01$, mean = 6.1) participants. For testing H2c, TRT performance between groups was compared with a one-way ANOVA, which showed no significant difference ($F < 1$) in mean score between the TBI group (mean = 59.1%, $SD = 4.3$), NH group (mean = 59.4%, $SD = 5.1$), and the HI group (mean = 59.1%, $SD = 4.3$).

Summary

The results of this study within the group of adults with TBI revealed significantly lower SNRs at 50% compared with 84% intelligibility and a significantly lower SNR for the single-talker masker compared with fluctuating noise at 50% intelligibility. The MPD, PPD, and PPD latency were all larger at the 50% compared with 84% intelligibility level, indicating higher listening effort over a longer period of time at lower intelligibility. However, contrary to what we observed in our previous studies, no effect of masker type on the pupil responses was found. The intelligibility effect was also observed for all *subjective ratings*, showing that at 50% intelligibility, participants experienced more effort, observed worse performance, and reported to quit listening more often than at 84% intelligibility. Interestingly, participants reported a higher quitting rate in the single-talker masker conditions compared with the fluctuating noise conditions. These findings indicate that previously reported intelligibility and masker type effects were shown for performance, but contrary to our expectations (H1), a main effect of masker type on the PPD was not replicated in listeners with TBI.

The results of the TBI group were statistically analyzed against results of the age-matched NH and HI groups. A group difference was observed for the SRTs that were slightly worse (higher) for the TBI group compared with the NH group, which confirms H2a. The PPD did not differ significantly between groups suggesting no difference in actual processing load while listening to speech in noise, which does not confirm H2b. Remarkably, the current result for the TBI group showed a negative correlation between SRT and PPD, showing participants with a worse SRTs to have overall smaller PPDs. The absence of a significant difference between the TBI, NH, and HI groups for the TRT suggests that the ability to complete sentences in the visual

modality was not affected for this group of TBI participants, confirming H2c. In addition, the absence of a difference in TRT scores provided no indication of a deviation in cognitive capacity for the TBI participants compared with the NH and HI groups.

Discussion

H1. Replication of Findings

This study revealed that the previously observed main effects of intelligibility level on the SRT and the pupil responses were replicated in the TBI group. A main effect of masker type on SRT as previously shown for NH participants was not observed in the TBI group. In our previous studies, we found a slightly but significantly better SRT for speech presented with a single-talker masker in the background compared with the condition with fluctuating noise. In the TBI group, this effect was only observed in the 50% intelligibility condition. These results are generally consistent with previous studies including NH participants without reported neurological problems (Koelewijn, Zekveld, Festen, Rönnberg, et al., 2012). Still, in our previous studies, a main effect of masker type on the pupil dilation response was found, with a larger response for the single-talker masker compared with the fluctuating masker at 50% and 84% intelligibility level (Koelewijn et al., 2014; Koelewijn, Zekveld, Festen, Rönnberg, et al., 2012). This effect was not found in the TBI group.

People with TBI are known to be easily distracted (Kaipio et al., 2000) and often report the inability to suppress irrelevant sounds (Lipowski, 1975). They report being constantly aware of every sound around them, which is often described as annoying or confusing. High distractibility may hinder the processing of incoming auditory information implying that it cannot be sufficiently sorted out and filtered accordingly. Hence, high level of distractibility in the TBI group might have overshadowed any performance benefit for the single-talker masker over the fluctuating noise. In all, these results may indicate disrupted attentional processing required during speech processing for the TBI group compared with that of both NH and HI groups without any reported neurological problems (Koelewijn et al., 2014; Koelewijn, Zekveld, Festen, Rönnberg, et al., 2012).

H2. Comparison of TBI Results With Those of NH and HI Listeners

The significantly higher (poorer) SRTs for the TBI group compared with the NH group in the absence of a difference in PTA confirms H2a. The results indicate that TBI participants have difficulties with listening to speech in noise despite a normal audiogram. Remarkably, no

significant differences were observed for the PPD between the TBI and NH group, indicating that there was no group effect of TBI on the PPD. Therefore, the idea that TBI would have a limiting effect on the deployment of cognitive resources, resulting in a smaller PPD (H2b), was not confirmed. Nevertheless, a trend was observed. This is further discussed in the following section.

Allocation of Effort

Figure 3 is an updated version of a previous figure (Koelewijn et al., 2014), which now includes the SRT and PPD data of the TBI group and age-matched NH and HI groups. Although no significant differences in PPD between the TBI and NH group were observed, the mean values of the PPDs measured for the TBI group were systematically smaller than those of the NH group for all conditions (see Figures 3). This is consistent with the previous observation showing that HI participants had a smaller pupil dilation response than NH participants (e.g., Zekveld et al., 2011) in listening conditions targeting 50% intelligibility. It is important to note, however, that the pupil dilation response in NH, HI, and probably also in TBI listeners depends on what intelligibility range is tested. This was shown by Ohlenforst et al. (2017). They examined the entire range from 0% to 100% sentence recognition performance and found that at higher levels of performance ($\geqslant 60\%$), pupil dilation was larger for HI than for NH

listeners. It is thus possible that the same interaction effect would have appeared in the TBI group if we had included a larger range of intelligibility levels in this study. This is an issue that needs to be addressed in future studies. Nevertheless, the question of why the pupil dilation response would be smaller for TBI group remains. Attentional problems, but also fatigue may underlie this phenomenon. This will be discussed in the following section.

The Possible Influence of Fatigue

Interestingly, there was a significant correlation between SRT and PPD in the TBI group. As shown in Figure 4, TBI participants with an overall higher (worse) SRT score showed a smaller PPD. This was observed while within participants a worse SRT for a relatively difficult listening condition normally related to a larger PPD (e.g., Zekveld et al., 2010). This result was not related to an increased peripheral hearing loss, as all TBI participants had a normal PTA.

The significant correlation might be related to diversity in levels of fatigue in the TBI group. High levels of listening effort over a longer period of time can result in fatigue (Hornsby, Naylor, & Bess, 2016). A substantial number (almost one third) of the people with TBI report abnormally high levels of fatigue (French et al., 2011; van Zomeren & van den Burg, 1985). To illustrate, 40% of a sample of Swedish stroke patients reported feeling fatigue *often to always* (Glader, Stegmayr, &
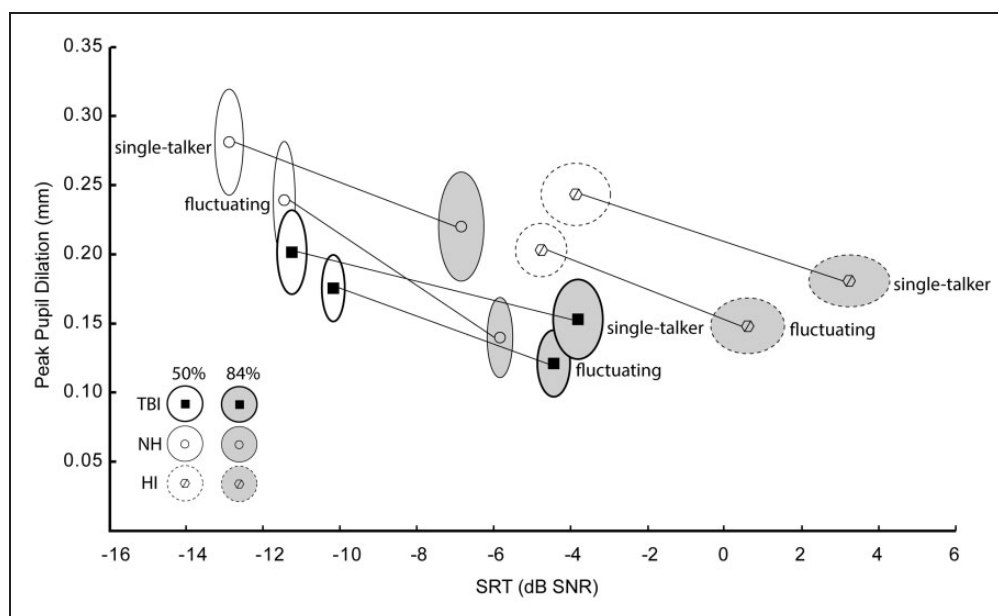

Figure 3. Average PPD as function of the average SRT for both masker type and intelligibility conditions for participants of the ABI, NH, and HI group. The ovals denote the standard errors of both the SRTs and PPDs. TBI = traumatic brain injury; NH = normal hearing; HI = hearing impaired.
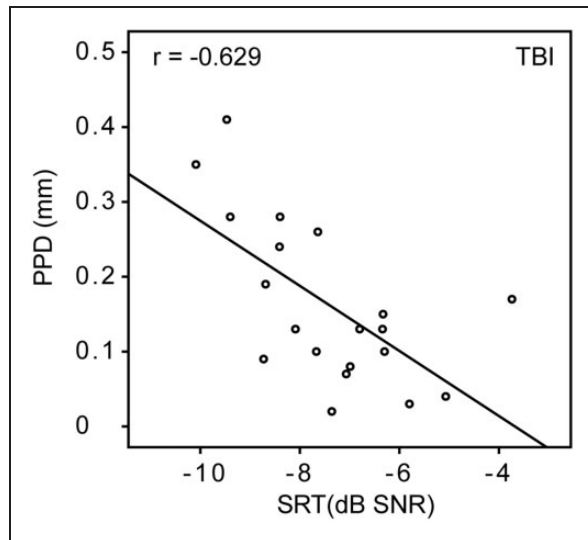
Figure 4. SRT (dB SNR) as a function of PPD both averaged over intelligibility level and masker type. TBI = traumatic brain injury; PPD = peak pupil dilation.

Asplund, 2002). According to FUEL (Pichora-Fuller et al., 2016), fatigue influences our processing capacity by affecting the way we evaluate our performance. High levels of fatigue could result in the preservation of resources by allocating less effort. This was suggested by Wang et al. (2017) who found that individuals with higher levels of fatigue had smaller task-related PPDs than individuals with lower levels of fatigue. The trend of overall lower PPDs in the TBI group (compared with the NH and HI group) and the significant correlation between SRT and PPD may indicate that fatigue limits the amount of resources a person can allocate to a certain task and may result in cognitive overload relatively early (i.e., higher (worse) SRTs) for people with smaller PPDs at a fixed intelligibility level (Figure 4; cf. Zekveld et al., 2011). As such, the high levels of fatigue could have served as a protective mechanism (cf. Wang et al., 2017), saving the listener from exploiting too much effort in difficult listening situations, hence the smaller pupil dilation compared with that of less fatigued NH listeners. This is speculation and deserves to be further examined in future research which should include measures of fatigue.

The Pupil Baseline

Finally, the current results showed an effect of intelligibility on the pupil baseline in the TBI group. In previous studies among NH and HI listeners using the same setup (Koelewijn, Zekveld, Festen, & Kramer, 2012; e.g., Zekveld et al., 2010), no effect of intelligibility on the pupil baseline was shown. A lower pupil baseline in the 84% intelligibility than in the 50% condition, as

observed in this study, could indicate a higher level of arousal at 50% intelligibility, which might reflect participants getting closer to depleting their cognitive resources in order to perform a more difficult task (Pichora-Fuller et al., 2016). Note that no significant correlations between baseline and SRT were observed indicating that it is unlikely that the absolute noise level used in the SRT task was the underlying cause of the effect of intelligibility on the baseline.

Limitations

Future research, if a sufficiently large sample of participants can be included, could incorporate 'type of traumatic brain damage' as a between-subject variable. However, it was beyond the goals of this study and was not achievable based on the small number of participants. In addition, future research should include independent and objective measures for distractibility and fatigue.

Conclusion

Participants with TBI in this study had normal PTAs and TRTs but showed difficulties with speech processing in noise, as reflected by worse SRTs compared with the NH group. This made them a unique group to study listening effort evoked by speech intelligibility in noise independently from differences in PTA. Previously reported effects of intelligibility on the SRT but not masker type were replicated in listeners with TBI. Most remarkably, in the TBI group, a significantly negative correlation between SRT and PPD was observed, indicating that given a fixed intelligibility level, higher (worse) SRTs related to smaller PPDs (Figure 4). Together with the absence of main effect of masker type for the TBI participants on SRT and the pupil dilation response, the results may suggest increased distractibility as a direct consequence of the TBI (Kaipio et al., 2000) or as suggested, related to high levels of fatigue (e.g., Boksem, Meijman, & Lorist, 2005). Explaining these observations in terms of depleted resources because of increased distractibility and fatigue is consistent with the FUEL (Pichora-Fuller et al., 2016) but still speculative. Further investigation is requested.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience*, *28*(1), 403–450. doi:10.1146/annurev.neuro.28.061604.135709

Ayasse, N. D., Lash, A., & Wingfield, A. (2017). Effort not speed characterizes comprehension of spoken sentences by older adults with mild hearing impairment. *Frontiers in Aging Neuroscience*, *8*, 296–312. doi:10.3389/fnagi.2016.00329

Bailey, I. L., & Lovie, J. E. (1980). The design and use of a new near-vision chart. *American Journal of Optometry and Physiological Optics*, *57*, 378–387. doi:10.1097/00006324-198006000-00011

Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2013). How linguistic closure and verbal working memory relate to speech recognition in noise—A review. *Trends in Amplification*, *17*(2), 75–93. doi:10.1177/1084713813495459

Boksem, M. A. S., Meijman, T. F., & Lorist, M. M. (2005). Effects of mental fatigue on attention: An ERP study. *Cognitive Brain Research*, *25*(1), 107–116. doi:10.1016/j.cogbrainres.2005.04.011

Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *The Journal of the Acoustical Society of America*, *88*(4), 1725–1736. doi:/10.1121/1.400247

French, L. M., Taber, K. H., Helmick, K., Hurley, R. A., & Warden, D. L. (2011). Traumatic brain injury in the military population. In *Combat and Operational Behavioral Health* (pp. 225–242). CreateSpace Independent Publishing Platform. doi:/10.1016/B978-0-323-54456-6.00005-0

Glader, E. L., Stegmayr, B., & Asplund, K. (2002). Poststroke fatigue: A 2-year follow-up study of stroke patients in Sweden. *Stroke*, *33*(5), 1327–1333. doi:10.1161/01.STR.0000014248.28711.D6

Hornsby, B. W. Y., Naylor, G., & Bess, F. H. (2016). A taxonomy of fatigue concepts and their relation to hearing loss. *Ear and Hearing*, *37*, 136S–144S. doi:10.1097/AUD.0000000000000289

Hyönä, J., Tommola, J., & Alaja, A. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *The Quarterly Journal of Experimental Psychology a, Human Experimental Psychology*, *48*(3), 598. doi:10.1080/14640749508401407

Kaipio, M. L., Cheour, M., Ceponiene, R., Öhman, J., Alku, P., & Näätänen, R. (2000). Increased distractibility in closed head injury as revealed by event-related potentials. *Neurophysiology*, *11*(7), 1463–1468. doi:10.1097/00001756-200005150-00021

Koelewijn, T., Zekveld, A. A., Festen, J. M., & Kramer, S. E. (2012). Pupil dilation uncovers extra listening effort in the presence of a single-talker masker. *Ear and Hearing*, *33*(2), 291. doi:10.1097/AUD.0b013e3182310019

Koelewijn, T., Zekveld, A. A., Festen, J. M., & Kramer, S. E. (2014). The influence of informational masking on speech perception and pupil response in adults with hearing impairment. *The Journal of the Acoustical Society of America*, *135*(3), 1596–1606. doi:10.1121/1.4863198

Koelewijn, T., Zekveld, A. A., Festen, J. M., Rönnberg, J., & Kramer, S. E. (2012). Processing load induced by informational masking is related to linguistic abilities. *International Journal of Otolaryngology*, *2012*(3), 1–11. doi:10.1155/2012/865731

Kramer, S. E., Kapteyn, T., Festen, J., & Kuik, D. (1997). Assessing aspects of auditory handicap by means of pupil dilatation. *International Journal of Audiology*, *36*(3), 155–164. doi:10.3109/00206099709071969

Kramer, S. E., Teunissen, C. E., & Zekveld, A. A. (2016). Cortisol, chromogranin A, and pupillary responses evoked by speech recognition tasks in normally hearing and hard-of-hearing listeners: A pilot study. *Ear and Hearing*, *37*, 126S–135S. doi:10.1097/AUD.0000000000000311

Kuchinsky, S. E., Ahlstrom, J. B., Cute, S. L., Humes, L. E., Dubno, J. R., & Eckert, M. A. (2014). Speech-perception training for older adults with hearing loss impacts word recognition and effort. *Psychophysiology*, *51*, 1046–1057. doi:10.1111/psyp.12242

Lipowski, Z. J. (1975). Sensory and information inputs overload: Behavioral effects. *Comprehensive Psychiatry*, *16*(3), 199–221. doi:10.1016/0010-440x(75)90047-4

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*, 467. doi:10.1121/1.1912375

Ohlenforst, B., Zekveld, A. A., Lunner, T., Wendt, D., Naylor, G., Wang, Y.,...Kramer, S. E. (2017). Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation. *Hearing Research*, *351*, 68–79. doi:10.1016/j.heares.2017.05.012

Pichora-Fuller, K. M., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E.,...Wingfield, A. (2016). Hearing impairment and cognitive energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing*, *37*(Suppl 1): 1–23. doi:10.1097/AUD.0000000000000312

Plomp, R., & Mimpen, A. (1979). Improving the reliability of testing the speech reception threshold for sentences. *International Journal of Audiology*, *18*(1), 43–52. doi:10.3109/00206097909072618

van Zomeren, A. H., & van den Burg, W. (1985). Residual complaints of patients two years after severe head injury. *Journal of Neurology, Neurosurgery, and Psychiatry*, *48*, 21–28. doi:10.1136/jnnp.48.1.21

Versfeld, N., Daalder, L., Festen, J., & Houtgast, T. (2000). Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *The Journal of the Acoustical Society of America*, *107*, 1671. doi:10.1121/1.428451

Wang, Y., Naylor, G., Kramer, S. E., Zekveld, A. A., Wendt, D., Ohlenforst, B., & Lunner, T. (2017). Relations between self-reported daily-life fatigue, hearing status, and pupil dilation during a speech perception in noise task. *Ear and Hearing*, *39*, 573–582. doi:10.1097/AUD.0000000000000512

Wendt, D., Koelewijn, T., Książek, P., Kramer, S. E., & Lunner, T. (2018). Toward a more comprehensive understanding of the impact of masker type and signal-to-noise ratio on the pupillary response while performing a speech-in-noise test. *Hearing Research*, *369*, 67–78. doi:10.1016/j.heares.2018.05.006

Werff, K. (2016). The application of the international classification of functioning, disability and health to functional auditory consequences of mild traumatic brain injury. *Seminars in Hearing*, *37*(03), 216–232. doi:10.1055/s-0036-1584409

Zekveld, A. A., Koelewijn, T., & Kramer, S. E. (2018). The pupil dilation response to auditory stimuli: Current state of knowledge. *Trends in Hearing*, *22*, 1–25. doi:10.1177/2331216518777174

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. *Journal of Speech, Language and Hearing Research*, *50*(3), 576. doi:10.1044/1092-4388(2007/040)

Zekveld, A., Kramer, S., & Festen, J. (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. *Ear and Hearing*, *31*(4), 480–490. doi: 10.1097/AUD.0b013e3181d4f251

Zekveld, A., Kramer, S., & Festen, J. (2011). Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response. *Ear and Hearing*, *32*, 498–510. doi:10.1097/AUD.0b013e31820512bb

Zhao, F., & Stephens, D. (2007). A critical review of King-Kopetzky syndrome: Hearing difficulties, but normal hearing? *Audiological Medicine*, *5*(2), 119–124. doi:10.1080/16513860701296421