

Objective Assessment of Listening Effort: Coregistration of Pupillometry and EEG

Kelly Miles^{1,2,3}, Catherine McMahon^{1,2}, Isabelle Boisvert^{1,2},
Ronny Ibrahim^{1,2}, Peter de Lissa^{2,4}, Petra Graham⁵, and
Björn Lyxell³

Trends in Hearing
Volume 21: 1–13
© The Author(s) 2017
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/2331216517706396
journals.sagepub.com/home/tia



Abstract

Listening to speech in noise is effortful, particularly for people with hearing impairment. While it is known that effort is related to a complex interplay between bottom-up and top-down processes, the cognitive and neurophysiological mechanisms contributing to effortful listening remain unknown. Therefore, a reliable physiological measure to assess effort remains elusive. This study aimed to determine whether pupil dilation and alpha power change, two physiological measures suggested to index listening effort, assess similar processes. Listening effort was manipulated by parametrically varying spectral resolution (16- and 6-channel noise vocoding) and speech reception thresholds (SRT; 50% and 80%) while 19 young, normal-hearing adults performed a speech recognition task in noise. Results of off-line sentence scoring showed discrepancies between the target SRTs and the true performance obtained during the speech recognition task. For example, in the SRT80% condition, participants scored an average of 64.7%. Participants' true performance levels were therefore used for subsequent statistical modelling. Results showed that both measures appeared to be sensitive to changes in spectral resolution (channel vocoding), while pupil dilation only was also significantly related to their true performance levels (%) and task accuracy (i.e., whether the response was correctly or partially recalled). The two measures were not correlated, suggesting they each may reflect different cognitive processes involved in listening effort. This combination of findings contributes to a growing body of research aiming to develop an objective measure of listening effort.

Keywords

alpha power, pupil dilation, listening effort, listening in noise, speech perception, spectral resolution

Date received: 15 June 2016; revised: 15 January 2017; accepted: 16 January 2017

Listening to speech in noise is a complex and effortful task, requiring a dynamic interplay between bottom-up and top-down processing (Arlinger, Lunner, Lyxell, & Pichora-Fuller, 2009; Rönnberg, Rudner, Foo, & Lunner, 2008; Zekveld, Heslenfeld, Festen, & Schoonhoven, 2006). Importantly, the effort required to listen to speech in noise is a commonly reported complaint in people with hearing impairment (Arlinger et al., 2009; Hawkins & Yacullo, 1984; Wouters & Berghe, 2001) that is not currently captured in standard clinical speech tests. Many factors may contribute to increased effort associated with listening to speech in noise, including age (Gosselin & Gagne, 2011; Larsby, Hällgren, Lyxell, & Arlinger, 2005; Tun, McCoy, & Wingfield, 2009) and cognitive influences such as working memory capacity and attention (Arlinger et al., 2009; Pichora-Fuller, 2006; Rudner, Lunner, Behrens, Thorén, & Rönnberg, 2012).

The term “listening effort” has been defined as “the mental exertion required to attend to, and understand, an auditory message” (p. 2) and has been studied from multiple perspectives (see McGarrigle et al., 2014, for a

¹Department of Linguistics, Australian Hearing Hub, Macquarie University, Sydney, Australia

²The HEARing Cooperative Research Centre, Melbourne, Australia

³Linnaeus Centre for HEaring And Deafness (HEAD), Swedish Institute for Disability Research, Linköping University, Sweden

⁴Department of Psychology, Australian Hearing Hub, Macquarie University, Sydney, Australia

⁵Department of Statistics, Australian Hearing Hub, Macquarie University, Sydney, Australia

Corresponding author:

Kelly Miles, Department of Linguistics, Macquarie University, Australian Hearing Hub 3.509, 16 University Avenue, Macquarie Park, NSW 2109, Australia.

Email: kelly.miles@mq.edu.au



review). Adverse health effects of prolonged mental effort, particularly with an effort-reward imbalance (Kuper, Singh-Manoux, Siegrist, & Marmot, 2002; Siegrist, 1996), have been linked to fatigue (Mehta & Agnew, 2012), cardiovascular strain (Peters et al., 1998), and stress (Hua et al., 2014). For listening-related effort in particular, adults with hearing loss report increased incidence of fatigue (Hua, Anderzén-Carlsson, Widén, Möller, & Lyxell, 2015; Kramer, Kapteyn, & Houtgast, 2006; Pichora-Fuller, 2003), are absent from work more frequently (Kramer et al., 2006), take longer to recover after work (Nachtegaal et al., 2009), and may withdraw from society (Weinstein & Ventry, 1982). There is also evidence that children with hearing loss experience greater fatigue than their normal-hearing peers, in part due to the effort required to listen to their teacher and interact with classmates, typically in acoustically poor classroom environments (Hornsby, Werfel, Camarata, & Bess, 2014).

Yet despite these negative health and social consequences of effortful listening, a reliable objective measurement of listening effort remains elusive (Bernarding, Strauss, Hannemann, Seidler, & Corona-Strauss, 2013). Current speech perception assessments only provide a crude estimation of the limitations of hearing impairment and do not typically consider the cognitive influences related to effort (Schneider, Pichora-Fuller, & Daneman, 2010; Wingfield, Tun, & McCoy, 2005) and the combination of, or interactions between, age and cognitive factors (Pichora-Fuller & Singh, 2006). Simultaneous evaluation of listening effort during speech recognition in noise could increase sensitivity of these assessments and guide device selection and settings as well as rehabilitation strategies.

A wide range of methods and tools have been used to explore listening effort that may better reflect the cognitive challenges which individuals with hearing loss face in real-world environments. Such measures have included subjective ratings (scales and questionnaires), dual tasks (performance measures on one task while the difficulty of a concurrent task varies), and physiologic measures such as changes in brain oscillations, pupillometry, skin conductance, and cortisol levels (see McGarrigle et al., 2014, for a comprehensive review). At present, pupil dilation and electroencephalogram (EEG) are the most-cited physiological measures that have the clinical potential to assess listening effort due to their noninvasiveness, increasing portability and user-friendliness (Badcock et al., 2013; Mele & Federici, 2012), and ability to be used during standard clinical speech perception assessments.

Changes in the pupillary response, which is under the physiological control of the locus coeruleus-norepinephrine system (Gilzenrat, Nieuwenhuis, Jepma, & Cohen, 2010), have been argued to reflect

increased processing load (Beatty & Wagoner, 1978; Granholm, Asarnow, Sarkin, & Dykes, 1996). Pupil size has been shown to be larger during sentence encoding when performance levels are low (e.g., 50% speech reception threshold [SRT]) in comparison to higher performance levels (e.g., 84% SRT; Koelewijn, Zekveld, Festen, Rönnberg, & Kramer, 2012; Kramer, Kapteyn, Festen, & Kuik, 1997; Zekveld, Festen, & Kramer, 2014; Zekveld, Heslenfeld, Johnsrude, Versfeld, & Kramer, 2014; Zekveld & Kramer, 2014; Zekveld, Kramer, & Festen, 2010). In addition to varying signal-to-noise ratios (SNRs) and SRTs, other speech stimuli manipulations have been shown to affect pupil dilation. For example, pupil size is larger when listening to speech in single-talker maskers than in fluctuating noise (Koelewijn et al., 2014; Koelewijn, Zekveld, Festen, Rönnberg, et al., 2012) or stationary noise (Koelewijn, Zekveld, Festen, & Kramer, 2012), or when the masking speech is the same gender as the speech signal (Zekveld, Rudner, Kramer, Lyzenga, & Rönnberg, 2014). Changing the complexity of the task through linguistic manipulation such as lexical competition (Kuchinsky et al., 2013) and sentence difficulty also increases the pupil size in the more linguistically challenging conditions (Piquado, Isaacowitz, & Wingfield, 2010; Wendt, Dau, & Hjortkjær, 2016). Degrading the spectral resolution of a speech stimulus through channel vocoding also increases pupil size (Winn, Edwards, & Litovsky, 2015). Collectively, these studies suggest that the pupillary response changes with task difficulty, which may reflect the increased effort associated with more challenging tasks.

Changes in brain oscillations is another physiological measure that has shown systematic changes associated with a wide variety of cognitive processes (Başar, Başar-Eroglu, Karakaş, & Schürmann, 2001; Herrmann, Fründ, & Lenz, 2010; Klimesch, 1996, 1999, 2012; Ward, 2003). For example, enhancement in the alpha frequency band (8–12 Hz) has been observed using working memory tasks with various types of stimuli, including syllables (Leiberg, Lutzenberger, & Kaiser, 2006) and single words (Karrasch, Laine, Rapinoja, & Krause, 2004; Pesonen, Björnberg, Hämäläinen, & Krause, 2006). Obleser, Wöstmann, Hellbernd, Wilsch, and Maess (2012) proposed that acoustic degradation and working memory load may similarly affect alpha oscillations due to the greater allocation of working memory resources required to comprehend an acoustically degraded signal (Piquado et al., 2010; Rabbitt, 1968; Wingfield et al., 2005). When memory load and signal degradation were most challenging, they demonstrated that alpha enhancement was superadditive, suggesting that the same alpha network might index both. Similar findings have been replicated using different sources of acoustic degradation during a

digit or word comprehension task (Petersen, Wöstmann, Obleser, Stenfelt, & Lunner, 2015; Wöstmann, Herrmann, Wilsch, & Obleser, 2015; cf. Strauß, Wöstmann, & Obleser, 2014, for review).

In our previous study (McMahon et al., 2016), we examined the changes in alpha power and pupil dilation during a speech recognition in noise task using noise-vocoded sentences (16 and 6 channels) and a four-talker babble-noise varying from -7 dB to $+7$ dB SNR. This demonstrated that the change in alpha power significantly declined with increasing SNRs for 16-channel vocoded sentences but remained relatively unchanged for 6-channel sentences. Pupil dilation showed a similar negative linear correlation for 16-channel vocoded sentences, which also was not observed for 6-channel vocoded sentences (instead showing a strong cubic relationship with SNR). Finally, changes in pupil dilation and alpha power were not correlated. However, differences in performance may explain the lack of correlation, as fixed SNRs were used across all participants. As such, for the present study, performance levels were fixed using individually obtained 50% and 80% SRTs for both 16-channel and 6-channel vocoded sentences. The signal-to-noise ratio was modulated to achieve these performance levels (and were therefore different across the participants). Manipulating two types of acoustic degradation approximately simulates listening in noise with a cochlear implant (Friesen, Shannon, Baskent, & Wing, 2001) and is relevant if objective measures of listening effort are to be applied in a clinical setting.

Alpha power and pupil dilation have both been proposed to be objective measures of listening effort with the possibility of implementation into clinical practice. Gaining a better understanding of how these physiological measures respond to changes in task difficulty within a population with normal hearing and cognition is important in order to interpret its behavior within an older population with hearing loss. The current study aimed to determine how pupil dilation and alpha oscillations change when parametrically varying both spectral degradation (using 16-channel and 6-channel noise vocoding) and speech recognition performance. Simultaneous recordings of pupil dilation and alpha oscillations will determine whether the measures are associated and might enable insight into whether they index the same aspect of listening effort.

Materials and Methods

Participants

Twenty-seven normal-hearing monolingual English-speaking participants were recruited for the study. Two participants did not attend all testing sessions and were therefore excluded. As one of the main aims of this study

was to assess whether EEG and pupil dilation correlated with each other, only participants who had 65% accepted trials in both measures were included ($n = 19$). Participants (12 women, 7 men) had a mean age of 27 years ($SD = 4.28$, range = 22–34 years). All had distortion-product otoacoustic emissions between 1 and 4 kHz, consistent with typical hearing or a mild sensorineural loss only and were right handed as assessed by *The Assessment and Analysis of Handedness: The Edinburgh Inventory* (Oldfield, 1971).

Speech Materials and Background Noise

Bamford–Kowal–Bench sentences adapted for Australian English (BKB-A; Bench & Doyle, 1979) were recorded by a native Australian-English female speaker. The sentences and background noise (four-talker babble) were vocoded using custom MATLAB scripts, where the frequency range 506000 Hz was divided into 6- or 16-logarithmically spaced channels. The amplitude envelope was then extracted by taking the absolute value from the Hilbert transform from each channel. The extracted envelope was used to modulate noise with the same frequency band. Each band of noise was then recombined to produce the noise vocoded sentences and background noise. The root mean square levels of the sentences and background noise were equalized in MATLAB after vocoding.

Speech Reception Thresholds

Automated adaptive speech-in-noise software developed by the National Acoustic Laboratories was used to obtain SRTs (see Keidser, Dillon, Mejia, & Nguyen, 2013, for a comprehensive review of the algorithm). The adaptive test has been validated with similar speech materials to the current study, with participants with normal hearing ($n = 12$) and hearing loss ($n = 63$), showing a standard deviation of 1.27 dB and 1.24 dB, respectively (Keidser et al., 2013). This suggests that the test results are reliable. The BKB-A sentences were presented at 65 dB Sound Pressure Level (SPL), and the background noise was adaptively adjusted to obtain the SRT. The adaptive procedure consisted of three phases. *Phase 1*: 5-dB steps until four sentences were completed, including one reversal, *Phase 2*: 2-dB steps until a minimum of four sentences were completed, and the phase's standard error (SE) was 1 dB or below, and *Phase 3*: 1-dB steps until 16 sentences (from the end of Phase 2) were completed, with a SE of .80 or below. When the SE reached .80 or below, the test terminated and recorded the SNR.

A sound-attenuated room was used during the testing sessions, and the equipment was calibrated prior to each participant's arrival. The speaker was positioned at 1 m and 0° azimuth from the participant. Participants were

informed they would hear sentences in noise and were instructed to repeat back all of the words of the sentence they heard. Two SRTs (50% and 80%) for two vocoding conditions (16 and 6 channels) were collected, resulting in four conditions. Table 1 shows the mean and standard deviations of the SNRs for each SRT.

Physiological Measures

The pupil and EEG recordings were measured simultaneously during the speech recognition task in a sound-attenuated and magnetically shielded room. Each participant was assigned randomly to Block A or Block B. Each block contained 220 sentences divided into four conditions (6ch50%SRT, 6ch80%SRT, 16ch50% SRT, and 16ch80% SRT) in which the

sentences were swapped in each block. For example, in Block A, a sentence presented in the 6ch50%SRT condition was presented in the 16ch80%SRT condition in Block B. Sentences were randomized during presentation. Each participant's SNR obtained during the behavioral session was used to present the sentences in each of the conditions. The top panel of Figure 1 shows the presentation protocol. Participants were instructed to repeat the sentence at the offset of the noise. Responses were recorded using a voice recorder and video camera set up directly in front of them to capture their face during recording allowing more accurate scoring of their responses at a later time. The sentences were scored at the word level (using the standard BKB/A scoring criteria) by a native Australian-English speaker and the percentage correct was averaged for each condition.

Table 1. Means and Standard Deviations of SNR (dB), True Performance Obtained During the Physiological Testing Session (%), and Pearson's r Correlation Coefficients of Pupil Dilation and Alpha Power, Presented by SRT and Channel Vocoding (First-Two Columns), and Channel Vocoding (Collapsed Across SRT) and SRT (Collapsed Across Channel Vocoding) in the Last Two Columns.

SRT %	50		80		–	–	50	80
Channel vocoding	6	16	6	16	6	16	–	–
SNR dB	0.4 (1.7)	–1.7 (1.4)	3.7 (1.9)	0.7 (1.7)	2.1 (2.4)	–0.7 (1.9)	–0.6 (2.0)	2.0 (2.4)
True performance %	39.2 (16.5)	53.0 (15.8)	60.6 (15.0)	68.8 (13.5)	49.9 (19.0)	60.9 (16.6)	46.1 (17.4)	64.7 (14.7)
Pearson's r	0.02 (0.14)	0.08 (0.20)	–0.04 (0.15)	–0.01 (0.17)	–0.01 (0.09)	0.03 (0.11)	0.06 (0.13)	–0.04 (0.08)

Note. SNR = signal-to-noise ratios; SRT = speech reception thresholds.

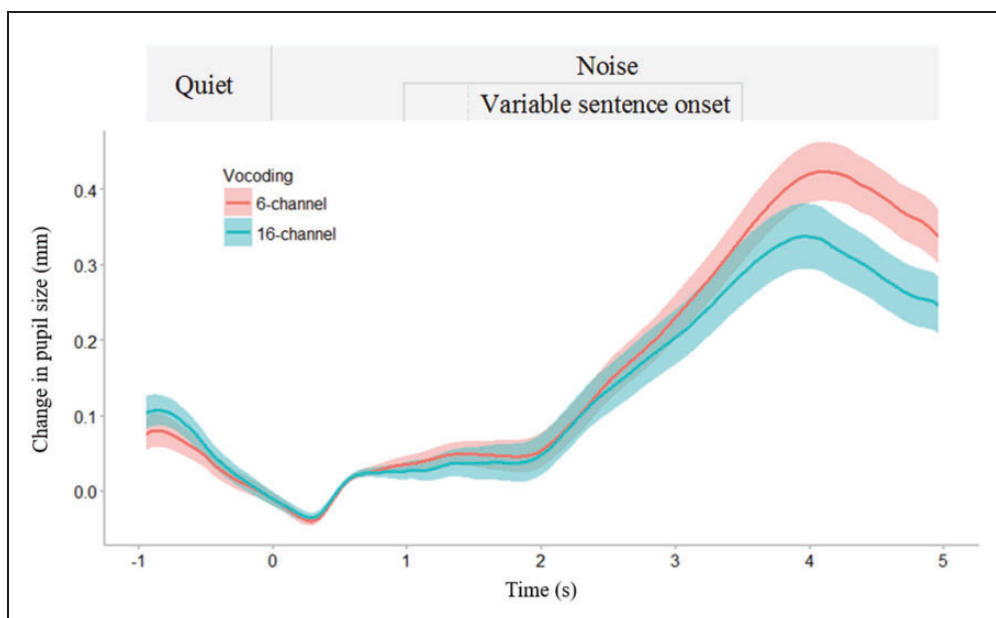


Figure 1. Average pupil size over time for all trials and participants, for 16- and 6-channel vocoding. The 0s timepoint refers to the beginning of noise. All sentences finished at the 3.5s timepoint. Shading represents ± 1 SE of the mean. The top panel represents the presentation protocol.

Pupil Recording

The right pupil was recorded using an SR Research Eyelink 1000 tower mount system at a 1,000 Hz sampling rate. Stimuli were presented through Experiment Builder software v1.10.1241. Prior to the task, the equipment was calibrated using a 9-point calibration grid on the screen. Pupil activity was recorded continuously until the experiment terminated. Off-line, single trials were processed with Dataviewer software (version 1.11.1) and compiled into single-trial pupil-diameter waveforms (−1 s to 5 s) for further processing and analyses using customized MATLAB scripts. Blink identification was determined on a trial-by-trial basis as pupil sample sizes smaller than three standard deviations below the mean pupil diameter. Trials containing more than 15% of the trial samples detected in blink were rejected. In the remaining accepted trials, linear interpolation was used to reduce lost samples and artifacts from blinks (Siegle, Ichikawa, & Steinhauer, 2008; Zekveld et al., 2010). Samples affected by blinks were interpolated from 66 ms preceding the onset of a blink to 132 ms following the offset of a blink. Data were smoothed using a 5-point moving average. Trials were then averaged across conditions for each participant. Regions of interest included baseline in noise (0–1 s) and the encoding period (2–6 s).

For each trial, relative percent change was calculated as maximum pupil size during encoding minus mean pupil size during baseline in noise, divided by the mean baseline in noise. This was then multiplied by 100 in order to report percent change from baseline. See Figure 1 for an example of the pupil response (shown in millimeters, not percent change) during the experiment.

EEG Recording

The continuous EEG was recorded with a 32-channel SynampsII Neuroscan amplifier. Thirty electrodes were positioned on the scalp in a standard 10 to 20 configuration (FP1 and FP2 were disabled as the participants rested their foreheads on the eyetracker support). The ground electrode was located between Fz and FPz electrodes. Electrical activity was recorded from both mastoids, with M1 set as the online reference. Ocular movement was recorded with bipolar electrodes placed at the outer canthi and above and below the left eye. Electrode impedances were kept below 5 k Ω . The signal from the scalp electrodes was amplified at 1,000 Hz, band-pass filtered between .01 and 100 Hz, and notch filtered online at 50 Hz.

Ocular artifact rejection was performed using Neuroscan software using a standard ocular reduction algorithm. Postprocessing was conducted in Fieldtrip-MATLAB (Oostenveld et al., 2011). The EEG data were epoched from −1 s to 5 s avoiding stimulus boundary artifacts caused by the filtering process. A two pass

reversal Butterworth filter with cutoff frequencies of 0.5 Hz to 45 Hz was applied to remove any drifts and high frequency noise that might occur. Band-pass filtering was used instead of high-pass filtering as in Obleser et al. (2012). Trials containing a variance exceeding 300 μV^2 were removed from further analyses. Trials were then two-pass band-pass filtered between 8 and 12 Hz to extract the alpha oscillation. The absolute value of the alpha band was extracted from the parietal electrodes (P3, P4, and Pz) during the encoding period (1 s duration finishing 200 ms before the end of the sentence) and baseline in noise (300–800 ms after the noise onset) on a trial by trial basis. See Figure 2 for a time–frequency representation of the EEG activity.

For each trial, relative percent change was calculated as mean alpha power during encoding minus mean alpha power during baseline, divided by mean alpha power during baseline. This was then multiplied by 100 in order to report percent change from baseline.

Statistical Methods

As the range of absolute values was considerably different between pupil dilation and alpha power change, relative percent change from baseline was used for all statistical analyses in order to facilitate comparisons between the two measures.

All analyses were performed in R version 3.2.1 using the nlme package (Pinheiro, Bates, DebRoy, & Sarkar, 2014). Linear mixed-effects (LME) models with a random intercept for individual were used for all analyses to control for repeated measures over different levels of the factors on individuals. Models for pupil size and alpha power were first assessed by including interactions. If there was no significant interaction, the main effects model was presented. *p* Values less than .05 were considered significant for all analyses.

Four LME regression models were developed to determine the effects of the task parameters on pupil dilation and alpha power (see Table 2). The first model assessed the effect of SRT (50% and 80%) and vocoding (16 and 6 channels) on these measures. Variability in performance for the targeted SRTs existed within the physiological experiment, despite using a validated adaptive method in the behavioral session to obtain these SRTs. This was particularly evident for the 80% SRT, where participants' true performance on the sentence recall task ranged from an average of 36 to 93% (see Table 2 *True performance %* for mean performance obtained for all conditions). To account for this variability, a second LME regression model was introduced, with true performance levels (i.e., individual speech recognition scores) and vocoding as predictor variables. A third LME model was developed to assess whether task accuracy influenced the measures. This was done because

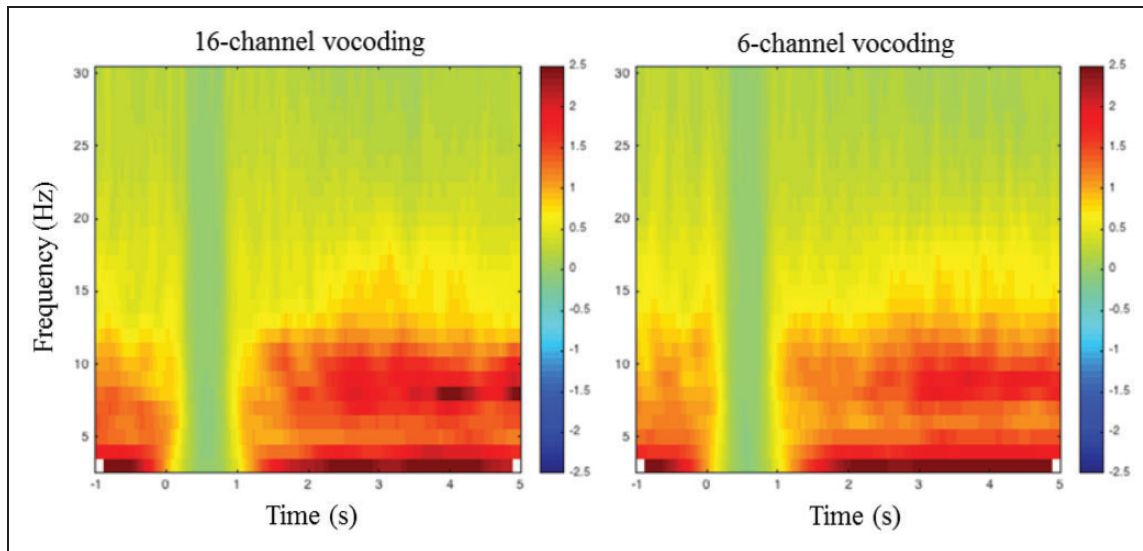


Figure 2. Time–frequency representation of the EEG activity averaged across all participants, in the parietal region, for 16- and 6-channel vocoding. The time–frequency representations are relative to the activity occurring during the 1 s of noise beginning at the 0 s timepoint. All sentences finished at the 3.5 s timepoint.

Table 2. Results for the Linear Mixed-Effects Models.

Model	Pupil size			Alpha power		
	Estimate	95% CI	<i>p</i> Value	Estimate	95% CI	<i>p</i> Value
Relative change in pupil/alpha						
Intercept	12.156	[9.145, 15.077]	<.001	135.910	[84.345, 187.475]	.000
SRT	0.045	[−0.773, 0.863]	.915	14.818	[−12.472, 42.108]	.287
Vocoding	1.491	[0.673, 2.308]	<.001	−29.991	[−57.285, −2.696]	.031
Relative change in pupil/alpha						
Intercept	14.908	[11.544, 18.271]	<.001	134.693	[62.225, 207.160]	.003
True performance	−0.046	[−0.072, −0.019]	<.001	0.140	[−0.722, 1.001]	.751
Vocoding	0.986	[0.119, 1.852]	.026	−28.261	[−57.310, 0.788]	.057
Relative change in pupil/alpha						
Intercept	11.191	[8.242, 14.138]	<.001	133.096	[81.166, 185.026]	<.001
Task accuracy	1.612	[0.676, 2.549]	<.001	7.661	[−23.206, 38.528]	.626
Vocoding	1.571	[0.644, 2.499]	<.001	−17.833	[−48.336, 12.670]	.252
Relative change in pupil/alpha						
Intercept	12.864	[9.956, 15.774]	<.001	130.475	[82.717, 178.233]	<.001
SNR	0.018	[−0.153, 0.189]	.836	−3.460	[−9.112, 2.191]	.230

Note. Reference levels: 50% SRT, 16-channel vocoding, correct recall (task accuracy). True performance (i.e., the actual percentage of speech recognized during the physiological testing session) was modelled in addition to SRT, as off-line sentence scoring was shown to deviate from target SRTs. SRT = speech reception thresholds; CI = confidence interval; SNR = signal-to-noise ratios.

the reason for an incorrect or not recalled response is variable and generally unknown. For example, if a participant is inattentive during a trial and does not recall a sentence, the physiological response may be different compared with if the participant invested effort to hear a sentence that was too challenging to perceive. Removing incorrect or not recalled sentences may

reduce variability in the measures. Therefore, only correct or partially recalled sentences were analyzed in the third model. Finally, a fourth model was constructed to assess whether changes in pupil size and alpha power were not merely due to changes in SNR (i.e., varying loudness across the conditions). Individual SNRs for each participant and condition were used in the model.

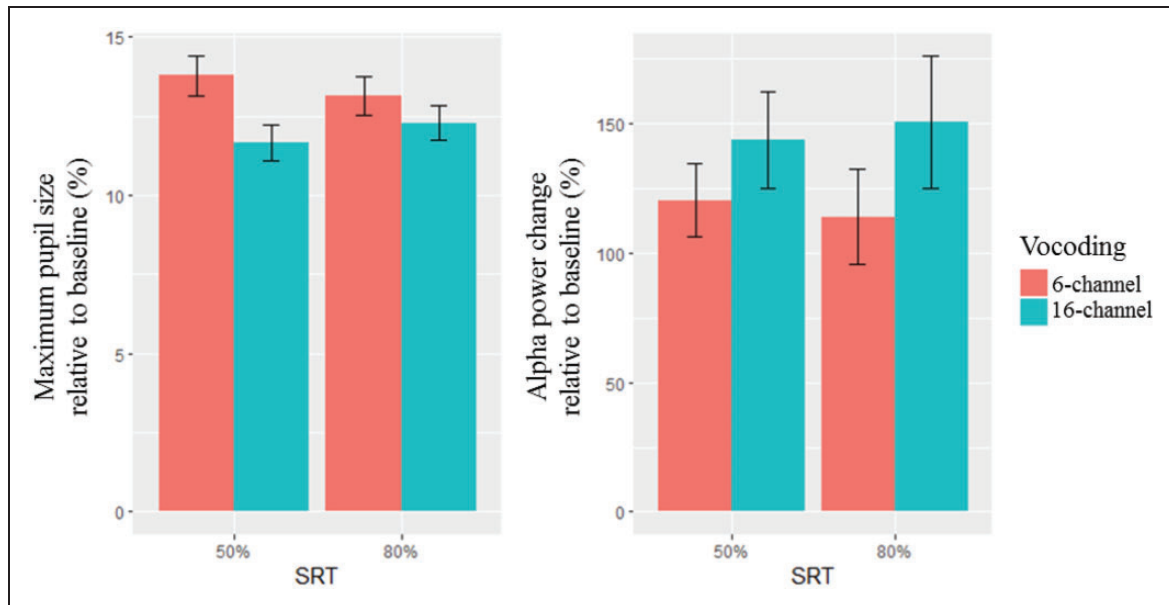


Figure 3. Mean \pm 1 SE of maximum pupil size and alpha power change relative to baseline, by SRT and channel vocoding.

To account for repeated measures on individuals, correlations presented in the results section are the average of the correlation coefficients calculated for each participant (average of the 19 coefficients for each of the 8 conditions). As there were unequal trials across the pupil and alpha measures, only those trials that were accepted in both measures were used in the correlation analyses ($n = 3,019$).

Results

SRT and Vocoding

To determine the effect of SRT and vocoding on changes in pupil dilation or alpha power, LME regression models with SRT and vocoding as predictor variables were developed. For pupil dilation, there was no significant interaction term ($p = .12$). A main effects model indicated no effect of SRT ($p = .91$), and a significant effect of vocoding ($p < .01$) on pupil size, which was 1.49% larger in the 6-channel condition compared with the 16-channel condition. For alpha power, there was no significant interaction term ($p = .62$). A main effects model indicated no effect of SRT ($p = .29$), and a significant effect of vocoding ($p = .03$). Alpha power change was -30.0% lower in the 6-channel condition compared with the 16-channel condition. Figure 3 shows the mean of maximum pupil size and alpha power change relative to baseline.

True Performance and Vocoding

As previously discussed, the true performance obtained on the speech recognition task during the physiological

testing session greatly differed from the targeted 50% and 80% SRTs (see Table 2). To determine whether this had any bearing on the physiological measures, LME regressions models with true performance (%) and vocoding as predictor variables were developed.

For pupil dilation, there was no significant interaction term ($p = .88$). A main effects model indicated a significant effect of performance level ($p < .01$) and vocoding ($p = .03$). As performance levels increased (toward 100%), the pupil size decreased by -0.05% . For alpha power, a LME regression model with performance level and vocoding as predictor variables indicated there was no significant interaction term ($p = .88$). A main effects model indicated no effect of performance level ($p = .75$). The inclusion of the participants' true performance level caused a loss of significance of channel vocoding ($p = .06$) in alpha power compared with the SRT and vocoding model. A likelihood ratio test suggested no benefit of including performance levels in the model (performance level and vocoding: log likelihood $-24,869.65$ vs. vocoding alone: log likelihood $-24,869.79$). Removing performance level indicated that vocoding was significant ($p = .03$). Alpha power was 29.9% greater in the 16-channel condition compared with 6-channel condition.

Task Accuracy and Vocoding

A LME model including task accuracy (partially correct versus correct sentence recall) and channel vocoding was developed. For pupil dilation, there was no significant interaction term ($p = .38$). A main effects model showed a significant effect of task accuracy ($p < .01$) and a

significant effect of vocoding ($p < .01$). Pupil size for partially correct sentence recall was 1.61% larger than correctly recalled sentences and 1.57% larger for 6-channel vocoded sentences compared with 16-channel vocoded sentences. For alpha power, there was no significant interaction term ($p = .17$). A main effects model showed no effect of task accuracy ($p = .63$) and no effect of vocoding ($p = .25$).

SNR

Finally, to determine whether SNR was influencing pupil size or alpha power, SNR was entered into an LME as a predictor variable, however this showed no significant effect on pupil size ($p = .84$) or alpha power ($p = .23$).

Correlation Between Pupil Size and Alpha Power Change

At the individual level, pupil size was not significantly correlated with alpha power change for any of the correlations ($n = 19$, $p > .05$ for all correlations; see Table 1 for means and *SDs*), including collapsing across all conditions ($n = 19$, mean $r = .01$, $SD = .08$, $p > .05$). Figure 4 shows individual Pearson's r coefficients for participants by vocoding.

To assess whether the lack of correlation between the measures may be due to intraindividual variability,

intraclass correlations (ICC) were conducted. ICC estimates and their 95% confidence intervals were calculated using the ICC package in R (Wolak, Fairbairn, & Paulsen, 2012). To balance the data set for ICC, the minimum number of trials available per subject, per condition, was first determined ($n = 19$). The results of the ICC analysis show a weak-to-strong degree of reliability for the alpha measurements (Table 3) and a very high degree of intraindividual reliability for the pupil measurements (Table 4). Alpha power was more variable than the pupil dilation measure; however, even in the conditions where there was strong reliability across both measures (e.g., 16 channels), the two measures were not correlated (Table 1).

Discussion

The aim of this study was to examine the effects of increasing listening effort during a sentence recognition-in-noise task on pupil dilation and alpha power change. The second aim was to determine whether these physiological measures that were recorded simultaneously were correlated. This would suggest that each respond to the same aspect of listening effort which may, itself, comprise multiple components (see Pichora-Fuller et al., 2016 for review). Listening effort was manipulated by parametrically varying the spectral content of the signal using noise vocoding (16 and 6 channels) and



Figure 4. Pearson's r correlation coefficients of pupil dilation and alpha power by participant and channel vocoding.

Table 3. Intraclass Correlation Coefficients (ICC) Assessing Intraindividual Reliability for Alpha Power.

SRT %	50		80		–	–	50	80
Channel vocoding	6	16	6	16	6	16	–	–
ICC	.38	.63	.22	.49	.50	.76	.73	.63
<i>p</i> Value	.055	<.001	.198	.012	.008	<.001	<.001	<.001
95% CI	[–0.11, 0.72]	[0.34, 0.83]	[–0.40, 0.65]	[0.08, 0.77]	[0.12, 0.77]	[0.58, 0.89]	[0.53, 0.88]	[0.35, 0.83]

Note. SRT = speech reception thresholds; CI = confidence interval.

Table 4. Intraclass Correlation Coefficients (ICC) Assessing Intraindividual Reliability for Pupil Dilation.

SRT %	50		80		–	–	50	80
Channel vocoding	6	16	6	16	6	16	–	–
ICC	.82	.84	.85	.85	.91	.92	.90	.92
<i>p</i> Value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
95% CI	[0.68, 0.92]	[0.71, 0.93]	[0.73, 0.93]	[0.73, 0.93]	[0.85, 0.96]	[0.86, 0.96]	[0.82, 0.95]	[0.85, 0.96]

Note. SRT = speech reception thresholds; CI = confidence interval.

performance (50% and 80% SRT). Specified SRTs were chosen rather than a fixed SNR to account for cognitive differences within the participant population (Souza & Arehart, 2015) and to investigate whether the large variability in alpha power change and pupil dilation change across the participants as reported by McMahon et al. (2016) was due to the influence of performance.

In the SRT and vocoding model, the more spectrally degraded 6-channel vocoded sentences elicited greater pupil dilation compared with the 16-channel sentences. This finding is consistent with previous studies where decreasing the spectral resolution of the signal (i.e., decreasing the number of channels in vocoded speech) systematically increases pupil diameter (Winn et al., 2015), suggesting listening was more effortful.

Alpha power was greater in the less spectrally degraded 16-channel condition, consistent with McMahon et al. (2016) which used the same noise-vocoded sentences and four-talker babble background noise. This finding diverges from studies using less complex linguistic stimuli which have shown that decreased acoustic quality enhances alpha power (digit task: Obleser et al., 2012; Wöstmann et al., 2015; word comprehension: Becker, Pefkou, Michel, & Hervais-Adelman, 2013; Obleser & Weisz, 2012). This could suggest that spectrally degraded sentences influence the alpha network differently due to the increased linguistic complexity. In the easier 16-channel condition, because of the better spectral quality, there may be less dependency on the semantic context to recognize a sentence; whereas in the more spectrally degraded 6-channel condition, there is a greater need to rely on semantic context to fill in the gaps. However, before semantic processing is engaged, at least some phonemes must be recognized in the incoming speech signal. This lower level acoustic or phonemic processing may be more demanding in the 6-channel condition, perhaps decreasing the possibility for more automated semantic recognition. Greater alpha power in the 16-channel condition may therefore reflect task-irrelevant inhibition due to the automaticity of sentence processing when the signal was clearer, while reduced alpha power in the 6-channel condition may reflect ongoing active processing (Klimesch, 2012; Weisz, Hartmann, Müller, & Obleser, 2011) due to the poorer signal quality.

It was anticipated that the pupil would be sensitive to changes in SRT, with greater increases in pupil size expected in the more cognitively demanding 50% SRT compared with 80% SRT condition. Contrary to expectations, however, this was not the case. During the physiological session, the true performance levels obtained for each SRT varied substantially from the target SRTs of 50% and 80%. On average, the true performance difference was only 18.6% between conditions, compared with the target SRTs which differed by 30% (i.e., the performance difference between 50% and 80% SRT). It is therefore possible that the narrower range of the true performance contributed to the nonsignificant change in pupil dilation when modelling SRT.

Using SRTs is common practice to assess speech recognition in noise both clinically and for research purposes (Best, Keidser, Buchholz, & Freeston, 2015; Lunner, 2003; Smits & Festen, 2013). Numerous pupilometry studies have varied task difficulty by using an adaptive method to reach a desired SRT while recoding the pupil response. In the current study, each individual's SNR was fixed during the physiological session in order to randomize stimuli presentation, as per Zekveld, Heslenfeld et al. (2014). To do this, participants' SNRs for 50% and 80% SRT were obtained in a prior behavioral session (see Methods section). The adaptive test has been validated with similar speech materials for participants with normal hearing and hearing loss and showed less variability than the current study (Keidser et al., 2013). The higher variability in the current study may have resulted from spectrally degrading the sentence materials (i.e., noise vocoding) or the different application of the SRTs across test sessions. Further, given that the behavioral session was considerably shorter in duration than the physiological session, the discrepancy between true performance and the target SRTs across the two sessions may be influenced by the participants' varying levels of motivation and fatigue between sessions. Irrespective of the cause of the variability, true performance levels were used in an alternative LME regression model, instead of the 50% and 80% SRTs.

True performance levels had no effect on alpha power, further suggesting that potential between-subjects variability in performance when sentences are presented in fixed SNRs (McMahon et al., 2016) is not likely to be

influencing the change in alpha power. In line with previous studies (Kramer et al., 1997; Zekveld et al., 2010), increasing performance significantly decreased pupil size. Moreover, consistent with Winn et al. (2015), decreasing the spectral quality of the signal elicited greater pupil dilation, even when accounting for individuals' true performance, demonstrating that performance alone does not fully capture the effort required to process a spectrally degraded signal. This is particularly evident in studies showing that even when task accuracy is high (~100%), continuing to degrade the spectral quality of the signal increases effort, as reflected in greater pupil size (Winn et al., 2015).

To further assess the relationship between the two physiological measures and task accuracy, only partially and correctly recalled sentences were examined. As incorrect sentence recall (or an absent response) could be due to many factors, including attention and misperception (Kuchinsky et al., 2013), they were excluded from the analysis. Pupil size was significantly larger in the 6-channel condition, and also for partially recalled sentences, however there was no significant interaction between the two effects. Therefore, while decreased spectral quality and partially recalled sentences appear to increase listening effort, when sentences are accurately recalled (correct response), there is no difference in listening effort between vocoding conditions, as indexed by pupil size. Unlike the pupil response, there was no difference in alpha power across the levels of task accuracy. This is in line with Obleser and colleagues (2012) who found task accuracy and alpha power were not correlated, although task accuracy was relatively high in their study (91–100%) and responses were either correct or incorrect. Removing the incorrect responses in the current study ($n = 888$), revealed that channel vocoding no longer appeared to influence alpha power. However, this is possibly due to the decreased statistical power when removing incorrectly recalled sentences.

Consistent with previous studies examining speech recognition in noise, pupil size and alpha power were not significantly correlated within individuals (McMahon et al., 2016). This lack of correlation has been similarly reported in the reading domain (Scharinger, Kammerer, & Gerjets, 2015). As speculated by McMahon et al. (2016), this may be due to attention mechanisms, such as individuals using different encoding and modifying strategies (cf., Power and Petersen, 2013), the measures themselves being under the control of different attentional networks (cf. Corbetta, Patel, & Shulman, 2008, for review of the different attention networks), or that each are encoding different aspects of listening effort.

Population parameters, such as age, hearing, and cognitive ability, may interact with the pupil and alpha response, which may explain some of the variability which exists in the current literature. For example,

Petersen et al. (2015) found that alpha power breaks down in participants with moderate hearing loss when signal degradation and working memory capacity is at its most challenging level. Zekveld et al. (2011) have also demonstrated that the relationship between a smaller pupil response (which would reflect a decreased cognitive load) and increasing speech intelligibility was indeed weaker in people with hearing impairment. The pupil and alpha response when hearing is impeded either by internal acoustic degradation (such as a hearing loss) versus external acoustic degradation (such as presenting a degraded signal to individuals with or without hearing impairment) may therefore not be entirely comparable.

Limitations of the current study include confining the analyses of alpha power change to the parietal area. While majority of the studies outlined in this article found enhanced alpha activity in this location, it may be too restrictive given the added complexity of processing spectrally degraded sentences in noise. Furthermore, alpha band (8–12 Hz) may be too coarse, averaging two (or more) parallel processes. For example, 8 to 10 Hz is suggested to relate to attentional processes, while 10 to 12 Hz may be related to linguistic-type activity such as semantic processing (cf. Klimesch, 1999, 2012, for review), where each selectively synchronize or desynchronize based on the stimulus. Future studies may wish to consider whole-head analysis and narrow-band frequencies which may provide further insight into different aspects of listening effort.

Conclusion

Both changes in pupil dilation and alpha power have been suggested to index listening effort. Understanding how these measures behave, and interact, during simultaneous measurement provides insight into what aspects of listening effort they may each be encoding. This will go a long way toward disentangling the multifaceted nature of listening effort and will improve the prospect of using them to complement standard hearing assessments. Here, we began to address these issues in a young, normal-hearing population to better understand how they operate when sensory input is not compromised by hearing impairment, and cognition is robust. Both measures appeared to be sensitive to changes in spectral resolution, while pupil dilation provided further information about performance levels and task accuracy. Further, the measures were not correlated, suggesting they may be sensitive, or respond differently to, the different aspects of “mental exertion” that comprise listening effort.

Authors' Note

The Departments of Linguistics, Psychology, and Statistics are part of the Australian Hearing Hub, an initiative of Macquarie

University, that brings together Australia's leading hearing and health-care organizations to collaborate on research projects.

Acknowledgments

This study was supported by the Macquarie University Research Excellence Scheme, and the HEARing Cooperative Research Centre, established and supported under the Business Cooperative Research Centres Programme of the Australian Government. The authors thank Jorg Buchholz for assisting with protocol development, Nille Elise Kepp for help preparing the test materials, and Jaime Undurraga and Keiran Thompson for programming advice. The Departments of Linguistics, Psychology, and Statistics at Macquarie University as well as the HEARing Cooperative Research Centre are part of the Australian Hearing Hub, an initiative of Macquarie University, that brings together Australia's leading hearing and healthcare organizations to collaborate on research projects.

Declaration of Conflicting Interests

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

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The Australian Hearing Hub, an initiative of Macquarie University and the Australian Government, provided financial support for the publication of this research article and the HEARing Cooperative Research Centre, established and supported under the Business Cooperative Research Centres Programme of the Australian Government.

References

- Arlinger, S., Lunner, T., Lyxell, B., & Pichora-Fuller, K. M. (2009). The emergence of cognitive hearing science. *Scandinavian Journal of Psychology, 50*(5), 371–384.
- Badcock, N. A., Mousikou, P., Mahajan, Y., de Lissa, P., Thie, J., McArthur, G. (2013). Validation of the Emotiv EPOC® EEG gaming system for measuring research quality auditory ERPs. *PeerJ, 1*, e38.
- Başar, E., Başar-Eroglu, C., Karakaş, S., & Schürmann, M. (2001). Gamma, alpha, delta, and theta oscillations govern cognitive processes. *International Journal of Psychophysiology, 39*(2), 241–248.
- Beatty, J., & Wagoner, B. L. (1978). Pupillometric signs of brain activation vary with level of cognitive processing. *Science, 199*, 1216–1218.
- Becker, R., Pefkou, M., Michel, C. M., & Hervais-Adelman, A. G. (2013). Left temporal alpha-band activity reflects single word intelligibility. *Frontiers in Systems Neuroscience, 7*, 121.
- Bernarding, C., Strauss, D. J., Hannemann, R., Seidler, H., & Corona-Strauss, F. I. (2013). Neural correlates of listening effort related factors: Influence of age and hearing impairment. *Brain Research Bulletin, 91*, 21–30.
- Best, V., Keidser, G., Buchholz, J. M., & Freeston, K. (2015). An examination of speech reception thresholds measured in a simulated reverberant cafeteria environment. *International Journal of Audiology, 54*(10), 682–690.
- Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron, 58*(3), 306–324.
- Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *The Journal of the Acoustical Society of America, 110*(2), 1150–1163.
- Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience, 10*(2), 252–269.
- Gosselin, P. A., & Gagne, J.-P. (2011). Older adults expend more listening effort than young adults recognizing speech in noise. *Journal of Speech, Language, and Hearing Research, 54*(3), 944–958.
- Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology, 33*, 457–461.
- Hawkins, D. B., & Yacullo, W. S. (1984). Signal-to-noise ratio advantage of binaural hearing aids and directional microphones under different levels of reverberation. *Journal of Speech and Hearing Disorders, 49*(3), 278–286.
- Herrmann, C. S., Fründ, I., & Lenz, D. (2010). Human gamma-band activity: A review on cognitive and behavioral correlates and network models. *Neuroscience & Biobehavioral Reviews, 34*(7), 981–992.
- Hornsby, B. W., Werfel, K., Camarata, S., & Bess, F. H. (2014). Subjective fatigue in children with hearing loss: Some preliminary findings. *American Journal of Audiology, 23*(1), 129–134.
- Hua, H., Anderzén-Carlsson, A., Widén, S., Möller, C., & Lyxell, B. (2015). Conceptions of working life among employees with mild-moderate aided hearing impairment: A phenomenographic study. *International Journal of Audiology, 54*(11), 873–880.
- Hua, H., Emilsson, M., Ellis, R., Widén, S., Möller, C., Lyxell, B. (2014). Cognitive skills and the effect of noise on perceived effort in employees with aided hearing impairment and normal hearing. *Noise and Health, 16*(69), 79–88.
- Karrasch, M., Laine, M., Rapinoja, P., & Krause, C. M. (2004). Effects of normal aging on event-related desynchronization/synchronization during a memory task in humans. *Neuroscience letters, 366*(1), 18–23.
- Keidser, G., Dillon, H., Mejia, J., & Nguyen, C.-V. (2013). An algorithm that administers adaptive speech-in-noise testing to a specified reliability at selectable points on the psychometric function. *International Journal of Audiology, 52*(11), 795–800.
- Klimesch, W. (1996). Memory processes, brain oscillations and EEG synchronization. *International Journal of Psychophysiology, 24*(1), 61–100.
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews, 29*(2), 169–195.
- Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences, 16*(12), 606–617.

- Koelwijn, 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), 291–300.
- Koelwijn, 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.
- Koelwijn, T., Zekveld, A. A., Festen, J. M., Rönnerberg, J., & Kramer, S. E. (2012). Processing load induced by informational masking is related to linguistic abilities. *International Journal of Otolaryngology, 2012*, 1–11.
- Kramer, S. E., Kapteyn, T. S., Festen, J. M., & Kuik, D. J. (1997). Assessing aspects of auditory handicap by means of pupil dilatation. *International Journal of Audiology, 36*(3), 155–164.
- Kramer, S. E., Kapteyn, T. S., & Houtgast, T. (2006). Occupational performance: Comparing normally-hearing and hearing-impaired employees using the Amsterdam Checklist for Hearing and Work. *International Journal of Audiology, 45*(9), 503–512.
- Kuchinsky, S. E., Ahlstrom, J. B., Vaden, K. I., Cude, S. L., Humes, L. E., Dubno, J. R., ... Eckert, M. A. (2013). Pupil size varies with word listening and response selection difficulty in older adults with hearing loss. *Psychophysiology, 50*(1), 23–34.
- Kuper, H., Singh-Manoux, A., Siegrist, J., & Marmot, M. (2002). When reciprocity fails: Effort–reward imbalance in relation to coronary heart disease and health functioning within the Whitehall II study. *Occupational and Environmental Medicine, 59*(11), 777–784.
- Larsby, B., Hällgren, M., Lyxell, B., & Arlinger, S. (2005). Cognitive performance and perceived effort in speech processing tasks: Effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. *International Journal of Audiology, 44*(3), 131–143.
- Leiberg, S., Lutzenberger, W., & Kaiser, J. (2006). Effects of memory load on cortical oscillatory activity during auditory pattern working memory. *Brain research, 1120*(1), 131–140.
- Lunner, T. (2003). Cognitive function in relation to hearing aid use. *International Journal of Audiology, 42*, Suppl1, 49–58.
- McMahon, C. M., Boisvert, I., de Lissa, P., Granger, L., Ibrahim, R., Lo, C. Y., ... Graham, P. L. (2016). Monitoring alpha oscillations and pupil dilation across the performance-intensity function. *Frontiers in Psychology, 7*, 745.
- Mehta, R. K., & Agnew, M. J. (2012). Influence of mental workload on muscle endurance, fatigue, and recovery during intermittent static work. *European Journal of Applied Physiology, 112*(8), 2891–2902.
- Mele, M. L., & Federici, S. (2012). Gaze and eye-tracking solutions for psychological research. *Cognitive Processing, 13*(1), 261–265.
- Nachtegaal, J., Kuik, D. J., Anema, J. R., Goverts, S. T., Festen, J. M., Kramer, S. E. (2009). Hearing status, need for recovery after work, and psychosocial work characteristics: Results from an internet-based national survey on hearing. *International Journal of Audiology, 48*(10), 684–691.
- Obleser, J., Wöstmann, M., Hellbernd, N., Wilsch, A., & Maess, B. (2012). Adverse listening conditions and memory load drive a common alpha oscillatory network. *The Journal of Neuroscience, 32*(36), 12376–12383.
- Obleser, J., & Weisz, N. (2012). Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cerebral cortex, 22*(11), 2466–2477.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.
- Pesonen, M., Björnberg, C. H., Hämäläinen, H., & Krause, C. M. (2006). Brain oscillatory 1–30Hz EEG ERD/ERS responses during the different stages of an auditory memory search task. *Neuroscience letters, 399*(1), 45–50.
- Peters, M. L., Godaert, G. L., Ballieux, R. E., van Vliet, M., Willemsen, J. J., Sweep, F. C., ... Heijnen, C. J. (1998). Cardiovascular and endocrine responses to experimental stress: Effects of mental effort and controllability. *Psychoneuroendocrinology, 23*(1), 1–17.
- Petersen, E. B., Wöstmann, M., Obleser, J., Stenfelt, S., & Lunner, T. (2015). Hearing loss impacts neural alpha oscillations under adverse listening conditions. *Frontiers in Psychology, 6*, 177.
- Pichora-Fuller, K. (2003). Cognitive aging and auditory information processing. *International Journal of Audiology, 42*(Suppl 2), 26–32.
- Pichora-Fuller, M. K. (2006). Perceptual effort and apparent cognitive decline: Implications for audiologic rehabilitation. *Seminars in Hearing, 27*(4), 284–293.
- Pichora-Fuller, K., & Singh, G. (2006). Effects of age on auditory and cognitive processing: Implications for hearing aid fitting and audiologic rehabilitation. *Trends in Amplification, 10*(1), 29–59.
- Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., ... Naylor, G. (2016). Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing, 37*, 5S–27S.
- Pichora-Fuller, M. K. (2006). Perceptual effort and apparent cognitive decline: implications for audiologic rehabilitation. *Seminars in Hearing, 27*(4), 284–293.
- Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2014). R Core Team (2014) nlme: Linear and nonlinear mixed effects models. R package version 3.1-117. Retrieved from <http://CRAN.R-project.org/package=nlme>.
- Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology, 47*(3), 560–569.
- Power, J. D., & Petersen, S. E. (2013). Control-related systems in the human brain. *Current Opinion in Neurobiology, 23*, 223–228.
- Rabbitt, P. M. (1968). Channel-capacity, intelligibility and immediate memory. *The Quarterly Journal of Experimental Psychology, 20*(3), 241–248.
- Rönnerberg, J., Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology, 47*(Suppl 2), 99–105.
- Rudner, M., Lunner, T., Behrens, T., Thorén, E. S., & Rönnerberg, J. (2012). Working memory capacity may

- influence perceived effort during aided speech recognition in noise. *Journal of the American Academy of Audiology*, 23(8), 577–589.
- Scharinger, C., Kammerer, Y., & Gerjets, P. (2015). Pupil dilation and EEG alpha frequency band power reveal load on executive functions for link-selection processes during text reading. *PLoS One*, 10(6), e0130608.
- Schneider, B. A., Pichora-Fuller, K., & Daneman, M. (2010). Effects of senescent changes in audition and cognition on spoken language comprehension. In S. Gordon-Salant, D. R. Frisina, A. N. Popper, & R. R. Fay (Eds.), *Springer handbook of auditory research: The aging auditory system* (vol. 34, pp. 167–210). New York, NY: Springer.
- Siegle, G. J., Ichikawa, N., & Steinhauer, S. (2008). Blink before and after you think: Blinks occur prior to and following cognitive load indexed by pupillary responses. *Psychophysiology*, 45(5), 679–687.
- Siegrist, J. (1996). Adverse health effects of high-effort/low-reward conditions. *Journal of Occupational Health Psychology*, 1(1), 27–41.
- Smits, C., & Festen, J. M. (2013). The interpretation of speech reception threshold data in normal-hearing and hearing-impaired listeners: II. *Fluctuating noise*. *The Journal of the Acoustical Society of America*, 133(5), 3004–3015.
- Souza, P., & Arehart, K. (2015). Robust relationship between reading span and speech recognition in noise. *International Journal of Audiology*, 54(10), 705–713.
- Strauß, A., Wöstmann, M., & Obleser, J. (2014). Cortical alpha oscillations as a tool for auditory selective inhibition. *Frontiers in Human Neuroscience*, 8, 350.
- Tun, P. A., McCoy, S., & Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and Aging*, 24(3), 761.
- Ward, L. M. (2003). Synchronous neural oscillations and cognitive processes. *Trends in Cognitive Sciences*, 7(12), 553–559.
- Weinstein, B. E., & Ventry, I. M. (1982). Hearing impairment and social isolation in the elderly. *Journal of Speech, Language, and Hearing Research*, 25(4), 593–599.
- Weisz, N., Hartmann, T., Müller, N., & Obleser, J. (2011). Alpha rhythms in audition: cognitive and clinical perspectives. *Frontiers in Psychology*, 2, 73.
- Wendt, D., Dau, T., & Hjortkjær, J. (2016). Impact of background noise and sentence complexity on processing demands during sentence comprehension. *Frontiers in Psychology*, 7, 345.
- Wingfield, A., Tun, P. A., & McCoy, S. L. (2005). Hearing loss in older adulthood what it is and how it interacts with cognitive performance. *Current Directions in Psychological Science*, 14(3), 144–148.
- Winn, M. B., Edwards, J. R., & Litovsky, R. Y. (2015). The impact of auditory spectral resolution on listening effort revealed by pupil dilation. *Ear and Hearing*, 36(4), 153–165.
- Wolak, M. E., Fairbairn, D. J., & Paulsen, Y. R. (2012). Guidelines for estimating repeatability. *Methods in Ecology and Evolution*, 3(1), 129–137.
- Wöstmann, M., Herrmann, B., Wilsch, A., & Obleser, J. (2015). Neural alpha dynamics in younger and older listeners reflect acoustic challenges and predictive benefits. *The Journal of Neuroscience*, 35(4), 1458–1467.
- Wouters, J., & Berghe, J. V. (2001). Speech recognition in noise for cochlear implantees with a two-microphone monaural adaptive noise reduction system. *Ear and Hearing*, 22(5), 420–430.
- Zekveld, A. A., Heslenfeld, D. J., Festen, J. M., & Schoonhoven, R. (2006). Top-down and bottom-up processes in speech comprehension. *Neuroimage*, 32(4), 1826–1836.
- Zekveld, A. A., Heslenfeld, D. J., Johnsrude, I. S., Versfeld, N. J., & Kramer, S. E. (2014). The eye as a window to the listening brain: Neural correlates of pupil size as a measure of cognitive listening load. *Neuroimage*, 101, 76–86.
- Zekveld, A. A., & Kramer, S. E. (2014). Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology*, 51(3), 277–284.
- Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. *Ear and Hearing*, 31(4), 480–490.
- Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2011). Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response. *Ear and Hearing*, 32(4), 498–510.
- Zekveld, A. A., Kramer, S. E., Kessens, J. M., Vlaming, M. S., & Houtgast, T. (2009). The influence of age, hearing, and working memory on the speech comprehension benefit derived from an automatic speech recognition system. *Ear and Hearing*, 30(2), 262–272.
- Zekveld, A. A., Rudner, M., Kramer, S. E., Lyzenga, J., & Rönnerberg, J. (2014). Cognitive processing load during listening is reduced more by decreasing voice similarity than by increasing spatial separation between target and masker speech. *Frontiers in Neuroscience*, 8, 88.