Special Issue Paper

QJEP

Quarterly Journal of Experimental Psychology
1–18

DOI: 10.1177/17470218251316797
qjep.sagepub.com
Sage

The time course of cognitive effort during disrupted speech

Andrew M Burleson[1,2] and Pamela E Souza[1]

Abstract

Listeners often find themselves in scenarios where speech is disrupted, misperceived, or otherwise difficult to recognise. In these situations, many individuals report exerting additional effort to understand speech, even when repairing speech may be difficult or impossible. This investigation aimed to characterise cognitive efforts across time during both sentence listening and a post-sentence retention interval by observing the pupillary response of participants with normal to borderline-normal hearing in response to two interrupted speech conditions: sentences interrupted by gaps of silence or bursts of noise. The pupillary response serves as a measure of the cumulative resources devoted to task completion. Both interruption conditions resulted in significantly greater levels of pupil dilation than the uninterrupted speech condition. Just prior to the end of a sentence, trials periodically interrupted by bursts of noise elicited greater pupil dilation than the silent-interrupted condition. Compared to the uninterrupted condition, both interruption conditions resulted in increased dilation after sentence end but before repetition, possibly reflecting sustained processing demands. Understanding pupil dilation as a marker of cognitive effort is important for clinicians and researchers when assessing the additional effort exerted by listeners with hearing loss who may use cochlear implants or hearing aids. Even when successful perceptual repair is unlikely, listeners may continue to exert increased effort when processing misperceived speech, which could cause them to miss upcoming speech or may contribute to heightened listening fatigue.

Keywords

Cognitive effort; listening effort; interrupted speech; pupillometry

Received: 29 January 2024; revised: 8 January 2025; accepted: 10 January 2025

Introduction

Listeners frequently find themselves in situations where interruptions—due to background noise, competing talkers, or environmental sounds—make speech difficult to understand. This challenge is particularly important for listeners with hearing loss (Başkent, 2010; Shafiro et al., 2015, 2016) who expend considerable effort processing disrupted or fragmented speech (Ohlenforst, Zekveld, Jansma, et al., 2017; Zekveld et al., 2018). As speech disruptions and misperceptions accumulate, processing demands rise to compensate for inaudible or misperceived speech. Yet, even after expending additional cognitive resources, listeners often fail to understand interrupted speech. Whether or not perceptual repair is successful, it still requires extra resources from the listener (Winn & Teece, 2021, 2022), which can affect the perception of upcoming speech (Winn, 2024). These difficulties can negatively impact a listener's quality of life, leading to social withdrawal (Demorest & Erdman, 1987; Edwards, 2007), professional limitations (Danermark & Gellerstedt, 2004; Kramer et al., 1997; Nachtegaal et al., 2009), and even cortical reorganisation and cognitive change (Dietrich et al., 2001; Glick & Sharma, 2020). Thus, understanding the cognitive resources required when processing degraded or unclear speech—where repair is uncertain or incomplete—is an important facet of understanding communication impairments.

In some circumstances, listeners may compensate for incomplete signals by relying on contextual cues or

[1]Hearing Aid Laboratory, Department of Communication Sciences and Disorders Evanston, Northwestern University, Evanston, IL, USA
[2]Emerging Auditory Research laboratory, Massachusetts Eye and Ear, Harvard Medical School, Boston, MA, USA

Corresponding author:
Andrew M Burleson, Emerging Auditory Research laboratory, Massachusetts Eye and Ear, Harvard Medical School, 243 Charles Street, Boston, MA, 02114, USA.
Email: ABurleson@meei.harvard.edu

cognitive and linguistic strategies to fill in missing speech, a process known as perceptual restoration (Warren, 1970). The success of this process depends on several signal- and listener-related factors that influence the recovery and repair of missing speech (Warren, 1970). For example, performance improves with better spectrotemporal fidelity (Clarke et al., 2016; Fogerty et al., 2015), shorter or fewer interruptions (Bologna et al., 2018; Kidd & Humes, 2012; Miller & Licklider, 1950; Shafiro et al., 2015; Wang & Humes, 2010), and noise-filled gaps over silent ones (Bashford et al., 1996; Bashford & Warren, 1987; Bologna et al., 2019; Nagaraj & Magimairaj, 2017; Powers & Wilcox, 1977). In addition, listeners with higher working memory capacity and better lexical knowledge restore more missing speech (Benard & Başkent, 2014; Burleson & Souza, 2022; Millman & Mattys, 2017; Nagaraj & Magimairaj, 2017).

However, these investigations also reveal that compensatory mechanisms do not always result in successful perceptual restoration. When the auditory signal is too degraded, interruptions are too long or frequent, or the listener cannot determine a clear path to repair, they may struggle to piece together the remaining fragments. In such cases, listeners may only grasp the "gist" of the spoken content at best, while at worst, they may recognise few, if any, words, leading to an inefficient use of cognitive resources (Rönnberg et al., 2013, 2021). Finally, the cognitive resources expended during difficult listening tasks can also vary widely, even when intelligibility scores appear similar across conditions (Ohlenforst, Zekveld, Lunner, et al., 2017), highlighting the role of resource deployment during even unsuccessful perceptual repair.

Rather than focusing on speech repair alone, recent findings emphasise the need to investigate how listeners manage higher-order resource devotion, or effort, when complete perceptual restoration is not feasible. This reframing aligns with models highlighting the neurobiological processes of effort and effort withdrawal in response to a consistent listening situation, such as a noisy restaurant. In these circumstances, listeners weigh the costs and benefits of continued devotion of effort as the likelihood of successful restoration diminishes, ultimately deciding whether or not to withdraw resources based on the perceived value or reward in these conditions (e.g., enjoying participation in a conversation) (Eckert et al., 2016). This investigation explores how listeners navigate cognitive effort during attempts to repair interrupted speech, even when they may not restore missing speech information. Examining these processes provides insight into how individuals cope with incomplete auditory input and the limits of perceptual restoration.

The cognitive resources required to process degraded speech are often overlooked in clinical settings, even for listeners with normal hearing. Listeners with clinically normal or near-normal hearing still struggle with interrupted speech (Bologna et al., 2018; Fogerty et al., 2015; Miller & Licklider, 1950; Powers & Wilcox, 1977; Warren, 1970), a difficulty that becomes particularly evident when the missing speech is so degraded that it may not be restored. Speech-recognition difficulty for patients with clinically normal hearing creates a challenging situation for both clinicians and patients, especially in noisy environments (Liberman & Kujawa, 2017; Tremblay et al., 2015). Clinical audiologists often recommend modifying the environment (e.g., reducing distance from the speaker) and adjusting conversation strategies (e.g., asking for rephrased clarification) to improve speech understanding. However, these recommendations focus on signal modification rather than the additional effort required for processing disrupted speech. Formal assessments of the cognitive effort required for processing degraded speech are uncommon in clinical practice. An exception is the use of self-report measures to assess effortful processing, such as the Speech and Spatial Qualities questionnaire (Gatehouse & Noble, 2004) or the Listening Effort Questionnaire (Hughes et al., 2019). However, such questionnaires are rarely used, and the cognitive demands of processing disrupted or misperceived speech often go unrecognised, depriving clinicians of valuable insight into their patients' listening challenges.

Processing demands require the listener to exert additional cognitive effort, which refers to listener-directed energy related to task demands, attentional engagement, and other listener-related factors such as fatigue (Kahneman, 1973; Pichora-Fuller et al., 2016). While the term "listening effort" is common, we use "cognitive effort" here to encompass task-related engagement beyond sentence listening *per se*. More challenging tasks require more cognitive resources during task completion. Cognitive effort expenditure is thought to be nonlinear, with both easy and very challenging tasks resulting in minimal effort deployed (Brehm & Self, 1989; Richter et al., 2016). The former is due to the ease of the task; the latter is likely due to disengagement, effort withdrawal, or "giving up" (Eckert et al., 2016).

Pupillometry offers an indirect method for quantifying cognitive effort, providing a continuous, real-time measure via pupil dilation. This dilation is thought to mirror the cognitive resources devoted to the task, with increased dilation indicating greater effort when more cognitive resources are engaged (Ahern & Beatty, 1979; Beatty, 1982; Vogelzang et al., 2016). It has been suggested that the link between effort exertion and pupil dilation is rooted in the autonomic nervous system and error monitoring in the cingulo-opercular network (Eckert et al., 2016; Vaden et al., 2013, 2017), physiologically linking this increase in dilation to the region known as the locus coeruleus (Aston-Jones & Cohen, 2005; Murphy et al., 2011) and broader cortical activity (Reimer et al., 2016).

Research also suggests that cognitive effort is devoted both during and after speech listening (Nees, 2016; Piquado et al., 2010; Winn et al., 2015; Winn & Moore, 2018). Measuring cognitive effort during two important time periods—sentence listening and the retention interval immediately afterward—offers insight into how listeners manage challenging conditions, even when speech repair is unlikely or impossible. This approach shifts the focus from the resources required for successful perceptual repair to understanding the sustained allocation of cognitive effort during interrupted speech, where complete restoration of the speech signal is unlikely or impossible.

Cognitive effort tends to increase as listeners begin listening to sentences and extends through the listening window, especially for difficult tasks, and the role of cognitive effort has been established for a range of auditory stimuli (Zekveld et al., 2018). For example, differences in effort emerge during listening tasks with changes in masker type (Koelewijn, Zekveld, Festen, & Kramer, 2012; Wendt et al., 2018), intelligibility (Porretta & Tucker, 2019; Zekveld et al., 2010), and hearing loss (Zekveld et al., 2011). Furthermore, as signal fidelity worsens, the cognitive effort deployed to process speech systematically increases (Winn et al., 2015). In addition to intelligibility, audibility, and degradation, linguistic factors, such as syntactic ambiguity and lexical frequency, can modulate the cognitive effort required to process speech (Engelhardt et al., 2010; Kuchinke et al., 2007). Cognitive resource use increases during sentence listening, especially under difficult conditions—even when repair may not be feasible and whether or not it is successful (Winn & Teece, 2021, 2022).

When speech is degraded, listeners continue to process the stimuli even after the entire sentence has concluded, a period known as the retention interval. The retention interval, occurring after the stimulus ends but before the participant responds, is considered a window for potential ongoing explicit reprocessing and repair (Winn, 2024; Winn & Teece, 2021, 2022). This sustained cognitive effort occurs when the processing requirements of a task remain incomplete. That is, processing an easy task (e.g., uninterrupted speech) can be completed almost entirely during sentence presentation, whereas a challenging task (e.g., interrupted speech) demands ongoing effort after the sentence has ended. For example, cognitive effort returns rapidly to baseline after highly intelligible sentences but remains elevated for unintelligible ones (Zekveld & Kramer, 2014). Similarly, for vocoded sentences, the pupillary response returns to baseline more quickly for correct responses than when they are incorrect (Winn et al., 2015), and there is prolonged dilation for the simultaneous presentation of two sentences (Koelewijn et al., 2015). Alternatively, the ongoing effort devoted to repair during the retention interval may be withdrawn should the task be deemed too difficult, or the perceived benefit be too low (Eckert et al., 2016). Together, these findings suggest that listeners may continue to process speech during the retention interval. However, it remains unclear to what extent cognitive effort is sustained during the retention interval when speech is disrupted and a path to full perceptual repair is unlikely or impossible.

This investigation aimed to explore the level of cognitive effort for two interruption conditions compared to an uninterrupted condition: sentences interrupted by periodic gaps of silence or bursts of speech-shaped noise, across two distinct time windows. The first window was during sentence listening, where data suggest increasing signal-degradation leads to increased cognitive effort (Winn et al., 2015). The second was the retention interval, during which listeners may attempt to repair the disrupted sentence (Gianakas et al., 2022; Winn et al., 2015; Winn & Moore, 2018). We predicted that the level of cognitive effort for both interruption conditions would be similar during the listening window, as the degree of signal degradation for the two conditions was comparable, and pupil dilation has been shown to be related to signal degradation (Winn et al., 2015). During the retention interval, we predicted that the level of cognitive effort would be greater for both interruption conditions compared to the uninterrupted condition, as listeners remained devoted to perceptual repair.

Method

Participants

This study included 17 participants, aged 22–64 years (average age = 42.1 years, *SD* = 14.1 years), who had normal or borderline normal hearing. A power analysis conducted with an assumed Cohen's effect size ($f^2 = 0.3$) with a power of 0.8 indicated that a sample size of 20 would be adequately powered (see the Discussion section for further consideration of sample size). To be eligible, participants needed to speak English as their primary language, be at least 18 years old, have symmetrical hearing thresholds (no difference greater than 10 dB HL at any frequency) without conductive loss (air-bone gap < 10 dB HL at any frequency), and have up to a mild hearing loss (defined as no single audiometric threshold exceeding 40 dB HL from 250 to 6,000 Hz; Figure 1). We excluded participants with moderate or poorer hearing loss because our study focused on the cognitive effort required for compensating an audible speech signal, rather than examining how listeners would respond to poor audibility of the signal itself (see the Discussion section for further consideration of hearing loss and perceptual restoration).

Participants performed better than 20/50 on a Snellen vision chart without correction, indicating visual acuity appropriate for this pupillometry experiment. Participants were not permitted to wear hard contact lenses or glasses during pupillometry as the reflection from the lenses can distort the data. None of the participants had a history of

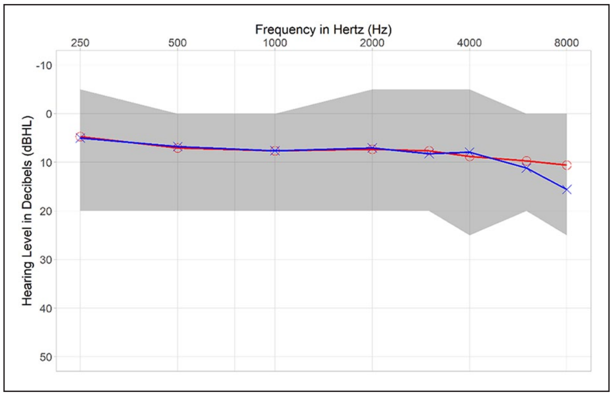

Figure 1. Audiometric thresholds for the 17 participants. Left ear thresholds are represented by the blue X, and right ear thresholds are represented by the red circles. The shaded area reflects the range of audiometric thresholds for this cohort.

eye or vision problems such as cataracts or colorblindness. All participants denied having dyslexia, diabetes mellitus, or a history of neurological disease. No participants wore heavy eye makeup during testing. The institutional review board of Northwestern University approved the study, and all participants participated in a written informed consent process and had adequate opportunity to ask questions prior to participation. Participants were compensated at an hourly rate for taking part in the study.

Stimuli

Sentence sets were composed of the Perceptually Robust English Sentence Test: Open Set (PRESTO), a sentence corpus designed to detect individual variation in neurocognitive abilities during speech recognition (J. L. Gilbert et al., 2013). The PRESTO sentence sets were designed to be balanced for talker gender, number of key words, average log key word frequency, and average word familiarity. Example sentences include, "*The* ***nearest synagogue*** *may not be within* ***walking distance***," and "*The* ***jaw operates*** *by* ***using antagonistic muscles***" (scored key words in bold). The original files were retrieved from a CD, stored on a computer hard drive as a .wav file, and had a sampling rate of 44.1 kHz with identical left and right channels. Sentences similar in length were selected from all sentences in the corpus. The average length of all selected sentences was 3,526 ms (*SD* = 526 ms). The sentences underwent conversion from digital to analogue using the on-board sound card. They were then presented using a Focusrite Scarlett 2i2 pre-amp and output through a single PreSonus Eris e3.5 speaker in front of the listener. The output was maintained at a 72-dB A-weighted average root mean square (RMS; LAeq) level, with peak A-weighted levels (LAFmax) measured at 82 dB at 1 m, or at the listener's seated position. Level measures were obtained using Brüel & Kjær Type 2250 sound-level metre at the level of the eye tracker's fixed chin rest (see subheading titled "*Eye Tracker*").

All sentences. Sentences in this investigation consisted of three listening conditions of 25 sentences each (75 total). Participants listened to uninterrupted sentences, sentences interrupted by gaps of silence, and sentences interrupted by bursts of speech-shaped noise.

Silent-gap sentences. Interrupted sentence processing was completed using a custom R2020a MATLAB script. Interrupted sentences were gated using a 50% duty cycle square wave at a 2.0-Hz rate, resulting in alternating sections of 250-ms speech and 250 ms of a silent interruption. To reduce spectral splatter, 10 ms cosine on- and off-ramps were applied to all interruption sites. All sentences were RMS normalised after silent-gap insertion. The aforementioned procedure resulted in 50 sentences of alternating 250 ms of speech and 250 ms of silence, 25 of which were used for the silent-gap-interrupted condition, and no further processing was required.

Noise-burst sentences. The remaining 25 silent-gap-interrupted sentences then had speech-shaped noise inserted into the gap. To generate the speech-shaped noise, the full set of 75 sentences were combined, and a Fourier transform was obtained, randomising all spectral components and phases. Next, they were converted back into the time domain using an inverse Fourier transform. Segments of the speech-shaped noise were interrupted in an inverse manner compared to the speech sentences described earlier, resulting in alternating segments of 250 ms of silence, followed by 250 ms of speech-shaped noise. The noise-burst stimuli were then linearly added to the remaining 25 pre-processed sentences, interleaving the speech and noise segments (creating alternating segments of 250 ms of speech and 250 ms of noise) with 10 ms cosine on- and off-ramps. The noise-burst-interrupted sentences were RMS normalised prior to noise-burst insertion (see the "Silent-gap sentences" section) so that the level of the speech itself was equivalent in the silent-gap- and noise-burst-interrupted sentences. The noise burst was 10 dB louder than the speech segments (−10 dB SNR), as smaller SNRs result in poorer perceptual restoration performance (Bashford et al., 1996), resulting in a slightly higher overall level (approximately 7 dB) for the noise-interrupted sentences.

Eye tracker

Data were collected using an SR Research EyeLink 1000 Plus eye tracker hardware with a customised experimental paradigm designed in Experiment Builder (SR Research Ltd., version 2.3.1). The EyeLink 1000, an infrared video-based eye-tracking device, captured pupil dilation in the right eye of the participant as listeners viewed a monitor occupying most of their visual field. Listeners were seated in a low-reverberation, dimly-lit research room. The luminance of the room was the same for all participants and monitored to minimise pupil dilation that naturally occurs with changes in luminance, consistent with other pupillometry test settings (Wagner et al., 2016). The luminance of the display monitor was not higher than the room lighting to further minimise the effect of luminance on pupil diameter (Winn et al., 2018).

Cognitive and linguistic measures

Some literature suggests working memory capacity and lexical knowledge to be the most relevant abilities for perceptual restoration (Benard & Başkent, 2014; Burleson & Souza, 2022; Millman & Mattys, 2017; Nagaraj & Magimairaj, 2017). All participants in this investigation were part of a larger cohort who had previously completed a battery of cognitive and linguistic measures. Because these data were not used as individual predictors of pupil dilation in this investigation (see "*Statistical Considerations*") and because the primary aim of the present report was the time course of cognitive effort during interrupted speech within the context of cognitive and linguistic abilities (Engelhardt et al., 2010; Kuchinke et al., 2007), these data will be presented in a descriptive manner to illustrate the range of cognitive and linguistic abilities of this cohort.

The Reading Span Task (RST; Rönnberg et al., 2013) assesses a participant's complex working memory—the ability to concurrently reprocess, store, and accurately respond to complex stimuli (Baddeley, 2000; Daneman & Carpenter, 1980). Participants were first asked to comprehend a series of visually presented sentences and determine if the sentence makes sense (with half being incoherent). Then, participants were randomly asked to recall *either* the first or last word of each sentence in any order, and the number of words correctly identified was scored. Sentences presented to participants increased from two to five per sequence, and higher scores indicate better performance.

The Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II), Vocabulary Subtest assesses a participant's lexical knowledge of word meaning, accessing receptive vocabulary knowledge, and is common in neuropsychological assessment (Wechsler, 2009). Single words were presented, and participants were asked to define or describe the word. Following the administration manual, scoring ranged from zero to two points for each response, and the task was discontinued after three consecutive scores of zero. The task gets more difficult as the words become more abstract and less common. Raw scores were transformed to *t*-scores using the administration manual, and higher scores indicate greater lexical knowledge.

QuickSIN

The Quick Speech-in-Noise (QuickSIN) test was administered to assess listeners' ability to recognise speech in background noise (Killion et al., 2004). The QuickSIN test presents sentences in four-talker babble noise, with the signal-to-noise ratio (SNR) decreasing in 5-dB steps from 25 dB to 0 dB. Each list consists of six sentences, and the participant is asked to repeat each sentence they hear. The SNR loss was calculated by subtracting the number of correctly identified key words from a constant value of 25.5. Two lists were presented to each listener at 70 dBA, and the average SNR loss of both lists represented their score. This measure provides a quantifiable assessment of the participant's speech-recognition performance in noisy environments, reflecting real-world listening challenges, and is related to perceptual restoration performance (Vijayasarathy et al., 2021).

Procedure

All participants in this investigation completed audiometric evaluation, including pure tone air- and bone-conduction thresholds, QuickSIN, the RST, and the WASI-II Vocabulary Subtest (Killion et al., 2004; Rönnberg et al., 2013; Wechsler, 2009). Participants were brought to the room with the eye tracker where they completed the informed consent process with the investigator. Next, participants were seated in front of the eye tracker, and the level of the fixed chin rest and chair height were adjusted until the participant felt comfortable. The eye tracker was then focused manually and calibrated using a 9-point calibration procedure prior to task administration. Verbal responses were recorded using a tabletop microphone (MXL ProCon Series 1 AC-404USB). Participants were provided with written and verbal instructions for the task. Participants were encouraged to take breaks as needed in addition to scheduled breaks to minimise the effects of fatigue.

Five unscored practice trials were administered prior to the experimental task, and participants were encouraged to ask any clarification questions during these trials. During the task, participants were instructed to listen to the sentence but to wait 5 s (i.e., the retention interval) before the screen prompted, "repeat sentence." Listeners were encouraged to try to fill in missing speech, and to guess if unsure. Key words were scored for exact matches (i.e., plurals and tenses had to match). After the five practice trials, listeners moved on to the uninterrupted condition. The uninterrupted sentences are perceptually simpler, and this task was completed first to allow participants to acclimate to the test environment prior to the more difficult interrupted sentences. Subsequently, the order of the silent-gap- and noise-burst-interrupted sentences (the comparison of primary interest, grouped by condition) was randomised across participants. Because blinks may be identified, removed, and interpolated prior to analysis (Klingner et al., 2011) no explicit instruction was given about not blinking. However, participants were encouraged to blink at the end of each trial, similar to an investigation from Wagner et al. (2016). Testing duration was 1 to 1.5 hr.

Each trial consisted of a baseline phase, a listening phase, a retention interval, a repetition phase, and a return to baseline phase. Every trial began with a 1-s blank screen followed by 4 s of a black fixation cross on a grey screen. Next, a 0.5-s, 1,000-Hz alert beep was presented 1 s prior to sentence onset so the participant could anticipate when to listen (Winn & Moore, 2018). After the sentence was played, participants waited 5 s (i.e., the retention interval) before the screen displayed, "repeat sentence." The experimenter scored key words during repetition. Finally, a 6-s delay allowed the participant's pupil to return to baseline. Afterwards, the screen read, "break, blink as needed," and participants were encouraged to take a break during this time. The experimenter then manually advanced to the next trial. At the end of each condition, there was a longer, mandatory break to minimise participant fatigue. The fixation cross was always present on the screen except during the blank screen, task instruction, and calibration. Between each trial, the eye-tracking software corrected for drift, and the eye tracker was re-calibrated using the 9-point process described earlier if needed (e.g., if the participant lifted their head from the chin rest).

Statistical considerations

Summary statistics were calculated for participants, including key word percent correct and results of cognitive and linguistic measures. After the pupil-dilation analysis (see the "Pupil diameter analysis" section), pupillometry data were further stratified by key word percent correct for later visualisation. Given the number of possible predictors (e.g., RST and WASI-II Vocabulary) and inherent individual variability in pupil dilation (Winn et al., 2018) relative to the number of participants, linear regression analyses investigating the predictive value of individual cognitive and linguistic ability were deemed prone to type I error and were not conducted.

Pupil diameter analysis. Trials were individually inspected for missing data and to identify if automatically coded blinks were correct using the SR Research programme EyeLink Data Viewer (SR Research Ltd., version 4.3.1). Data were grouped by participant and condition and exported for analysis. Down sampling, data preparation, filtering, interpolation, window summary statistics, and generalised additive mixed modelling (GAMM) were conducted using the R statistical programme (version 4.4.1; R Core Team, 2024).

The relative value of the pupil diameter was captured using a sampling rate of 1,000 Hz. Missing data, including blinks, were given a value of −1. Blinks were defined as a period of activity with pupil data missing for three or more consecutive samples. The data were down sampled using the median and a time bin size of 50 ms so that meaningful pupil dilation change could be detected (van Rij et al., 2019). The first trials of a new condition may result in an abnormally large pupil dilation, so the first five trials from each condition were removed (Wendt et al., 2016), resulting in a total of 60 trials per participant. Trials with more than 20% of the data missing were removed prior to analysis, resulting in the removal of 15.7% of trials, consistent with existing literature (Mukai et al., 2018; Porretta & Tucker, 2019; van Rij et al., 2019). The final analysis included 287 noise-interrupted trials, 290 silent-interrupted trials, and 283 uninterrupted trials. The data were then median filtered with a rolling window size of 11 before linear interpolation across missing data such as blinks (Jackson & Sirois, 2009). Finally, the data were baseline aligned (individually by trial) for the 1 s just before sentence onset, or from 4.5 to 5.5 s. Summary statistics including mean and peak dilation by condition and peak dilation latency were extracted from pre-determined time windows. These windows were chosen to begin 1 s after relevant analysis onset and extend long enough for the pupil to dilate but shorter than the mean sentence length (Winn et al., 2018). The data were then visualised for interpretation.

The data were analysed using GAMM, a nonlinear analysis which can be used for analysis of pupil dilation differences. GAMM is a powerful tool that can account for type I errors that occur as a result of autocorrelation between points, a common issue for the slowly changing pupil during pupillometry experiments (Baayen et al., 2018). GAMM can also include nonlinear interactions using two or more numeric predictors (van Rij et al., 2019). Furthermore, GAMM permits direct comparison of the time course of pupil dilation between conditions and visualisation of significant differences in latency between conditions through the use of factor smooths, which allow for the shape of the nonlinear pupil dilation variable to vary by participant (for additional information on smooth terms, see the article by Wood, 2017). Unlike pre-determined time windows, GAMM analysis quantifies pupil diameter differences without *a priori* assumptions, can model complex and nonlinear interactions, and can utilise nonlinear random effects (van Rij et al., 2019). For our analyses, the interruption conditions were entered as predictors for the pupil diameter in the model. In addition, random smooths were entered both for time in milliseconds and for each listener, allowing time to have a different effect for each listener. Each model produced summary statistics including parametric coefficients and smooth terms. In addition, latency comparisons between conditions were visualised

Table 1. Results of cognitive and linguistic measures, the Reading Span Task (RST) measured in percent correct, and the WASI-II Vocabulary Subtest, which is a *t*-score.

	Mean (*SD*)	Range	Skew	Kurtosis
RST	53.9% (18.8)	23–90	−0.41	−0.27
WASI-II Vocabulary	65.4 (12.12)	42–80	−0.45	−0.96
Silent-gap percent correct	30.6% (10.4)	13–55	−0.008	0.49
Noise-burst percent correct	40.8% (10.8)	28–61	0.71	−0.75
Uninterrupted percent correct	96.7% (4.2)	85–100	−1.8	2.85

Key word accuracy for each listening condition is also shown here.

when conditions resulted in significantly different amounts of pupil dilation (i.e., excursion from zero).

The model formula used for the analysis of the entire time series is as follows:

$$pupil_{ij} \sim \beta_0 + \beta_1 \cdot condition_{ij} + s(time_{ij}, condition_{ij}, k = 60) + s(time_{ij}, listener_i, bs = "fs", m = 1) + \rho \cdot AR.start + \epsilon_{ij}$$

In this model, Pupil_{ij} is the pupil response for the *j*-th observation from the *i*-th listener, β_0 is the intercept, β_1 is the coefficient for the fixed effect of interruption condition, and s(time_{ij}, condition_{ij}, $k=60$) represents the smooth term for time, varying by interruption condition, with $k=60$ basis functions. The term s(time_{ij}, listener_i, bs="fs," m=1) represents the random smooth term for time within each listener, using a factor smooth with $m=1$. Autocorrelation was accounted for in the model by extracting the autocorrelation parameter ρ from an earlier model using the function start_value_rho. Autocorrelation was aligned with the start of each trial, denoted by a start_event marker used to reset autocorrelation calculation. Smoothing parameters were estimated using the "*fREML*" method, with a discrete parameter set to TRUE for computational efficiency. The scaled-t distribution family="*scat*" was used to model the variance in the data. Finally, ϵ_{ij} denotes the residual error term.

After onset-aligned visualisation of pupil dilation across the entire trial, analysis was divided into two different time windows: pupil dilations that occurred during the sentence-listening window (i.e., sentence-onset analysis) and those that occurred during the retention interval (i.e., sentence-offset analysis). The listening window was defined as 6.5 s through 9 s, and the retention interval was defined as 11 s through 13 s.

Results

Participants, cognitive and linguistic measures, and perceptual restoration scores

The average age for the 17 participants was 42.1 (*SD*=14.1) and ranged from 22 to 64 years. The average education duration was 17.9 years (*SD*=1.8), with a range of 14–20 years, and the QuickSIN average was 0.14 SNR loss (*SD*=1.41) with a range −2 to 4, representing a range from normal to mildly impaired difficulty in noise. See Table 1 for the results of the cognitive and linguistic measures and perceptual restoration scores.

The RST scores for this cohort are in agreement with published data in a similar cohort using the same measure (Friedman & Miyake, 2005; Füllgrabe et al., 2015), are generally above scores of older listeners with hearing loss (Souza et al., 2015), and are 1 standard deviation below an online administration of the same task (Burleson & Souza, 2022). The WASI-II scores at 65.4 are approximately 1.5 standard deviations above the administration manual's average score, with transformed *t*-values representing a mean of 50 and a standard deviation of 10 (Wechsler, 2009). These data are consistent with those from a sample of participants with a mean of 18 years of education, or a master's degree. Similarly, the perceptual restoration scores for this cohort are in line with existing literature using similar interruption paradigms, with participants scoring significantly better with sentences periodically interrupted by bursts of noise rather than gaps of silence (Bashford et al., 1996; Bologna et al., 2018; Burleson & Souza, 2022; Nagaraj & Magimairaj, 2017) although they did not perform as well as some others, likely due to a difference in sentence corpus complexity (Powers & Wilcox, 1977; Saija et al., 2014) or the experimental nature of pupillometry studies (e.g., head fixation).

Listening window (sentence onset) and retention interval (sentence offset) analysis

First, data were aligned to stimulus onset and visualised (Figure 2). The *a priori* mean/max window for mean and peak dilation, along with peak dilation latency, occurs between 6.5 s and 9 s (Table 2). Overall, silent-gap- and noise-burst-interrupted sentences had the greatest and very similar levels of pupil dilation and peak dilation latency, and uninterrupted sentences had the least amount of dilation with an earlier peak dilation latency.

Next, data were aligned to stimulus offset. The retention interval occurred after the sentence ended and, in this investigation, lasted for 5 s. We recall that the retention interval is thought to be a period during which degraded sentence repair occurs (Nees, 2016; Piquado et al., 2010; Winn, 2016; Winn & Teece, 2021, 2022). For mean and peak dilation along with peak dilation latency, a mean/max window was restricted to 11 s–13 s (Table 2). Again, silent- and noise-interrupted sentences had similar levels of mean and peak dilation and had greater levels of mean dilation than the uninterrupted sentences. All three conditions had similar peak dilation latencies.

Full-trial GAMM analysis

Figure 3 illustrates the estimated difference in pupil size for the three conditions during the retention interval, with significant differences indicated when the confidence interval extends above or below the zero line, complementing the summary statistics (Table 3). For the GAMM analysis, the entire time series was used without the need to predefine time windows, as GAMMs are designed to capture dynamic changes across time. Noise-interrupted sentences resulted in greater pupil dilation than the silent-gap-interrupted sentences for a brief period near the end of the listening window (8.4–9 s). Both interruption conditions led to significantly more pupil dilation than uninterrupted sentences across most of the trial time frame. Both noise- and silent-interrupted sentences elicited a larger pupillary response than the uninterrupted sentences from

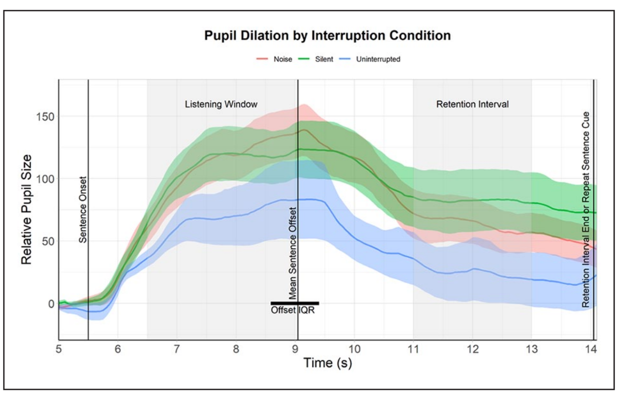

Figure 2. Relative pupil size across time for the silent-interrupted, noise-interrupted, and uninterrupted sentences. Timing markers across trials are described within the figure. The interquartile range (IQR) for sentence offset is represented by the black bar near the mean sentence offset.

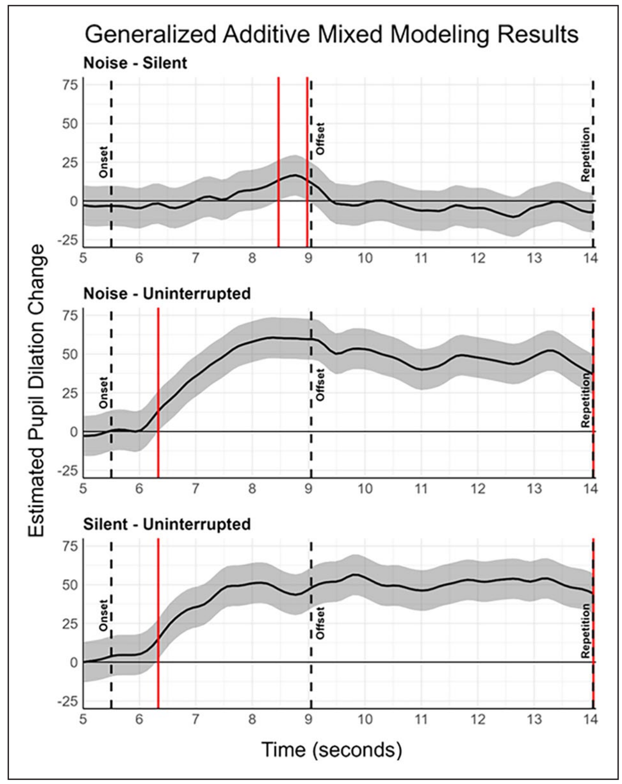

Figure 3. Visualisation of estimated pupil differences from the GAMM results for the three conditions across the entire trial window. Significant excursion from zero indicates a difference in pupil dilation between the compared conditions. Latency of differences is highlighted in red. Note the overlap with participant repetition and the significant difference latency overlap at the end of Panels 2 and 3.

Table 2. Mean pupil dilation, peak dilation, and peak dilation latency by time window (standard deviation in parenthesis).

Mean/Max windows		Mean dilation	Peak dilation	Peak dilation latency (ms)
Baseline window	Silent	0 (33)	79 (53)	
	Noise	0 (35)	87 (73)	
	Uninterrupted	0 (32)	94 (82)	
Listening window	Silent	105 (126)	285 (161)	8,010 (738)
	Noise	114 (121)	302 (166)	8,076 (725)
	Uninterrupted	53 (101)	244 (130)	7,827 (801)
Retention interval	Silent	75 (146)	283 (172)	11,985 (629)
	Noise	64 (136)	270 (154)	11,896 (619)
	Uninterrupted	8 (122)	221 (124)	11,999 (633)

For the listening window, values were aligned to sentence onset, and for the retention interval, they were aligned to sentence offset. Note that peak dilation values reflect the mean of the individual peak dilations across participants for each condition.

Table 3. Relative to the "Noise" condition (intercept), results for the generalised additive mixed model, reporting the standard error (*SE*), *t*-values, and *p*-values for the parametric coefficients and estimated degrees of freedom (E*df*), reference degrees of freedom (Ref.*df*), F, and *p*-values for the smooth terms and random effects.

GAMM Results		Estimate	*SE*	*t*-value	*p*-value
Parametric coefficients	Intercept (noise)	71.0	13.0	5.43	<**.001**
	Silent	0.09	6.0	0.015	.988
	Uninterrupted	−35.8	6.0	−5.97	<**.001**
		E*df*	Ref.*df*	F-value	*p*-Value
Smooth terms	Smooth: Silent	43.75	52.83	8.279	<**.001**
	Smooth: Noise	42.19	51.65	6.70	<**.001**
	Smooth: Uninterrupted	42.87	52.17	7.45	<**.001**
	Random effect: Listener over time	133.82	153	267.74	<**.001**

The "Silent" and "Uninterrupted" conditions are shown as contrasts relative to the noise-interrupted condition. Bolded values indicate p < 0.001; all p-values below 0.05 are considered significant.

shortly after sentence onset at 6.3 s until sentence repetition at 14 s.

Listening window, retention interval, and trial difficulty

Because pupil dilation can be modulated by trial difficulty, we opted to visualise pupil dilation by both time and percent correct across the entire trial. This approach allows for a more nuanced understanding of cognitive effort (pupil dilation) in relation to the extent of perceptual restoration (percent correct). Figure 4 illustrates pupil dilation as a function of percent correct for silent-gap and noise-burst sentences. In the visualisation, dark blue/purple denotes small pupil dilation, and green/yellow represents greater pupil dilation. Consistent with the previous visualisations, pupil dilation increases shortly after stimulus onset in both interrupted conditions. During the listening window, there is a high amount of dilation for trials for both interruption conditions where participants performed well. However, a high degree of dilation appears to occur for silent-gap sentences where participants performed more poorly, getting fewer than 40% of the key words correct. Relative dilation between the two interruption conditions appears similar in the retention interval, consistent with the GAMM analysis. Performance was near ceiling for uninterrupted sentences, and these data were not presented.

Discussion

When speech is disrupted, listeners face a challenging listening experience, especially when the pathway to perceptual repair is unclear or difficult, or when large portions of the speech signal are misperceived. Even when listeners are capable of deploying compensatory mechanisms, they may still struggle to repair or restore missing speech. Previous research has shown a connection between higher-order cognitive and linguistic processes and perceptual restoration (Benard & Başkent, 2014; Burleson & Souza, 2022; Millman & Mattys, 2017; Nagaraj & Magimairaj, 2017) indicating that both repair and attempted repair require cognitive resources. However, deploying these resources demands cognitive effort, and the amount of effort exerted can fluctuate over time.

This investigation aimed to evaluate the relative cognitive effort required when speech is periodically disrupted and difficult to repair, using two types of signal interruption. Our findings indicate that interrupted speech requires more effort, as evidenced by greater pupil dilation, than uninterrupted speech across nearly the entire trial. Pupil dilation was also greater during the listening window than during the retention interval across all listening conditions, suggesting greater cognitive effort during this period. In addition, we found that pupil dilation was significantly higher just before sentence end for speech interrupted by noise than silent interruptions. This may reflect differences in processing between the two interruption types, the perceived differences in difficulty between the two conditions, or the nature of interruptions themselves. Overall, these results indicate that listeners expend additional, sustained cognitive effort when speech is difficult to understand or repair—particularly when the route to perceptual restoration may not be possible.

These findings have important implications for clinicians and researchers, especially in the context of patients with hearing loss or cochlear implants who struggle with speech recognition, particularly in noisy environments. In such situations, speech may be fragmented and misperceived by the listener, leading to greater cognitive resource devotion despite minimal gains in speech recognition. In other words, listeners may expend significant cognitive effort with little reward, which could contribute to frustration or fatigue, potentially worsening the effects of listening fatigue, such as social isolation and a poorer quality of life (Danermark & Gellerstedt, 2004; Demorest & Erdman, 1987; Edwards, 2007; Kramer et al., 1997; Nachtegaal et al., 2009).

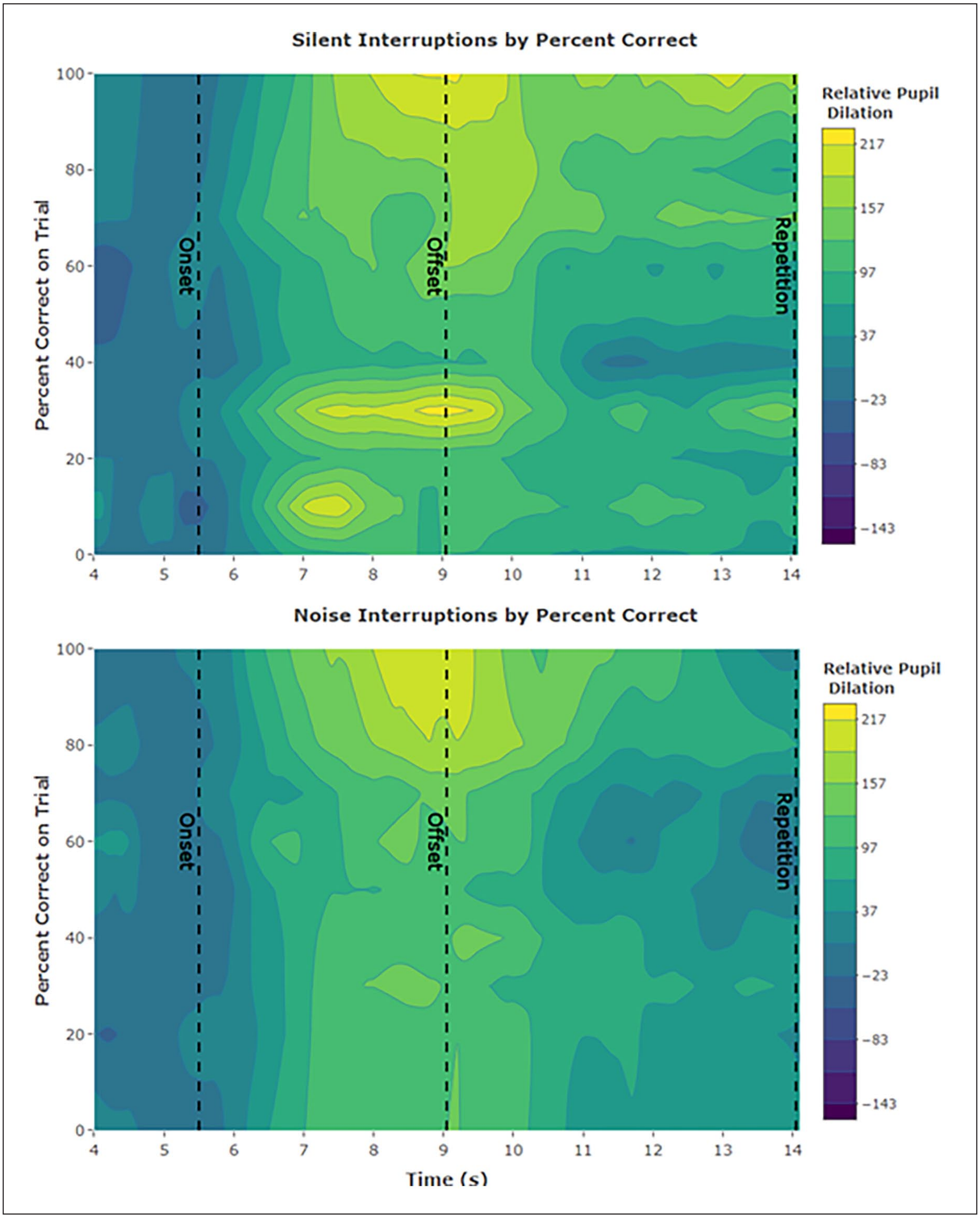

Figure 4. Contour plot representing relative pupil dilation by key word percent correct across the entire trial. Dark blue/purple represents minimal pupil dilation, while yellow represents greater pupil dilation. Sentence onset, offset, and participant repetition markers are delineated by labelled dashed lines.

Cognitive effort during the listening window

There is a large and consistent body of literature indicating that stimulus degradation significantly influences the amount of cognitive effort devoted to a task, as demonstrated by pupillometry (Koelewijn, Zekveld, Festen, & Kramer, 2012; Koelewijn, Zekveld, Festen, Rönnberg, & Kramer, 2012; Winn et al., 2015; Zekveld et al., 2010).

However, when listening to very degraded or difficult auditory stimuli, listeners may experience a “dropout” effect, reducing the amount of cognitive effort devoted to the task (Eckert et al., 2016). This can result in a nonlinear, inverted U-shaped response, where listeners devote minimal cognitive effort to both very easy and very difficult tasks (Kramer et al., 2016; Ohlenforst, Zekveld, Lunner, et al., 2017; Zekveld & Kramer, 2014). Previous investigations adjusted background noise or reduced the spectral fidelity of the target speech, demonstrating that lower intelligibility (i.e., more degradation) led to an increase in cognitive effort (Wendt et al., 2016; Winn et al., 2015).

During the listening window, we predicted that pupil dilation, reflecting cognitive effort, would be comparable between the two interrupted sentence conditions. This prediction was based on the similar level of speech degradation in both conditions, with about 50% of the original speech signal remaining. However, our results did not support this prediction. Pupil dilation was slightly but significantly elevated just before the end of sentence presentation in the noise-interrupted condition compared to the silent-interrupted condition.

The increased effort for noise-interrupted sentences during the listening window could be due to several factors. First, the addition of a louder sound, such as a noise burst, is known to increase pupil dilation (Antikainen & Niemi, 1983). In this investigation, the process of creating the noise interruptions resulted in an approximately 7-dB increase in the overall speech signal, which may explain the elevated pupil dilation during the listening window for this condition. Second, while an alerting tone was used prior to sentence onset in all three conditions to minimise any surprise caused by the stimulus itself, it is possible that the noise bursts were perceived as unexpected, which is also known to increase pupil dilation (Damsma & van Rijn, 2017; Gilzenrat et al., 2010; Marois et al., 2018, 2020; Qiyuan et al., 1985).

However, the difference between the silent-gap and noise-burst interruption during the listening window may be related to differences in processing between the two auditory signals. When speech is periodically interrupted by noise, the speech is perceived as continuous *behind* the noise burst, which has been shown to boost perceptual restoration scores (Bashford et al., 1996; Bashford & Warren, 1987; Bologna et al., 2019; Nagaraj & Magimairaj, 2017; Powers & Wilcox, 1977; Saija et al., 2014). This effect persists as long as the noise burst itself is more intense than the speech, effectively acting as a speech masker (Bashford et al., 1992, 1996; Warren & Obusek, 1971).

In the current investigation, both pupil dilation and performance were higher for the noise-interrupted condition. While larger pupil dilation has been typically associated with *poorer* performance (Miles et al., 2017; Zekveld et al., 2010) (which is consistent with the relationship between interrupted and uninterrupted sentences), the increased dilation during the listening window in the noise-interrupted condition is consistent with the notion that speech-recognition scores are not necessarily related to the amount of cognitive effort devoted to a task (Winn & Teece, 2021, 2022). In other words, non-auditory aspects of the signal can lead to improved performance *and* additional effort. For example, the perception that speech is continuous behind the noise bursts may trigger additional lexical candidates that “fill in the gap” between speech fragments (Akeroyd, 2008; Cooke, 2006). This, in turn, may enable listeners to connect speech fragments and resolve the interruption (Luce & Pisoni, 1998; Taler et al., 2010). By doing so, listeners may unearth contextual clues that support sentence understanding, such as lexical prediction (Altmann & Mirković, 2009; Brothers & Kuperberg, 2021; Federmeier & Kutas, 1999; Kuperberg & Jaeger, 2016). Next, while both interruption conditions provide listeners with access to similar cues, such as temporal cues (e.g., vowel position when available, amplitude modulation of speech patterns) and spectral cues (e.g., transient consonants), those cues may be used more effectively in the noise-burst condition. This may result in greater pupil dilation, or cognitive effort, in this condition (Fogerty et al., 2015). Evidence also shows that pupil dilation increases with increasing linguistic processing demands, suggesting that the greater pupil dilation for noise-interrupted sentences may reflect the extra effort needed for additional linguistic processing (Ahern & Beatty, 1979; Engelhardt et al., 2010; Kuchinke et al., 2007; Kuchinsky et al., 2013; Vogelzang et al., 2016; Wendt et al., 2016).

Consistent with existing literature, listeners exhibited poorer recognition of silent-gap-interrupted sentences, yet this poorer performance did not translate into higher pupil dilation during the listening window (Miles et al., 2017; Zekveld et al., 2010). Listening to speech interrupted by gaps of silence is uncommon in daily life, at least when compared to the more prevalent experience of listening to speech interrupted by noise, such as a noisy restaurant. Although novel stimuli can elicit greater pupil dilation (Johansson & Balkenius, 2018; Marois et al., 2020), no such effect was apparent between the interruption conditions in this investigation. The pattern of results suggests that for the silent-gap sentences, lower levels of cognitive effort achieved relatively low recognition performance. One plausible explanation is that when sentences are interrupted by silent gaps, the onset of the interruption may be misperceived as a word boundary, limiting lexical activation and possible lexical candidates to fill in the missing gap (Mattys & Clark, 2002). Listening to silent-gap sentences may be akin to listening to a foreign language, where speech information is clearly audible but not matched to known words stored in a listener’s long-term memory, reducing recognition and limiting the utility of additional cues such as context and prediction.

The dilation increase during the listening window for the noise-interrupted sentences in the current investigation may also be related to attempts at chunking or storing the auditory "glimpses" for later processing (Baddeley & Della Sala, 1996; Baddeley & Hitch, 1994; Cooke, 2006; A. C. Gilbert et al., 2014; Li & Loizou, 2007), consistent with current theories of degraded speech processing, such as the Ease of Language Understanding model (Rönnberg et al., 2013, 2021). This result is also consistent with data indicating that pupil dilation increases with increasing memory demands (until those demands are exceeded), such as increasing sequence length during successful digit recall (Kahneman & Beatty, 1966; Klingner et al., 2011; Piquado et al., 2010).

Cognitive effort during the retention interval

The retention interval is a time window where sustained pupil dilation may signal ongoing cognitive processing. For example, research by Kahneman and Beatty (1966) found that solving difficult math problems elicited more prolonged pupil dilation than easier ones and that the pupils of high-aptitude students returned to baseline more quickly than those of lower-aptitude students, especially when the problem was simple (Ahern & Beatty, 1979). Sustained post-stimulus pupil dilation may also indicate the amount of ongoing information storage and reprocessing, as evidenced by pupil dilation gradually decreasing as participants serially recite from memory a string of previously memorised digits (Beatty, 1982). Auditory data suggest that the pupillary response is sustained during the retention interval for more degraded stimuli (Koelewijn et al., 2015; Koelewijn, Zekveld, Festen, & Kramer, 2012; Winn et al., 2015; Zekveld et al., 2010) and for lower-context materials (Winn, 2016; Winn & Moore, 2018). Listeners may improve accuracy by using the quiet-retention interval as additional processing time (Gianakas et al., 2022). Previous literature supports the notion that prolonged pupil dilation that extends into the retention interval reflects additional ongoing processing (Francis et al., 2016; Koelewijn et al., 2017; Wendt et al., 2018), and that this ongoing processing continues even with semantic ambiguity (Kadem et al., 2020) or when misperceptions are correctly perceived (Winn & Teece, 2021, 2022) and may reflect differences in learning strategy (McHaney et al., 2021). This aligns with the current findings, suggesting that more difficult listening conditions result in sustained dilations or a shallower downwards slope than easier conditions. These notions are quite consistent with the present data, where both interrupted sentence conditions exhibited sustained pupil dilation compared to the uninterrupted condition. Indeed, pupil dilation for the uninterrupted sentences quickly returns to near baseline, requiring little to no ongoing processing beyond short-term memory of the sentence itself and no ongoing cognitive effort.

Both interrupted sentence conditions required sustained cognitive effort across nearly the entire retention interval, although the two interruption conditions were not significantly different from one another. This result is consistent with our prediction that the level of cognitive effort would be sustained for both interrupted conditions. While silent-gap-interrupted sentences pose more challenges for listeners, resulting in lower key word–correct scores (Bashford et al., 1996; Burleson & Souza, 2022; Nagaraj & Magimairaj, 2017; Powers & Wilcox, 1977), this did not result in significant dilation differences during the retention interval itself. Typically, the silent-interrupted conditions are perceived as less continuous because silent gaps are erroneously perceived as word boundaries, reducing overall lexical activation and requiring additional post-stimulus repair (Bashford et al., 1992; Bashford & Warren, 1987; Clarke et al., 2016; Winn & Teece, 2021). However, results for the more-difficult silent-gap sentences do not suggest that any additional post-stimulus repair occurred beyond what was observed for the noise-burst sentences.

The relatively similar levels of sustained effort for both conditions may reflect some post-stimulus repair or other cognitive processes (Francis et al., 2016; Koelewijn et al., 2017; Wendt et al., 2018). First, because the effort for both interrupted sentences was higher than that of the uninterrupted sentences, this suggests that some post-stimulus repair was attempted, even when perceptual repair was very difficult or impossible. This sustained effort points to the possible engagement of compensatory strategies as listeners attempt to make sense of the degraded auditory signal. Listeners may actively engage in lexical retrieval, a process that may demand ongoing cognitive resources even after the speech has ended (Altmann & Mirković, 2009; Rönnberg et al., 2013, 2021). This additional processing may continue even when speech cannot be repaired, or perceptual restoration is unlikely to take place, as listeners still attempt to resolve ambiguous speech misperceptions whether or not they are able to correctly repair them (Winn & Teece, 2021, 2022).

The sustained pupil dilation during the retention interval may also reflect the memory processes required to temporarily store the speech as listeners prepare to repeat it. While the pupillary response for the uninterrupted condition returns to near baseline during the retention interval, a level of sustained effort is still evident that may reflect maintenance of the sentence in short-term memory (Bönitz et al., 2021; Johnson et al., 2014; Kahneman & Wright, 1971). Although intact speech is likely less effortful to recall than misperceived speech, increases in pupil dilation above this level may reflect processes occurring in addition to retention in short-term memory.

Hearing difficulty, cognitive effort, and perceptual restoration. The present findings show that listeners continue to expend cognitive effort—even after the sentence has ended—when processing disrupted, misperceived speech. The combination of additional time *and* effort may cause listeners with hearing loss to fall behind as they work to repair misperceived speech, deploying additional effort just to keep up with the ongoing conversation, even when they accurately repair speech (Winn, 2024; Winn & Teece, 2021, 2022). Indeed, individuals with hearing loss benefit from brief pauses to "catch up" (Gianakas et al., 2022) and may misperceive upcoming speech if still actively repairing what was just heard (Winn, 2024). With these factors in mind, consider day-to-day conversation where information flows rapidly between interlocutors (Heldner & Edlund, 2010). Even the initiation of conversational repair, such as asking "what did you say?", usually occurs only about 1 s after the breakdown, a duration much longer than the retention interval used in many investigations (Kendrick, 2015). To further complicate matters, extended pauses between turns can be perceived as disagreement by a conversational partner (Roberts et al., 2006, 2011). This leaves little time for additional processing and heightens the pressure for listeners to respond.

If speech is not readily matched to long-term memory representations (Rönnberg et al., 2013), having extra processing time aids in resolving lexical competition and ambiguity (Farris-Trimble et al., 2014; Gianakas & Winn, 2019; McMurray et al., 2017). As lexical components are clarified, additional cues may become available to the listener during the retention interval. One such cue may be sentence prediction, which is less readily used by listeners with hearing loss during sentence listening alone (Winn & Moore, 2018). Moreover, accurate perception for listeners with hearing loss may depend on post-speech cognitive processes such as *postdiction* to cope with and recover from difficult listening situations (Rönnberg et al., 2019, 2021). Thus, abilities facilitating speech recovery *after* sentence presentation may rely on higher-order processes such as working memory (the ability to simultaneously store and reprocess information) and lexical knowledge (knowledge about words in a language). Although the use of individual abilities such as working memory to predict pupil dilation remains elusive (Winn et al., 2018), these abilities appear to support perceptual restoration in some previous investigations (Benard et al., 2014; Burleson & Souza, 2022; Nagaraj & Magimairaj, 2017) and may be a focus for future research. These compensatory processes are important for everyday listening, both for listeners with hearing loss and for those without hearing loss in challenging auditory environments.

Limitations. This study characterises the important role that cognitive effort plays during disrupted sentence listening with limited opportunity for speech repair. Investigations into cognitive effort, including the present study, are not without limitations. One limitation of the current investigation is the long retention interval of 5 s. Although the intention was to give listeners ample time for repair, retention intervals of this length are not common and may engage short-term memory processes above and beyond what is expected as listeners maintain speech before repeating it (Bönitz et al., 2021). In addition, given the amount of intraindividual variability inherent in pupil-dilation data (Winn et al., 2018), it was not possible to characterise how individual cognitive and linguistic predictors affect effort. Such a prediction might be possible with a larger cohort than is typical in most pupillometry literature. While the current investigation nearly met the number of participants indicated via power analysis, future investigations replicating the relationship between perceptual restoration and cognitive effort should aim to include a larger cohort and to keep in mind the relationship between effect size and statistical power. In addition, stimuli that can differentiate between sentences with high repair (high dilation, high performance) and low repair (high dilation, low performance), as well as stimuli that were too easy (low dilation, high performance) or too difficult (low dilation, low performance) may more clearly elucidate the components that support perceptual restoration tasks. These areas are interesting opportunities for future work.

Conclusion

The current investigation indicates that pupil dilation, a marker of cognitive effort, is greater for speech interrupted by gaps of noise or silence than for uninterrupted speech. The data reveal a larger pupillary response for speech interrupted by noise just before the sentence ends, but also that pupil dilation is similar and sustained across the retention interval for both interruption conditions. These data provide insights into the time course of processing disrupted speech and demonstrate that elevated and sustained levels of pupil dilation, which may reflect cognitive effort, persist when speech is misperceived. This misperception may inadvertently cause disruption of upcoming speech recognition (Winn, 2024; Winn & Teece, 2021, 2022), causing additional difficulty in conversation and real-world everyday listening. Understanding the time course of effort and repair during both sentence listening and the retention interval sheds light on the additional effort devoted to misperceived speech for listeners with normal hearing or mild hearing loss. The increased effort required for disrupted or degraded speech may explain why some listeners experience listening fatigue when speech is misperceived or disrupted.

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.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was partially supported by NIH (R01 DC006014).

ORCID iD

Andrew M Burleson https://orcid.org/0000-0002-4857-3583

References

Ahern, S., & Beatty, J. (1979). Pupillary responses during information processing vary with Scholastic Aptitude Test scores. *Science*, *205*(4412), 1289–1292.

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology*, *47*(Suppl. 2), S53–S71.

Altmann, G. T., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, *33*(4), 583–609.

Antikainen, J., & Niemi, P. (1983). Neuroticism and the pupillary response to a brief exposure to noise. *Biological Psychology*, *17*(2–3), 131–135.

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.

Baayen, R. H., van Rij, J., De Cat, C., & Wood, S. (2018). Autocorrelated errors in experimental data in the language sciences: Some solutions offered by Generalized Additive Mixed Models. In D. Speelman, K. Heylen, D. Geeraerts (Eds.), *Mixed-effects regression models in linguistics* (pp. 49–69). Springer International Publishing.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423.

Baddeley, A. D., & Della Sala, S. (1996). Working memory and executive control. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1346), 1397–1404.

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology*, *8*(4), 485.

Bashford, J. A., Riener, K. R., & Warren, R. M. (1992). Increasing the intelligibility of speech through multiple phonemic restorations. *Perception & Psychophysics*, *51*(3), 211–217.

Bashford, J. A., & Warren, R. M. (1987). Multiple phonemic restorations follow the rules for auditory induction. *Perception & Psychophysics*, *42*(2), 114–121.

Bashford, J. A., Warren, R. M., & Brown, C. A. (1996). Use of speech-modulated noise adds strong "bottom-up" cues for phonemic restoration. *Perception & Psychophysics*, *58*(3), 342–350.

Başkent, D. (2010). Phonemic restoration in sensorineural hearing loss does not depend on baseline speech perception scores. *The Journal of the Acoustical Society of America*, *128*(4), EL169–EL174.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*(2), 276–292.

Benard, M. R., & Başkent, D. (2014). Perceptual learning of temporally interrupted spectrally degraded speech. *The Journal of the Acoustical Society of America*, *136*(3), 1344–1351.

Benard, M. R., Mensink, J. S., & Başkent, D. (2014). Individual differences in top-down restoration of interrupted speech: Links to linguistic and cognitive abilities. *The Journal of the Acoustical Society of America*, *135*(2), EL88–EL94.

Bologna, W. J., Vaden, K. I., Jr., Ahlstrom, J. B., & Dubno, J. R. (2018). Age effects on perceptual organization of speech: Contributions of glimpsing, phonemic restoration, and speech segregation. *The Journal of the Acoustical Society of America*, *144*(1), 267–281.

Bologna, W. J., Vaden, K. I., Jr., Ahlstrom, J. B., & Dubno, J. R. (2019). Age effects on the contributions of envelope and periodicity cues to recognition of interrupted speech in quiet and with a competing talker. *The Journal of the Acoustical Society of America*, *145*(3), EL173–EL178.

Bönitz, H., Lunner, T., Finke, M., Fiedler, L., Lyxell, B., Riis, S. K., & . . .Wendt, D. (2021). How do we allocate our resources when listening and memorizing speech in noise? A pupillometry study. *Ear and Hearing*, *42*(4), 846–859.

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology*, *40*(1), 109–131.

Brothers, T., & Kuperberg, G. R. (2021). Word predictability effects are linear, not logarithmic: Implications for probabilistic models of sentence comprehension. *Journal of Memory and Language*, *116*, 104174. https://doi.org/10.1016/j.jml.2020.104174

Burleson, A. M., & Souza, P. E. (2022). Cognitive and linguistic abilities and perceptual restoration of missing speech: Evidence from online assessment. *Frontiers in Psychology*, *13*, Article 1059192.

Clarke, J., Başkent, D., & Gaudrain, E. (2016). Pitch and spectral resolution: A systematic comparison of bottom-up cues for top-down repair of degraded speech. *The Journal of the Acoustical Society of America*, *139*(1), 395–405.

Cooke, M. (2006). A glimpsing model of speech perception in noise. *The Journal of the Acoustical Society of America*, *119*(3), 1562–1573.

Damsma, A., & van Rijn, H. (2017). Pupillary response indexes the metrical hierarchy of unattended rhythmic violations. *Brain and Cognition*, *111*, 95–103.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Danermark, B., & Gellerstedt, L. C. (2004). Social justice: Redistribution and recognition—A non-reductionist perspective on disability. *Disability & Society*, *19*(4), 339–353.

Demorest, M. E., & Erdman, S. A. (1987). Development of the communication profile for the hearing impaired. *Journal of Speech and Hearing Disorders*, *52*(2), 129–143.

Dietrich, V., Nieschalk, M., Stoll, W., Rajan, R., & Pantev, C. (2001). Cortical reorganization in patients with high frequency cochlear hearing loss. *Hearing Research*, *158*(1–2), 95–101.

Eckert, M. A., Teubner-Rhodes, S., & Vaden, K. I., Jr. (2016). Is listening in noise worth it? The neurobiology of speech recognition in challenging listening conditions. *Ear and Hearing*, *37*, 101S–110S.

Edwards, B. (2007). The future of hearing aid technology. *Trends in Amplification*, *11*(1), 31–45.

Engelhardt, P. E., Ferreira, F., & Patsenko, E. G. (2010). Pupillometry reveals processing load during spoken language comprehension. *Quarterly Journal of Experimental Psychology*, *63*(4), 639–645.

Farris-Trimble, A., McMurray, B., Cigrand, N., & Tomblin, J. B. (2014). The process of spoken word recognition in the face of signal degradation. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 308–327. https://doi.org/10.1037/a0034353

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, *41*(4), 469–495.

Fogerty, D., Ahlstrom, J. B., Bologna, W. J., & Dubno, J. R. (2015). Sentence intelligibility during segmental interruption and masking by speech-modulated noise: Effects of age and hearing loss. *The Journal of the Acoustical Society of America*, *137*(6), 3487–3501.

Francis, A. L., MacPherson, M. K., Chandrasekaran, B., & Alvar, A. M. (2016). Autonomic nervous system responses during perception of masked speech may Reflect constructs other than subjective listening effort. *Frontiers in Psychology*, *7*, Article 263.

Friedman, N. P., & Miyake, A. (2005). Comparison of four scoring methods for the reading span test. *Behavior Research Methods*, *37*(4), 581–590.

Füllgrabe, C., Moore, B. C., & Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition. *Frontiers in Aging Neuroscience*, *6*, Article 347.

Gatehouse, S., & Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *International Journal of Audiology*, *43*(2), 85–99. https://doi.org/10.1080/14992020400050014

Gianakas, S. P., Fitzgerald, M. B., & Winn, M. B. (2022). Identifying listeners whose speech intelligibility depends on a quiet extra moment after a sentence. *Journal of Speech, Language, and Hearing Research*, *65*(12), 4852–4865.

Gianakas, S. P., & Winn, M. B. (2019). Lexical bias in word recognition by cochlear implant listeners. *The Journal of the Acoustical Society of America*, *146*(5), 3373–3383. https://doi.org/10.1121/1.5132938

Gilbert, A. C., Boucher, V. J., & Jemel, B. (2014). Perceptual chunking and its effect on memory in speech processing: ERP and behavioral evidence. *Frontiers in Psychology*, *5*, Article 220.

Gilbert, J. L., Tamati, T. N., & Pisoni, D. B. (2013). Development, reliability, and validity of PRESTO: A new high-variability sentence recognition test. *Journal of the American Academy of Audiology*, *24*(1), 26–36.

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.

Glick, H. A., & Sharma, A. (2020). Cortical neuroplasticity and cognitive function in early-stage, mild-moderate hearing loss: Evidence of neurocognitive benefit from hearing aid use. *Frontiers in Neuroscience*, *14*, Article 93.

Heldner, M., & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. *Journal of Phonetics*, *38*(4), 555–568.

Hughes, S. E., Rapport, F., Watkins, A., Boisvert, I., McMahon, C. M., & Hutchings, H. A. (2019). Study protocol for the validation of a new patient-reported outcome measure (PROM) of listening effort in cochlear implantation: The Listening Effort Questionnaire-Cochlear Implant (LEQ-CI). *BMJ Open*, *9*(7), e028881.

Jackson, I., & Sirois, S. (2009). Infant cognition: Going full factorial with pupil dilation. *Developmental Science*, *12*(4), 670–679.

Johansson, B., & Balkenius, C. (2018). A computational model of pupil dilation. *Connection Science*, *30*(1), 5–19.

Johnson, E. L., Miller Singley, A. T., Peckham, A. D., Johnson, S. L., & Bunge, S. A. (2014). Task-evoked pupillometry provides a window into the development of short-term memory capacity. *Frontiers in Psychology*, *5*, Article 218.

Kadem, M., Herrmann, B., Rodd, J. M., & Johnsrude, I. S. (2020). Pupil dilation is sensitive to semantic ambiguity and acoustic degradation. *Trends in Hearing*, *24*, 2331216520964068.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Citeseer.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583–1585.

Kahneman, D., & Wright, P. (1971). Changes of pupil size and rehearsal strategies in a short-term memory task. *The Quarterly Journal of Experimental Psychology*, *23*(2), 187–196.

Kendrick, K. H. (2015). The intersection of turn-taking and repair: The timing of other-initiations of repair in conversation. *Frontiers in Psychology*, *6*, Article 250.

Kidd, G. R., & Humes, L. E. (2012). Effects of age and hearing loss on the recognition of interrupted words in isolation and in sentences. *The Journal of the Acoustical Society of America*, *131*(2), 1434–1448.

Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *116*(4), 2395–2405.

Klingner, J., Tversky, B., & Hanrahan, P. (2011). Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. *Psychophysiology*, *48*(3), 323–332.

Koelewijn, T., de Kluiver, H., Shinn-Cunningham, B. G., Zekveld, A. A., & Kramer, S. E. (2015). The pupil response reveals increased listening effort when it is difficult to focus attention. *Hearing Research*, *323*, 81–90.

Koelewijn, T., Versfeld, N. J., & Kramer, S. E. (2017). Effects of attention on the speech reception threshold and pupil response of people with impaired and normal hearing. *Hearing Research*, *354*, 56–63.

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), 291–300.

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*, 865731.

Kramer, S. E., Kapteyn, T. S., Festen, J. M., & Kuik, D. J. (1997). Assessing aspects of auditory handicap by means of pupil dilatation. *Audiology*, *36*(3), 155–164.

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.

Kuchinke, L., Võ, M. L.-H., Hofmann, M., & Jacobs, A. M. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology*, *65*(2), 132–140.

Kuchinsky, S. E., Ahlstrom, J. B., Vaden, K. I., Cute, 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.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32–59.

Li, N., & Loizou, P. C. (2007). Factors influencing glimpsing of speech in noise. *The Journal of the Acoustical Society of America*, *122*(2), 1165–1172.

Liberman, M. C., & Kujawa, S. G. (2017). Cochlear synaptopathy in acquired sensorineural hearing loss: Manifestations and mechanisms. *Hearing Research*, *349*, 138–147.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1.

Marois, A., Labonté, K., Parent, M., & Vachon, F. (2018). Eyes have ears: Indexing the orienting response to sound using pupillometry. *International Journal of Psychophysiology*, *123*, 152–162.

Marois, A., Pozzi, A., & Vachon, F. (2020). Assessing the role of stimulus novelty in the elicitation of the pupillary dilation response to irrelevant sound. *Auditory Perception & Cognition*, *3*(1–2), 1–17.

Mattys, S. L., & Clark, J. H. (2002). Lexical activity in speech processing: Evidence from pause detection. *Journal of Memory and Language*, *47*(3), 343–359.

McHaney, J. R., Tessmer, R., Roark, C. L., & Chandrasekaran, B. (2021). Working memory relates to individual differences in speech category learning: Insights from computational modeling and pupillometry. *Brain and Language*, *222*, 105010.

McMurray, B., Farris-Trimble, A., & Rigler, H. (2017). Waiting for lexical access: Cochlear implants or severely degraded input lead listeners to process speech less incrementally. *Cognition*, *169*, 147–164. https://doi.org/10.1016/j.cognition.2017.08.013

Miles, K., McMahon, C., Boisvert, I., Ibrahim, R., De Lissa, P., Graham, P., & Lyxell, B. (2017). Objective assessment of listening effort: Coregistration of pupillometry and EEG. *Trends in Hearing*, *21*, 2331216517706396.

Miller, G. A., & Licklider, J. C. (1950). The intelligibility of interrupted speech. *The Journal of the Acoustical Society of America*, *22*(2), 167–173.

Millman, R. E., & Mattys, S. L. (2017). Auditory verbal working memory as a predictor of speech perception in modulated maskers in listeners with normal hearing. *Journal of Speech, Language, and Hearing Research*, *60*(5), 1236–1245. https://doi.org/10.1044/2017_JSLHR-S-16-0105

Mukai, Y., Järvikivi, J., & Tucker, B. V. (May 30–June 1, 2018). *The effect of phonological-orthographic consistency on the processing of reduced and citation forms of Japanese words: Evidence from pupillometry*. Paper presented at the proceedings of the 2018 annual conference of the Canadian Linguistics Association in Saskatchewan, Canada.

Murphy, P. R., Robertson, I. H., Balsters, J. H., & O'Connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus–noradrenergic arousal function in humans. *Psychophysiology*, *48*(11), 1532–1543.

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.

Nagaraj, N. K., & Magimairaj, B. M. (2017). Role of working memory and lexical knowledge in perceptual restoration of interrupted speech. *The Journal of the Acoustical Society of America*, *142*(6), 3756–3766.

Nees, M. A. (2016). Have we forgotten auditory sensory memory? Retention intervals in studies of nonverbal auditory working memory. *Frontiers in Psychology*, *7*, 1892.

Ohlenforst, B., Zekveld, A. A., Jansma, E. P., Wang, Y., Naylor, G., Lorens, A., & . . .Kramer, S. E. (2017). Effects of hearing impairment and hearing aid amplification on listening effort: A systematic review. *Ear and Hearing*, *38*(3), 267–281.

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.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., . . .Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and Hearing*, *37*, 5S–27S. https://doi.org/10.1097/AUD.0000000000000312

Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*, *47*(3), 560–569.

Porretta, V., & Tucker, B. V. (2019). Eyes wide open: Pupillary response to a foreign accent varying in intelligibility. *Frontiers in Communication*, *4*, Article 8.

Powers, G. L., & Wilcox, J. C. (1977). Intelligibility of temporally interrupted speech with and without intervening noise. *The Journal of the Acoustical Society of America*, *61*(1), 195–199.

Qiyuan, J., Richer, F., Wagoner, B. L., & Beatty, J. (1985). The pupil and stimulus probability. *Psychophysiology*, *22*(5), 530–534.

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org

Reimer, J., McGinley, M. J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D. A., & Tolias, A. S. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nature Communications*, *7*(1), 13289.

Richter, M., Gendolla, G., & Wright, R. (2016). Three decades of research on motivational intensity theory: What we have learned about effort and what we still don't know. *Advances in Motivation Science*, *3*, 149–186.

Roberts, F., Francis, A. L., & Morgan, M. (2006). The interaction of inter-turn silence with prosodic cues in listener perceptions of "trouble" in conversation. *Speech Communication*, *48*(9), 1079–1093. https://doi.org/10.1016/j.specom.2006.02.001

Roberts, F., Margutti, P., & Takano, S. (2011). Judgments concerning the valence of inter-turn silence across speakers of American English, Italian, and Japanese. *Discourse Processes*, *48*(5), 331–354. https://doi.org/10.1080/0163853X.2011.558002

Rönnberg, J., Holmer, E., & Rudner, M. (2019). Cognitive hearing science and ease of language understanding. *International Journal of Audiology*, *58*(5), 247–261. https://doi.org/10.1080/14992027.2018.1551631

Rönnberg, J., Holmer, E., & Rudner, M. (2021). Cognitive hearing science: Three memory systems, two approaches, and the Ease of Language Understanding model. *Journal of Speech, Language, and Hearing Research*, *64*, 359–370.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., & . . .Pichora-Fuller, M. K. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, *7*, Article 31.

Saija, J. D., Akyürek, E. G., Andringa, T. C., & Başkent, D. (2014). Perceptual restoration of degraded speech is preserved with advancing age. *Journal of the Association for Research in Otolaryngology*, *15*(1), 139–148.

Shafiro, V., Sheft, S., & Risley, R. (2016). The intelligibility of interrupted and temporally altered speech: Effects of context, age, and hearing loss. *The Journal of the Acoustical Society of America*, *139*(1), 455–465.

Shafiro, V., Sheft, S., Risley, R., & Gygi, B. (2015). Effects of age and hearing loss on the intelligibility of interrupted speech. *The Journal of the Acoustical Society of America*, *137*(2), 745–756.

Souza, P., Arehart, K., & Neher, T. (2015). Working memory and hearing aid processing: Literature findings, future directions, and clinical applications. *Frontiers in Psychology*, *6*, Article 1894.

Taler, V., Aaron, G. P., Steinmetz, L. G., & Pisoni, D. B. (2010). Lexical neighborhood density effects on spoken word recognition and production in healthy aging. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *65*(5), 551–560.

Tremblay, K. L., Pinto, A., Fischer, M. E., Klein, B. E., Klein, R., Levy, S., & . . .Cruickshanks, K. J. (2015). Self-reported hearing difficulties among adults with normal audiograms: The Beaver Dam Offspring Study. *Ear and Hearing*, *36*(6), e290–e299.

Vaden, K. I., Jr., Kuchinsky, S. E., Cute, S. L., Ahlstrom, J. B., Dubno, J. R., & Eckert, M. A. (2013). The cingulo-opercular network provides word-recognition benefit. *Journal of Neuroscience*, *33*(48), 18979–18986.

Vaden, K. I., Jr., Teubner-Rhodes, S., Ahlstrom, J. B., Dubno, J. R., & Eckert, M. A. (2017). Cingulo-opercular activity affects incidental memory encoding for speech in noise. *NeuroImage*, *157*, 381–387.

van Rij, J., Hendriks, P., van Rijn, H., Baayen, R. H., & Wood, S. N. (2019). Analyzing the time course of pupillometric data. *Trends in Hearing*, *23*, 2331216519832483.

Vijayasarathy, S., Mohan, M., Nagalakshmi, P., & Barman, A. (2021). Speech perception in noise, gap detection and amplitude modulation detection in suspected hidden hearing loss. *Hearing, Balance and Communication*, *19*, 203–211.

Vogelzang, M., Hendriks, P., & van Rijn, H. (2016). Pupillary responses reflect ambiguity resolution in pronoun processing. *Language, Cognition and Neuroscience*, *31*(7), 876–885.

Wagner, A., Toffanin, P., & Başkent, D. (2016). The timing and effort of lexical access in natural and degraded speech. *Frontiers in Psychology*, *7*, Article 398.

Wang, X., & Humes, L. E. (2010). Factors influencing recognition of interrupted speech. *The Journal of the Acoustical Society of America*, *128*(4), 2100–2111.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, *167*(3917), 392–393.

Warren, R. M., & Obusek, C. J. (1971). Speech perception and phonemic restorations. *Perception & Psychophysics*, *9*(3), 358–362.

Wechsler, D. (2009). *Wechsler abbreviated scale of intelligence (WASI-II)*. Pearson.

Wendt, D., Dau, T., & Hjortkjær, J. (2016). Impact of background noise and sentence complexity on processing demands during sentence comprehension. *Frontiers in Psychology*, *7*, Article345.

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.

Winn, M. B. (2016). Rapid release from listening effort resulting from semantic context, and effects of spectral degradation and cochlear implants. *Trends in Hearing*, *20*, 2331216516669723.

Winn, M. B. (2024). The effort of repairing a misperceived word can impair perception of following words, especially for listeners with cochlear implants. *Ear and Hearing*, *45*(6), 1527–1541. https://doi.org/10.1097/AUD.0000000000001537

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), e153.

Winn, M. B., & Moore, A. N. (2018). Pupillometry reveals that context benefit in speech perception can be disrupted by later-occurring sounds, especially in listeners with cochlear implants. *Trends in Hearing*, *22*, 2331216518808962.

Winn, M. B., & Teece, K. H. (2021). Listening effort is not the same as speech intelligibility score. *Trends in Hearing*, *25*, 23312165211027688.

Winn, M. B., & Teece, K. H. (2022). Effortful listening despite correct responses: The cost of mental repair in sentence recognition by listeners with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *65*(10), 3966–3980.

Winn, M. B., Wendt, D., Koelewijn, T., & Kuchinsky, S. E. (2018). Best practices and advice for using pupillometry to measure listening effort: An introduction for those who want to get started. *Trends in Hearing*, *22*, 2331216518800869.

Wood, S. N. (2017). *Generalized additive models: An introduction with R*. CRC Press.

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

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.