

# Strategic Pauses Relieve Listeners from the Effort of Listening to Fast Speech: Data Limited and Resource Limited Processes in Narrative Recall by Adult Users of Cochlear Implants

Trends in Hearing  
Volume 27: 1–22  
© The Author(s) 2023  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/23312165231203514  
journals.sagepub.com/home/tia



Ryan M. O’Leary<sup>1</sup> , Jonathan Neukam<sup>2</sup> , Thomas A. Hansen<sup>1</sup> ,  
Alexander J. Kinney<sup>1</sup> , Nicole Capach<sup>2</sup> , Mario A. Svirsky<sup>2</sup> ,  
and Arthur Wingfield<sup>1</sup> 

## Abstract

Speech that has been artificially accelerated through time compression produces a notable deficit in recall of the speech content. This is especially so for adults with cochlear implants (CI). At the perceptual level, this deficit may be due to the sharply degraded CI signal, combined with the reduced richness of compressed speech. At the cognitive level, the rapidity of time-compressed speech can deprive the listener of the ordinarily available processing time present when speech is delivered at a normal speech rate. Two experiments are reported. Experiment 1 was conducted with 27 normal-hearing young adults as a proof-of-concept demonstration that restoring lost processing time by inserting silent pauses at linguistically salient points within a time-compressed narrative (“time-restoration”) returns recall accuracy to a level approximating that for a normal speech rate. Noise vocoder conditions with 10 and 6 channels reduced the effectiveness of time-restoration. Pupil dilation indicated that additional effort was expended by participants while attempting to process the time-compressed narratives, with the effortful demand on resources reduced with time restoration. In Experiment 2, 15 adult CI users tested with the same (unvocoder) materials showed a similar pattern of behavioral and pupillary responses, but with the notable exception that meaningful recovery of recall accuracy with time-restoration was limited to a subgroup of CI users identified by better working memory spans, and better word and sentence recognition scores. Results are discussed in terms of sensory-cognitive interactions in data-limited and resource-limited processes among adult users of cochlear implants.

## Keywords

cochlear implants, time-compressed speech, narrative recall, pupillometry, speech vocoding

Received 18 January 2023; Revised received 11 August 2023; accepted 8 September 2023

## Introduction

Cochlear implants (CIs) have seen increasing use as a means of restoring hearing for individuals whose hearing loss exceeds the restorative capacity of conventional hearing aids. It is the case, however, that the signal produced by CIs differs markedly from the spectral richness of natural speech. Although current CIs may contain as many as 22 intra-cochlear electrodes, CI users may not have spectral resolution (perceptually separated bands of frequency information) beyond the equivalent of four to eight frequency channels due to factors such as current spread and neural survival (Friesen et al., 2001; Fu & Nogaki, 2005; Perreau et al.,

2010). To put these numbers in perspective, although speech intelligibility improves as the number of spectral channels increases up to the maximum of 22 (Croghan et al., 2017),

<sup>1</sup>Department of Psychology, Brandeis University, Waltham, Massachusetts, USA

<sup>2</sup>Department of Otolaryngology, NYU Langone Medical Center, New York, New York, USA

### Corresponding Author:

Arthur Wingfield, Department of Psychology and Volen National Center, for Complex Systems, Brandeis University, Waltham, MA 02453, USA.  
Email: wingfiel@brandeis.edu



in CI simulations, improvement in word recognition begins to asymptote beyond eight channels (Faulkner et al., 2001).

Due in large measure to the spectrally limited nature of the CI signal, even when speech recognition is successful (as measured in the clinic by the ability to repeat words [e.g., CNC word lists; Luxford, 2001; Peterson & Lehiste, 1962] and short sentences [e.g., AzBio sentences; Spahr et al., 2012]), this success may not be matched in the presence of the perturbations and challenges as may be encountered in everyday listening. One particularly interesting perturbation is time-compressed speech, such as the auditory “small print” disclaimers that accompany TV and radio advertisements (Ji et al., 2013).

The most common method of time compression of speech is uniform compression using the sampling method. In this method, a speech processing algorithm periodically deletes small segments of speech and silent periods from the speech recording, abutting the remaining segments. The result is speech delivered in less than its original playing time, while leaving the original pitch (F0) unchanged, as well as maintaining the relative prosodic cues of inter- and intra-word timing patterns present in the original utterance that are important for speech understanding at the linguistic level, such as the relative lengthening of clause-final words and silent pauses that can signal that a clause boundary has been reached (Amichetti et al., 2021; Hoyte et al., 2009; Nooteboom, 1997; Shattuck-Hufnagel & Turk, 1996). Although time compression has a detrimental effect on comprehension and recall of speech by individuals with normal hearing, as might be expected, the effect is dramatically increased when combined with the degraded signal of a CI (Ji et al., 2013).

For normal-hearing adults, small degrees of time-compression of meaningful sentences (e.g., reproduction to 80–85% of the original playing time) generally allow for good comprehension. Increasing compression beyond this point produces a progressive decrease in speech comprehension. Depending on the linguistic complexity of the materials, the original speech rate, and the articulatory clarity of the speaker, major deficits in comprehension can be expected when a sentence is compressed to less than 50% of the original playing time (Wingfield et al., 2003, 2006).

An important feature of work conducted with normal-hearing adults has been a demonstration that comprehension and recall of time-compressed speech can be rescued by the insertion of lengthened silent periods at linguistically salient points (e.g., clause and sentence boundaries), within a speech sample to allow listener’s processing to “catch up” while listening to the compressed speech (Overman, 1971) at least so long as the degree of time compression is not too great (Wingfield et al., 1999). The placement of pauses is critical, such that the insertion of pauses at random points unrelated to the clausal structure yields little or no improvement in recall (Wingfield et al., 1999).

Even with the insertion of pauses at linguistically salient points it would seem intuitively likely that extremely high

degrees of time compression, especially if combined with the limited spectral richness of the CI signal, would result in a situation in which no amount of effort or availability of added processing time would allow speech to be understood. For such a situation, we borrow the term *data-limited process* from Norman and Bobrow (1975) to describe a case where no additional time or resources can rescue comprehension. By contrast, Norman and Bobrow (1975) employ the term *resource-limited process* as one in which performance will be limited only by the cognitive resources a listener is able or willing to apply to the task.

In the case of CI users, this distinction takes on not only theoretical but practical importance. This distinction is the major focus of the present study. We operationalize the threshold between a data-limited process and resource-limited process as the point on the continuum at which no benefit would be received from aiding processing time. In this regard, we also take advantage of the association between pupil dilation and cognitive effort (Zekveld et al., 2018) to examine the pattern of cognitive effort as the process shifts from a resource-limited process to one approaching a data-limited state.

It should be noted that in much of the cognitive literature the terms *effort* (often modified as *listening effort*, *cognitive effort*, *attentional effort*, or more generally, *processing effort*, depending on the authors and context) and *resources*, have often been used interchangeably. We follow Kahneman (1973) and Pichora-Fuller et al. (2016) in defining *effort* as the intentional allocation of attentional or cognitive resources to the accomplishment of a particular task or goal (see Peelle, 2018, for a detailed discussion of effort and neurocognitive resources).

A number of approaches have been taken to the measurement of processing effort in cognitive studies, many of which have recognized limitations. For example, dual-task studies have seen frequent use as a measure of processing effort based on the assumption that the effort needed for success on a primary speech task will be revealed in a performance decrement on a concurrent unrelated task (e.g., Naveh-Benjamin et al., 2005; Tun et al., 2009). Although useful (e.g., Pals et al., 2013), dual-task studies can be prone to trade-offs in the momentary attention given to each task that may complicate interpretation (Hegarty et al., 2000). Ratings of subjective effort, although having potential ecological interest, do not always correlate with dual-task results. This suggests that subjective ratings of effort and behavioral deficits in dual-task studies may measure different features of the effort construct (Abdel-Latif & Meister, 2022). Importantly, like dual task measures, effort ratings are an inherently off-line measure (see the review in McGarrigle et al., 2014).

By contrast with the above measures, pupillometry (the measurement of task-related changes in the size of the pupil of the eye) can serve as an objective, online physiological index of perceptual or cognitive effort. A special value of

the task-evoked pupillary response (TEPR) is that it can be measured independently, and does not interfere with, task performance. The use of pupillometry as a measure of effort is based on the finding that, in addition to a pupillary response to changes in ambient light and emotional arousal (Kim et al., 2000), the pupil has been shown to dilate while a person is engaged in tasks requiring perceptual or cognitive effort (Beatty & Lucero-Wagoner, 2000; Kahneman & Beatty, 1966).

In terms of speech comprehension, there is extensive literature reporting an increase in pupil dilation while individuals are processing sentences that increase in linguistic complexity (Ayasse & Wingfield, 2018; Just & Carpenter, 1993; Piquado et al., 2010). Increases in pupil size have also been reported while individuals are processing perceptually challenging speech due to acoustic masking, or mild-to-moderate hearing loss (Ayasse et al., 2017; Kuchinsky et al., 2014; Zekveld et al., 2011). Relevant to our present interests, such increases in pupil size have also been reported while processing spectrally impoverished speech associated with CIs (Winn & Teece, 2021; see also Winn et al., 2015).

## The Present Experiments

Two experiments are reported that examine the effects of time compression and time restoration on the recall of spoken narratives, and the associated differences in processing effort. Experiment 1 was conducted with normal-hearing young adults to serve as a best-case baseline to place in context the performance of the CI users tested in Experiment 2.

Our focus in these experiments departs from the more usual practice in CI outcome studies of testing the repetition of single words and short sentences. Instead, we used materials more closely matching the challenges of everyday communication. In this case, our behavioral measure was accuracy in recalling the content of multisentence spoken narratives.

Testing memory for narratives as opposed to, for example, tests of sentence repetition, is useful for three reasons. The first is pragmatic: in everyday life the listener, whether with normal hearing or one who hears via a CI, ordinarily wishes not only to perceive and understand narrative discourse but also to remember what they have heard. Second, there is a clear theoretical framework on which to build our understanding of resource-limited effects in narrative recall. Specifically, models such as that of Kahneman (1973) or the Framework for Understanding Effortful Listening (Pichora-Fuller et al., 2016) characterize cognitive processing as a limited resource system. Thus, if individuals show poorer memory for a narrative in the presence of acoustic challenge relative to easy-to-understand speech, this suggests that the increased demand for cognitive resources during listening in the former case will result in fewer resources available for effective memory encoding.

A final feature of using narrative passages as stimuli is their conventional structure of a setting, theme, and

resolution, that tells a story unfolding in time (Mandler, 1978). This narrative structure can allow the listener to make predictions about the kinds of information that will be presented at each point. This attribute characterizes much of everyday conversational speech. It may also have implications for individuals' distribution of processing effort as a narrative unfolds in time.

In Experiment 1, we used noise-band vocoding with 10 and 6 vocoder channels to simulate a range of spectral richness as might be available to CI users. In doing so, it is recognized that vocoding does not capture all features of the CI signal, such as perceptual "smearing" and frequency mismatch as may be experienced by postlingually deaf adults after implantation (Svirsky et al., 2021). Vocoding has nevertheless received wide use in simulation studies as an approximation to CI hearing that allows for control of the spectral information available to the listener.

Based on extant literature, we expect Experiment 1 to show progressively reduced recall accuracy for narratives heard with 10- and 6-channel vocoded speech relative to unprocessed clear speech. Additionally, we expect reduced accuracy for time-compressed speech relative to speech delivered at a normal speech rate. We also expect that, compared to nonvocoded time-compressed speech, there will be a recall advantage when time-compressed speech has had the lost processing time restored by inserting silent pauses at linguistically salient points in the narrative.

Central to this latter question is the extent to which the postulated recall decrement due to time compression can still be rescued by providing a space for additional processing time at syntactically salient points in a spoken narrative when the number of vocoder channels (spectral richness) with which the speech is delivered is varied. The point at which reduced recovery with fewer vocoder channels is observed could thus be seen as reflecting the boundary between resource-limited and data-limited processes. Experiment 2 was conducted with adult CI users tested for recall of the same time-compressed and time-restored materials, albeit without vocoding.

In addition to the behavioral measures of passage recall, in both experiments, pupil size was tracked throughout the course of each trial as participants were hearing a narrative passage and preparing to give their spoken recall. Although pupillometry has seen wide use as a measure of listening effort for sentence-length materials, there have been few studies that have tracked pupil size across an entire narrative passage.

## Experiment 1

### Method

**Participants.** Participants were 27 young adults (12 males and 15 females), with ages ranging from 18 to 26 ( $M = 19.72$  years,  $SD = 2.37$ ). The audiometric evaluation was carried

out for each participant using a Grason-Stadler AudioStar Pro clinical audiometer (Grason-Stadler, Inc., Madison, WI) using standard audiometric procedures in a sound-attenuating testing room. The participants had a mean better-ear pure tone threshold average (PTA) across 0.5, 1, 2, and 4 kHz of 8.00 dB HL ( $SD = 3.31$ ), placing them in a range considered to be clinically normal hearing for speech (Katz, 2002).

**Vocabulary knowledge.** Because our study involves linguistic materials we wished to ensure that potentially poorer performance on the recall task by CI users in Experiment 2 not be due to differences in linguistic ability between the CI users and the normal-hearing young adults. It has been shown, in this regard, that vocabulary can serve as a reliable proxy for general linguistic ability (Anderson & Freebody, 1981). For this purpose, all participants received a 20-item version of the Shipley vocabulary test (Shipley et al., 2009). The Shipley is a written multiple-choice test in which the participant is asked to indicate which of six listed words has the same or nearly the same meaning as a given target word. The group's mean score was 13.41 ( $SD = 2.10$ ).

**Working memory span.** Working memory span was assessed with the Reading Span (R-Span) task modified from Daneman and Carpenter (1980; see Stine and Hindman, 1994). In this task, participants read sets of sentences ranging in succession from one to five sentences. After each sentence, the participant says whether the sentence is true or false. After all of the sentences of the set have been presented, the participant's task is to recall the last word of each sentence in the order they were heard. Participants received three trials for any given number of sentences. The sets were presented using a stair-step procedure, where the number of sentences increased from one to five. All participants heard up to the length of three sentences, and the task stopped when a participant was unable to get at least two out of three trials at a particular set size. As such, the longer trials (length of four and five sentences) were presented only if participants were successful with the shorter trials. McCabe et al.'s (2010) scoring method was used, in which the working memory score is calculated as the total number of trials in which all sentence-final words were recalled correctly in the correct order. The group's mean R-Span score was 9.56 ( $SD = 3.03$ ).

**Executive Function/Inhibition.** Participants received the Flanker task (Eriksen & Eriksen, 1974) as an index of the inhibition component of executive function. Participants saw a fixation cross-centered on a computer screen for 500 ms. This was followed by a horizontal array of five equally sized and spaced arrows for 800 ms. Participants were instructed to attend to the central arrow and to ignore the four flanking arrows. The task was to press a left key for a left-facing central arrow and a right key for a right-facing central arrow. The flanking arrows either all point in the same direction as the target arrow (congruent trials) or in the opposite direction (incongruent trials). Participants received a total of 60 test trials: 30 congruent, and 30

incongruent. The conflict cost was measured as the mean latency to keypresses in the incongruent trials minus the congruent trials. The participants' mean conflict cost was 53.16 ms ( $SD = 46.13$ ).

All participants reported themselves to be native speakers of American English and no known medical issues that might interfere with the ability to perform the experimental tasks. Written informed consent was obtained from all participants according to a protocol approved by the Institutional Review Boards of Brandeis University and the New York University School of Medicine.

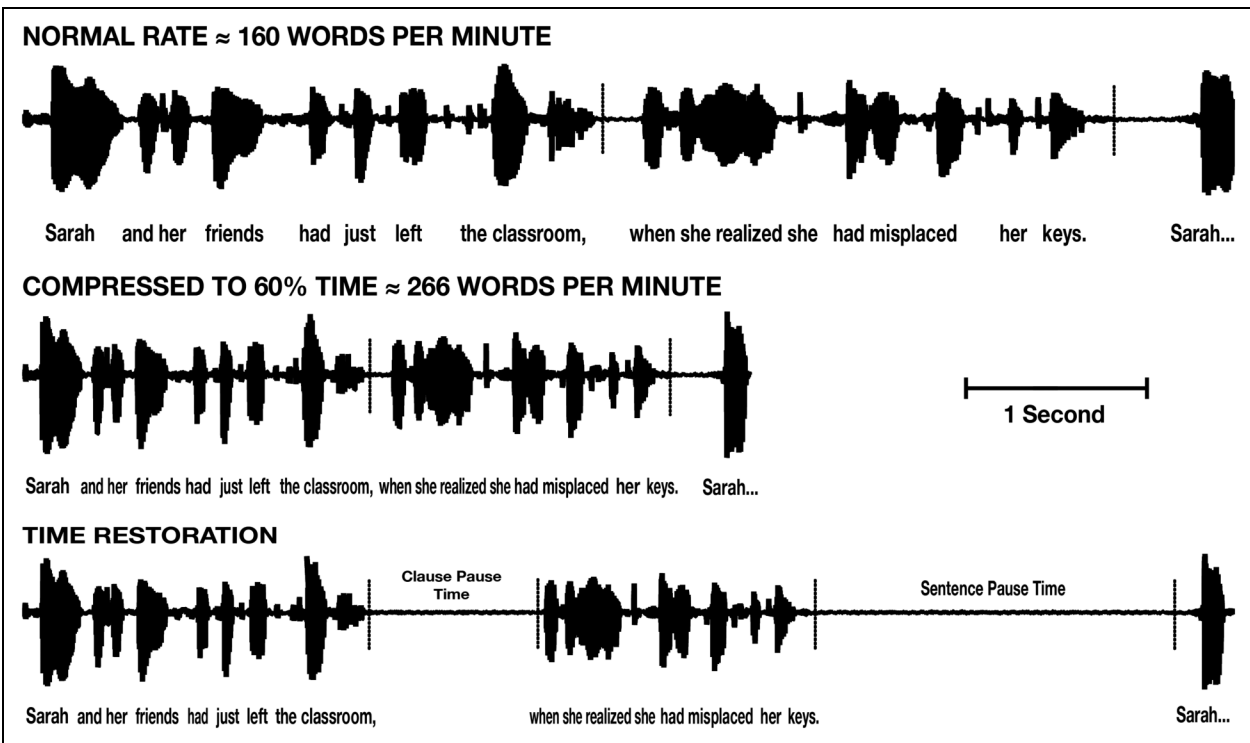
**Stimuli.** The test stimuli consisted of twenty-seven 67- to 97-word narrative passages ( $M = 81.80$  words). The narratives were brief vignettes representing simple stories. For instance, one passage discussed an American student with a summer job at a Scottish museum who learned more from talking with coworkers than from books from the library, while another discussed a high school athlete who won a swimming scholarship to a university where she trained hard and exceeded her best high school racing times.

The narratives contained four to five sentences and five within-sentence major clause boundaries. The passages were recorded by a male speaker of American English using natural prosody onto computer sound files using Sound Studio v2.2.4 (Felt Tip, Inc., New York, NY) that digitized (16-bit) at a sampling rate of 44.1 kHz. Recordings were equalized within and across narratives for root-mean-square (RMS) intensity using Praat (Boersma & Weenink, 2022). The recorded passage durations ranged from 23.44 s to 34.14 s ( $M = 28.32$  s) representing an average speech rate of 173.4 words per minute (wpm).

**Time compression.** Each of the 27 passages was time-compressed to be reproduced to 60% of their original playing time, equivalent to a mean speech rate of 288.71 wpm using Praat (Boersma & Weenink, 2022) that employs the pitch-synchronous overlap-and-add algorithm. This compression algorithm uses uniform linear compression such that both words and silent pauses were reduced equivalently to maintain the normal relative temporal pattern. There was no disturbance of pitch.

The effect of time compression can be seen in Figure 1 by comparing the original waveform of a sample sentence taken from one of the stimulus narratives (top panel) with the same sentence compressed to 60% of its original playing time (middle panel). One can see that the relative intra- and inter-word timing patterns are preserved even as the absolute durations are reduced.

**Time restoration.** A second version of each of the time-compressed passages was constructed in which silent periods were inserted into the narratives at clause and sentence boundaries. The placement of silent periods in these positions followed prior studies showing that pauses at linguistically salient points are more effective in aiding recall than placing pauses at random points with regard to the



**Figure 1.** Waveform of a sample sentence recorded at its original normal speech rate (top panel), the time-compressed version of the same sentence (middle panel), and the time-compressed version with silent periods inserted after the main clause and sentence ending to restore the lost processing time due to the time-compression (lower panel).

linguistic content (Wingfield et al., 1999). In order to respect the relative importance of the linguistic boundaries, the durations of the silent periods added at sentence boundaries were set at twice the duration of the silent periods added at clause boundaries.

The effect of time-restoration is illustrated in the bottom panel in Figure 1. One can see, for example, the longer inserted pause following the sentence-ending "...misplaced her keys," relative to the pause inserted after the clause that ends with the noun-phrase, "the classroom." As also seen in this illustration, the total duration of the silent periods was set such that the duration of the time-restored compressed passages was equal to the duration of the original un-compressed recordings (cf., Blumstein et al., 1985; Chodorow, 1979). The inserted silent periods were taken and extended from silent moments in the initial recordings to avoid any possible differences in the ambient background during speech and silent periods.

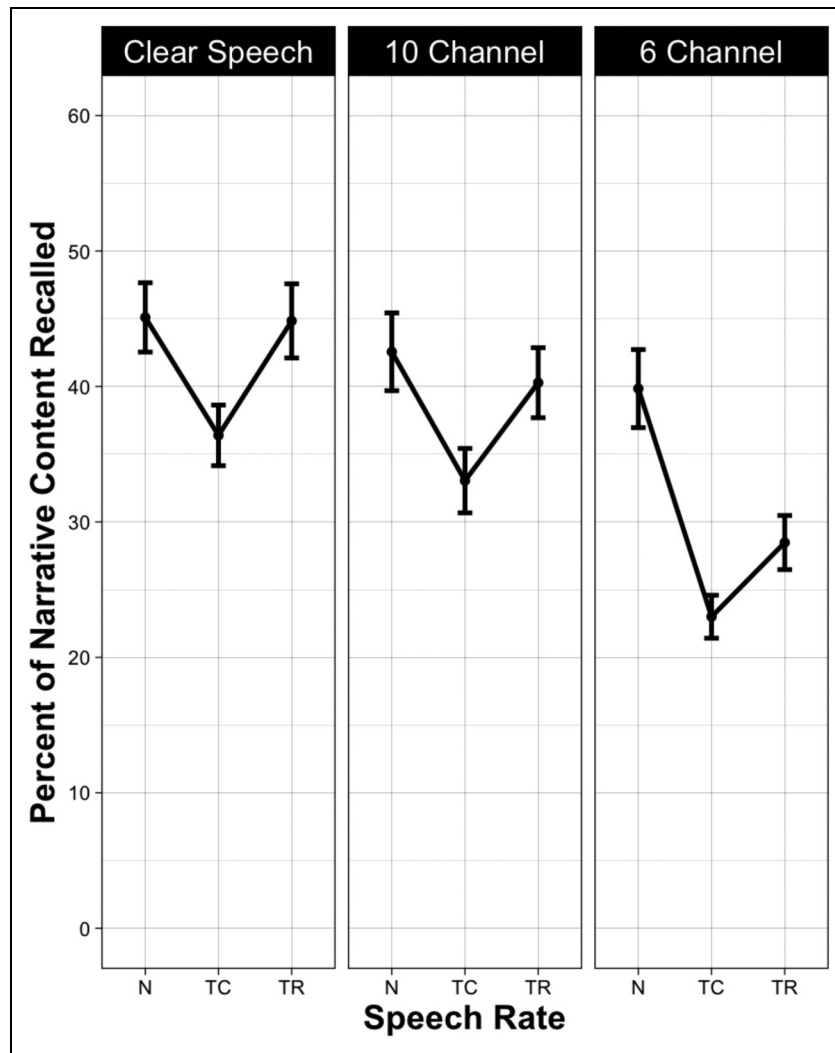
**Vocoding.** Each of the narratives in their original, time-compressed, and time-restored versions was processed to produce versions with 10 and 6 vocoder channels. Noise-band vocoding was conducted using MATLAB (MathWorks, Natick, MA) following the method described by Shannon et al. (1995). Broadband speech (141–7,000 Hz) was preemphasis filtered using a first-order Butterworth high pass filter with a cutoff frequency of 1,200 Hz (Wilson, 2004). Because it was a first-order filter, its roll-off was 6 dB/octave. This was band pass filtered

using a third-order Butterworth filter (18 dB/octave roll-off) into logarithmically spaced frequency bands, with the amplitude envelope of each frequency band extracted by Hilbert transform followed by a sixth-order Butterworth low pass filter at 300 Hz (36 dB/octave roll-off). The resultant signals for all of the bands were then recombined and low pass filtered with 7,000 Hz as the cut-off to produce the vocoded stimuli (Kong et al., 2015). The vocoded stimuli were scaled to match the RMS amplitude of the unprocessed stimuli.

### Procedures

Participants were told they would hear a series of brief narratives. Three seconds after a narrative ended, they received an on-screen visual prompt to recall aloud as much of the narrative as possible without guessing. They were further told that some of the narratives would be distorted in various ways, but they were to do their best to accurately recall what they had heard. Recall responses were audio-recorded for later transcription and scoring.

Each passage was preceded by a 2-s silent period to serve as a baseline for the pupillometry as will be described below. Each trial thus consisted of 2-s of silence for pupillary baseline acquisition, the stimulus narrative, a 3-s retention interval, and the participant's spoken recall. After giving their recall response, the participant pressed a key to initiate the next trial.



**Figure 2.** Mean percentage of narrative content (propositions) recalled from narratives presented in clear speech (left panel), with 10-channel vocoding (middle panel) and with 6-channel vocoding (right panel) when presented at their original normal speech rate (N), when time-compressed to 60% of their original playing time (TC) and when the lost processing time was restored by inserting silent periods at clause and sentence boundaries (TR). Error bars are one standard error.

Each participant heard all 27 passages: three in clear speech (unprocessed) at a normal speech rate, three in clear speech time-compressed at 60%, three in clear speech time-compressed with time-restoration, three 10-channel vocoded passages at a normal speech rate, three 10-channel time-compressed at 60%, three 10-channel time-compressed with time restoration, three 6-channel vocoded passages at a normal rate, three 6-channel time-compressed and three 6-channel time-compressed with time restoration. The particular passages heard in each of the conditions were counter-balanced across participants such that, by the end of the experiment, each passage had been heard an equal number of times in each of the nine experimental conditions.

The narratives were intermixed in the presentation, with the participants told before each passage what condition they would hear. Stimuli were presented binaurally over Eartone 3A insert

earphones (E-A-R Auditory systems, Aero Company, Indianapolis, IN) at 65 dB HL. Prior to the main experiment, participants received a three-part session to familiarize them with the sound of vocoded speech, the sound of time-compressed and time-restored speech, and the sound of speech that was both vocoded and time-compressed. Participants were also familiarized with the task of recalling these materials. None of the passages used in this familiarization session were used in the main experiment.

**Pupillometry.** Participants' moment-to-moment pupil responses were recorded throughout the experiment using an EyeLink 1000 Plus eye-tracker (SR Research, Ontario, Canada) using a standard nine-point calibration procedure. Pupil size data were acquired at a rate of 1,000 Hz, recorded using the SR Research Experiment Builder, and processed via MATLAB software (MathWorks, Natick, MA).

Participants placed their heads in a customized individually adjusted chin rest to reduce head movement and to maintain a distance of 60 cm from the EyeLink camera.

To further facilitate reliable pupil size measurements, participants were instructed to keep their eyes fixated on a continuously displayed 2 cm<sup>2</sup> black cross centrally located on a computer screen placed above the EyeLink camera. The computer screen was filled with a medium gray color to avoid ceiling or floor effects on the pupil size at baseline (Winn et al., 2018). Eye blinks were detected and removed using the algorithm described by Hershman et al. (2018) and filled by means of linear interpolation. Trials in which more than 25% of the data required interpolation were removed from analyses. This resulted in a loss of 22 trials (<5% in all conditions). The pupillometry signal was smoothed using a moving window average of 2 s and then down-sampled to 1-s bins.

Task-related pupil sizes were baseline corrected to account for non-task changes in the base pupil size as can occur across trials (Ayasse & Wingfield, 2020). This was accomplished by subtracting the mean pupil size measured over the last 1-s of the 2-s silent period preceding each sentence from the task-related pupil size measures (see Reilly et al., 2019, for data and a discussion on linear versus proportional baseline scaling). It should be noted that spontaneous pupillary fluctuations during a baseline period may add undesirable variability to the y-intercept of the subsequent trial, a potential that may be exacerbated by fewer experimental trials in each condition. A direct approach to protect against the effects of baseline fluctuations would be for the experimenter to examine pupil size in real-time and only initiate a trial when the participants' pupil size was stable (Winn et al., 2018). This was not done in the current experiment. However, examination of the data confirmed that there was no significant difference in baseline pupil size across experimental conditions ( $p = .988$ ).

**Adjusting for age differences.** When conducting pupillometry with older adults it is important to account for the smaller base pupil size and smaller dynamic range of older adults' pupil size (*senile miosis*; Bitsios et al., 1996; Guillon et al., 2016). This difference could lead one to underestimate the level of effort associated with a task by older relative to younger adults. Although all of the participants in Experiment 1 were young adults, we conducted the same adjustment process as would be used for the CI users in Experiment 2 that included older as well as younger adults.

Various methods may be considered to adjust pupillometry data for *senile miosis* when an experiment includes both young and older adults (cf., McLaughlin et al., 2022; Seropian et al., 2022; Winn et al., 2018). The method we chose was to represent task-related pupil sizes as a percentage of an individual's independently measured pupillary dynamic range (Piquado et al., 2010). In order to calculate the dynamic range of each participant's pupil, participants were presented with 60-s displays of a dark screen (0.05 fL) followed by a light screen (30.0 fL) at the beginning and end of the

experiment. The dynamic range for each participant was taken as the average of the range measures before and after the experiment. This allowed representation of the task-related pupillary response as a percentage ratio of each individual's dynamic range. This was calculated as  $(d_M - d_{min}) / (d_{max} - d_{min}) \times 100$ , where  $d_M$  was the participant's baseline-corrected pupil size at any given moment,  $d_{min}$  was the minimum constriction taken as the average pupil size over the last 10 s of viewing light screen, and  $d_{max}$  was the maximum dilatation taken as the average pupil size over the last 30 s of viewing the dark screen (e.g., Ayasse et al., 2017; Ayasse & Wingfield, 2018; see also Winn et al., 2018).

The light range adjustment is not intended to necessarily imply that there is a common circuitry between the light reflex and the response to a cognitive task (Zhao et al., 2019). This adjustment is rather an attempt to determine the *relative* maximum and minimum range of pupillary change to normalize individual differences.

## Results

**Statistical Analyses: Linear Mixed Effect Models (Recall Performance).** Narrative recall data were analyzed using linear mixed-effects models in R version 3.52 (LMEM's, lme4 package version 1.1–19; Bates et al., 2015, lmerTest package version 3.1–3; Kuznetsova et al., 2015) with participants and narratives included as random intercepts. Rate (normal rate, time-compression, time-restoration) and clarity (clear speech, 10 channel, and 6 channel) were included as factors of three levels and treated as fixed effects. A reverse selection approach was used to analyze models, starting with a maximal model including all variables and interactions. From the maximal model, likelihood ratio tests were used to contrast models and remove nonsignificant effects to find the most parsimonious model using a likelihood ratio test (anova function) (Bates et al., 2015). Reported  $p$ -values are the result of likelihood ratio tests between models unless stated otherwise.

**Narrative Recall.** Scoring was based on the percentage of propositions recalled correctly from each passage. Propositions (sometimes referred to as "idea units") were determined by representing each passage as a propositional text base using the Kintsch system (Kintsch & Van Dijk, 1978; Turner & Greene, 1978). This was instantiated using the Computerized Propositional Idea Density Rater program (Brown et al., 2008) that translates text into a detailed propositional breakdown using a Monty-Lingua (Liu, 2004) part-of-speech tagger. This tagger initially identifies all parts of speech, and then applies a set of corrective rules to determine a proposition count using IKVM software (see Frijters, 2004). Gist scoring was used in which credit was given for the recall of propositions that contained the same meaning as the original proposition (Turner & Greene, 1978).

The left panel of Figure 2 shows the mean percentage of propositions recalled when passages were heard with clear speech (i.e., nonvocoded) when presented at their original normal speech rate (N), when time-compressed to 60% of original playing time (TC), and when the time-compressed passages had the lost processing time restored by inserting silent periods at clause and sentence boundaries (TR). The middle and right panels show these data for passages heard with 10 and 6 vocoder channels, respectively.

The data shown in Figure 2 were analyzed using a linear mixed effect model in which likelihood ratio tests were used to contrast models and remove nonsignificant effects to determine the most parsimonious model. The final model contained Clarity, Rate, the interaction of Clarity and Rate, and Shipley Vocabulary as fixed effects with participants and narratives as random intercepts. The slope of Rate was allowed to vary by participants. The outcome is given in Table 1.

As would be expected, the accuracy of narrative recall was greatest when the narratives were heard with clear speech and progressively declined when heard with 10 and 6 channels, confirmed by a significant effect of stimulus clarity ( $p < .001$ ). The appearance of a decline in recall with time-compressed and time-restored speech was confirmed by a significant effect on speech rate ( $p < .001$ ).

Although overall recall accuracy differed between the three clarity conditions (clear speech, 10-channel vocoding, and 6-channel vocoding) it can be seen in Figure 2 that all three conditions showed a similar V-shaped function. However, when clarity was progressively reduced, the detrimental effect of time compression increased, and the benefit of time restoration decreased. Consistent with this observation, there was a significant Stimulus Clarity  $\times$  Rate interaction ( $p = .001$ ).

A series of pairwise planned comparisons confirmed that time compression had a detrimental effect on recall in all three stimulus clarity conditions ( $p < .001$  in all cases). As suggested by visual inspection of the left panel of Figure 2, when passages were heard with clear speech, time-restoration (TR) raised the level of recall accuracy for otherwise time-compressed speech 95% of the way back to baseline, a

level that did not differ significantly from the accuracy level for passages presented with their original normal (N) speech rate ( $p = .874$ ).

When heard with 10-channel vocoding (middle panel), time-restoration raised performance 76% of the way back to the normal speech-rate baseline ( $p = .002$ ). The appearance of a gap between the accuracy level for passages presented at the original normal speech rate and the compressed passages with time-restoration was not significant ( $p = .125$ ).

It can be seen that 6-channel vocoding (right panel) still allowed significant recovery for time-compressed speech with time-restoration when compared with time-compression without restoration ( $p = .005$ ). However, for this level of vocoding, time-restoration only brought the level of recall accuracy 31% of the way back to normal speech rate. This resulted in a significant difference still remaining between time-restored passages and the recall level when passages were presented at the original normal speech rate ( $p < .001$ ).

**Predictors of Individual Differences in Recall.** The effects of age, hearing acuity (PTA), Shipley vocabulary, R-Span, and Flanker scores on recall accuracy were examined in the mixed effects model. Shipley vocabulary score was a significant predictor of overall recall accuracy ( $p = .019$ ). For these normal-hearing young adults, none of the remaining measures (hearing acuity, age, R-Span, or Flanker) had a systematic effect on recall accuracy.

**Statistical Analyses: Generalized Additive Mixed Models (Pupillometry).** Generalized Additive Mixed Models (GAMMs) were used to model pupillometry data (van Rij et al., 2019). GAMMs are used to estimate nonlinear relationships (“smooths”) over a time course between multiple variables and covariates. GAMMs estimate fixed parametric terms between variables, as is done in LMEMs. GAMMs also estimate smooth terms that use smoothing splines to model nonlinear effects. These models have two advantages over growth curve analysis (GCA; Mirman, 2017). First, GAMMs make no assumption of symmetry around the midpoint of the pupillometric time series. Second, GAMMs can account for the within-trial autocorrelation between time points that may inflate Type 1 error rates in GCA (Huang & Snedeker, 2020). GAMM modeling was executed using the mgcv (Wood, 2017) and itsadug (van Rij et al., 2020) packages in R Studio.

**Pupillometry.** The pupillometry data are shown in Figure 3. In all cases, pupil sizes are adjusted for baseline and presented as a percentage of an individual’s pupillary dynamic range. Although the time restored and original speech rate conditions occupy the same duration, the time-compression condition does not. This is illustrated in Figure 3(A) which shows the mean adjusted pupil sizes as participants listened to narratives presented with the original normal speech rate, when

**Table 1.** Log-Likelihood Model Comparisons for Linear Mixed Effects Analysis of Normal-Hearing Young Adults’ Narrative Recall.

Predictor	$\chi^2$ <sup>a</sup>	$df$ <sup>b</sup>	$p$ <sup>c</sup>
Clarity	140.30	6	<b>&lt;.001</b>
Rate	65.56	6	<b>&lt;.001</b>
Clarity $\times$ rate	23.74	4	<b>&lt;.001</b>
Shipley vocabulary	5.53	1	<b>.019</b>

Notes.

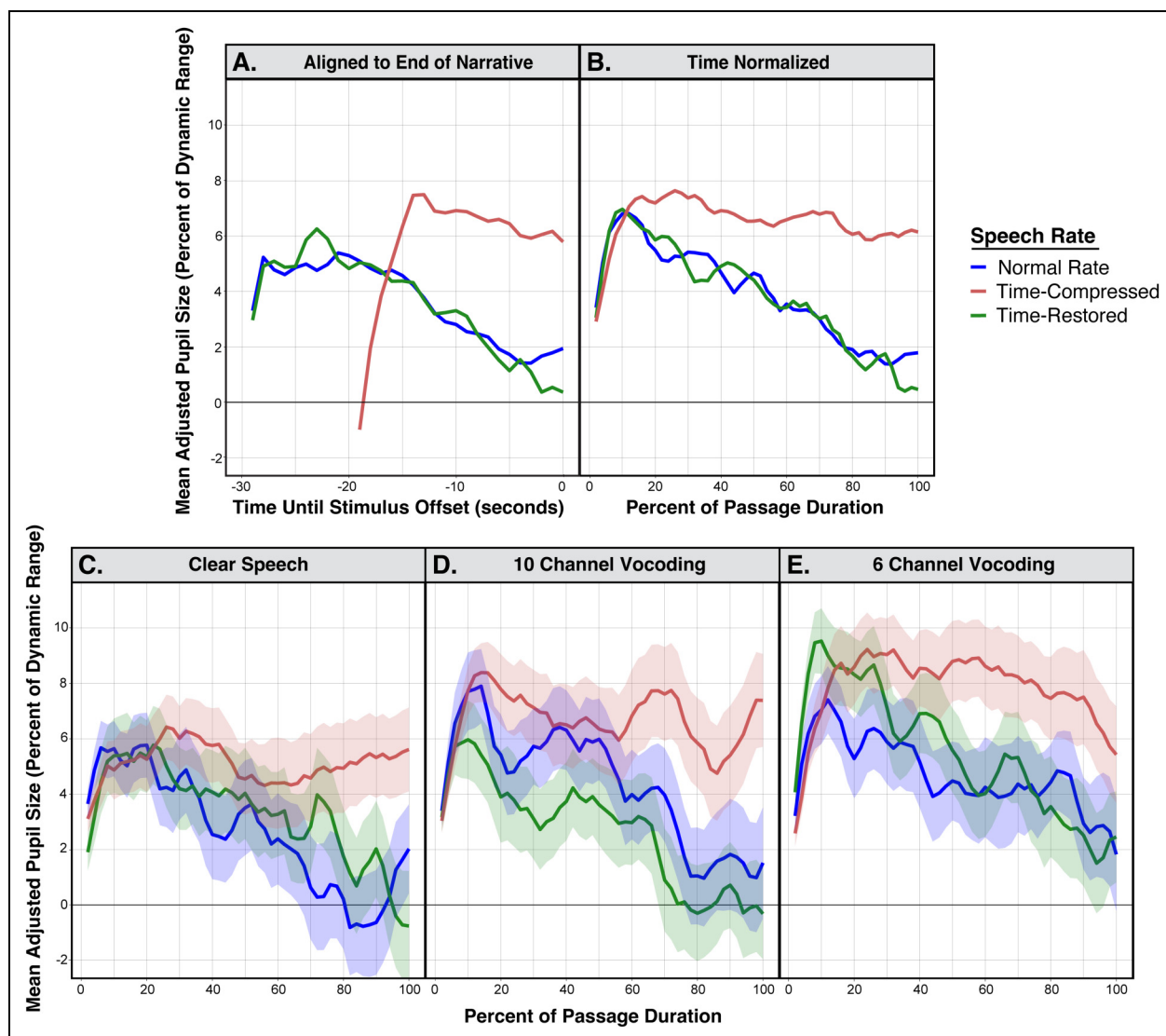
<sup>a</sup> $\chi^2$  value for comparisons of each step of the model.

<sup>b</sup>Degrees of freedom for the  $\chi^2$  test.

<sup>c</sup>Value of  $p$  reflects the significance of the change in model fit at each step of the model.

Significant values of  $p$  are indicated in bold.





**Figure 3.** Top two panels show the time course of mean pupil dilations, adjusted to baseline and to each individual's dynamic range while normal-hearing young adults listened to narrative passages. Data are shown for narratives that were time-compressed, presented at a normal speech rate, and in the time-restoration condition. Data are collapsed across vocoder conditions. (A) shows pupillary responses aligned to the end of the narrative passage, with time 0 indicating the end of the passage. (B) shows these data as a percentage of the duration of each passage. The three bottom panels show the mean adjusted pupil dilations separately for clear speech (C), 10-channel vocoding (D), and 6-channel vocoding (E) as a percentage of the passage duration. Error ribbons are one standard error.

time-compressed, and when time-compressed with time-restoration collapsed across clarity conditions. These data are aligned to the end of the narrative passage and trimmed to the mean stimulus length. Figure 2(B) shows these same data, this case time-normalized by presenting pupil sizes as a percentage of the duration of each passage. All remaining pupillometry figures will be shown time-normalized as a percentage of passage duration.

Two major features appear by visual inspection of the time series data shown in both Figure 3(A) and Figure 3(B). The first is that participants' adjusted pupil sizes progressively decline toward the latter part of the passages when heard at

the original normal speech rate, and when time-compressed passages were heard with time-restoration. The second major feature observed in Figure 3(A) and (B) is that the time-restoration condition reduces the level of processing effort indexed by pupil dilation to a level close to that for passages heard at a normal speech rate. These patterns are the same whether or not time normalization is employed.

The lower three panels show the data in Figure 3(B) decomposed to separate the pupillometry data for time-compressed speech, time-restored speech and the original normal speech rate for clear speech (Figure 3(C)), 10 channel vocoding (Figure 3(D)), and 6 channel vocoding

(Figure 3(E)). Despite increased variability in the time series curves due to fewer data points per condition, in all cases, the same general pattern is seen, in which time compression increases the level of processing effort that is then rescued by time-restoration. Within this trend one also observes a general increase in pupil size as spectral clarity was reduced over 10- and 6-vocoder channels.

The pupillometry data shown in Figure 3(C), (D), and (E) were analyzed using the above-described generalized additive mixed models. The final model included Clarity (clear speech, 10 channel vocoded, 6 channel vocoded) and Rate (normal rate, time-restored, time-compressed) as fixed effects, with participants and narratives as factor “smooths” (random effects). Autocorrelation was calculated between model residuals and factored into the final model thus allowing an autoregressive error parameter of the autocorrelation between time points in each trial. Inspection of the final model showed that specifying this parameter appropriately factored out the autocorrelation between residuals. An additive model was thus chosen because after addressing autocorrelation there was no significant interaction between Clarity and Rate, and the Akaike Information Criterion (AIC) value was lower for the final additive model when compared to the interactive model (AIC difference =  $-8.27$ ) suggesting that the simpler model was preferred. The GAMM summary results for the final model are shown in Table 2.

Inspection of Table 2 confirmed a significant effect of time compression on pupil dilation relative to the pupillary response to narratives heard at a normal speech rate ( $p < .001$ ). Inserted pauses within time-compressed speech reduced processing effort to levels comparable to that of normal speech rate. This was reflected in the absence of a significant difference in the pupillary responses to the time-restored and normal speech rate conditions ( $p = .604$ ). The pupillary response for the more severely degraded 6-channel vocoding was significantly larger than those observed for ( $p < .001$ ). The observed trend toward larger pupil sizes when speech was heard with 10-channel vocoding relative to clear speech was not significant ( $p = .149$ ).

*Time Regions of Significant Differences.* Table 2 provides a convenient summary of the overall effects of speech rate

**Table 2.** Generalized Additive Mixed Model Summary of Young Adults’ Pupillary Responses.

Parametric Coefficients	Estimate	SE	t-Value	p-Value
Intercept (normal rate: clear speech)	3.04	1.13	2.70	.007
Time-restored	-0.33	0.62	-0.52	.602
Time-compressed	2.31	0.62	3.70	<.001
10-Channel vocoding	0.94	0.65	1.44	.149
6-channel vocoding	2.32	0.65	3.56	<.001

and noise band vocoding on the pupillary response to the narrative passages. It is also important, however, to examine non-linear effects that may reveal specific points in time when significant differences between conditions occurred.

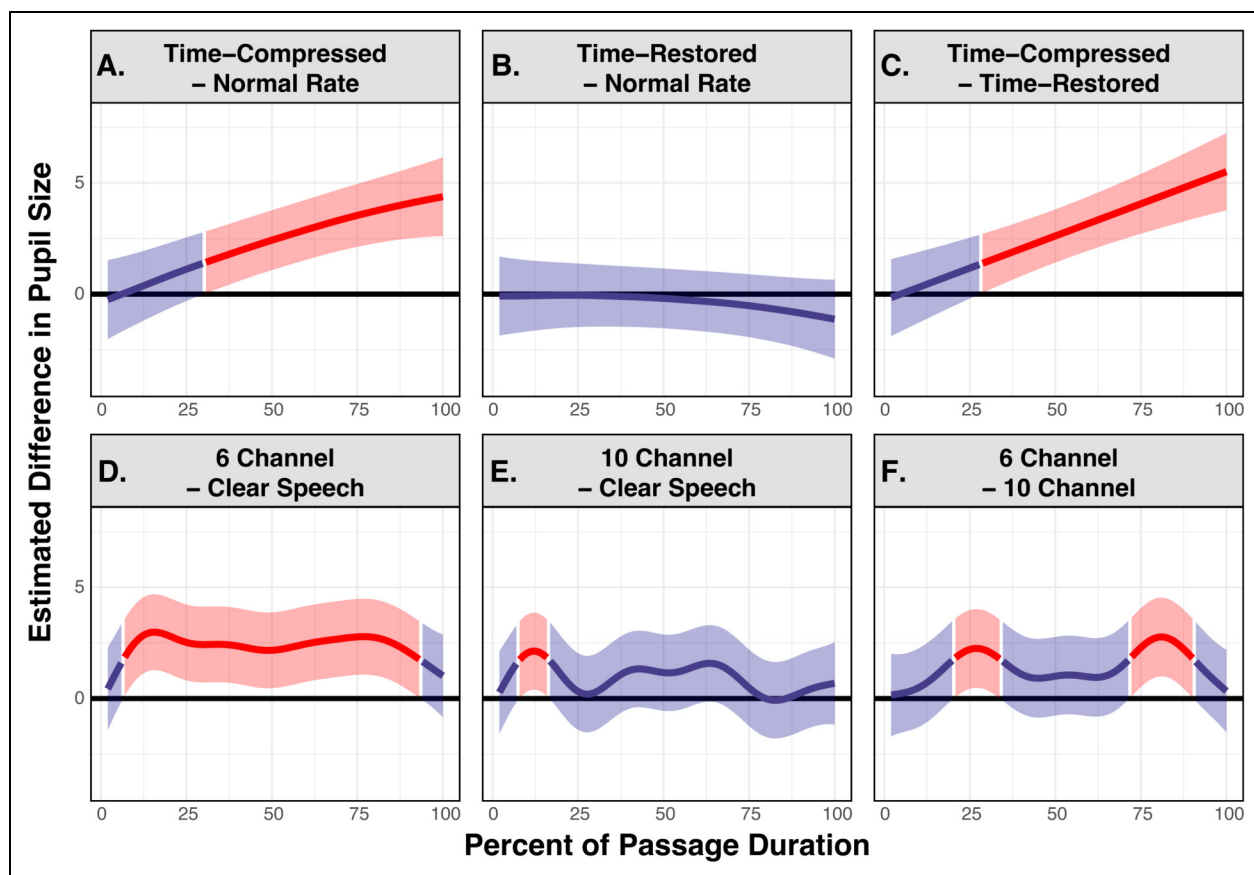
The panels in Figure 4 visualize the time series of GAMM-estimated differences in pupil size. In all panels, the y-axis represents the model-estimated difference in dynamic range adjusted pupil size between conditions, with the 0-point representing no difference. As such, the time points where the shaded regions do not overlap with the bold line at the 0 point on the y-axis represent a moment in the time series where the difference between conditions is statistically significant. We follow Cychoz et al. (2023) in showing the significant regions of the time series in red and the nonsignificant regions in blue using the tidymv package (Coretta, 2022).

The top three panels of Figure 4 display the difference curves estimated by the GAMM between participants’ pupil traces for passages when heard with normal speech rate subtracted from those heard with time-compression (Figure 4(A)), participants’ pupil traces for passages heard with at a normal speech rate subtracted from those heard with time-restoration (Figure 4(B)), and participants’ pupil traces for passages heard with time-restoration subtracted from those heard with time-compression (Figure 4(C)).

The effects of vocoding are shown in the bottom panels, displaying the difference curves estimated by the GAMM for participants’ pupil traces when passages were heard with clear speech subtracted from those heard with 6-channel vocoding (Figure 4(D)), participants’ pupil traces when passages were heard with clear speech subtracted from passages heard with 10-channel vocoding (Figure 4(E)), and participants’ pupil traces when passages were heard with 10-channel vocoding subtracted those heard with 6-channel vocoding (Figure 4(F)).

For the normal hearing young adults seen in Figure 4(A), pupil size when participants were listening to time-compressed speech was significantly larger than pupil size when participants listened to passages heard at a normal speech rate for 69.29% of the passages’ durations, with these significant differences starting at 30.71% after stimulus onset and continuing until the passages’ offset. Figure 4(B) confirms that there is no statistically detectable difference in pupil size when participants listened to passages heard with time-restoration as compared with the normal rate at any point of the passages’ durations. Consistent with this, Figure 4(C) shows that pupil size for passages heard time-compression was significantly larger than pupil size for passages heard with time-restoration for 71.27% of the passages’ durations, with that region of significance beginning at 28.73% after stimulus onset and lasting until the passages’ offset.

As can be seen in Figure 4(D), participants also had significantly larger pupil size when listening to passages that were



**Figure 4.** Generalized additive mixed model predicted difference curves of adjusted pupillary responses when young adults listened to narrative passages. The top panels show contrasts in predicted pupil size between conditions of time-compression versus normal rate (A), time-restored versus normal rate (B), and time-compressed versus time-restored (C). The bottom panels display contrasts in predicted pupil size between conditions of 6-channel vocoding versus clear speech (D), 10-channel vocoding versus clear speech (E), and 6-channel vocoding versus 10-channel vocoding (F). A higher estimated difference is an indication of the greater difference between compared conditions, with regions highlighted in red indicating time points where there is a statistically significant difference. Shaded ribbons are 95% confidence intervals.

heard with 6-channel vocoding as compared to clear speech for a total of 86.12% of the passages' durations, in the range from 6.95% through 93.07% of the passage' durations. In contrast, in Figure 4(E), pupil size when participants were listening to passages heard with 10-channel vocoding was only significantly larger than pupil size when participants were listening to clear speech for 7.92% of the passage. This difference occurred in a brief window toward the beginning of the passage from 7.94% through 15.86% of the passages' durations. Finally, when comparing the two vocoder conditions in Figure 4(F), pupil size was significantly larger when participants listened to a passage with 6-channel vocoding than when they listened to a passage heard with 10-channel vocoding for 30.69% of the passage. This difference occurred in two time windows from 20.81% to 33.68% and 72.28% to 90.10% of the passage's durations.

## Discussion

Experiment 1 replicates two well-known effects relating to recall of time-compressed clear speech by young adults with age-normal hearing. First, accelerating the rate of speech had a significant detrimental effect on speech recall (Foulke, 1971). Second, restoring the lost processing time at linguistically salient points in the time-compressed passages allowed recall accuracy to *return* to a level equivalent to the participant's baseline accuracy for normal, uncompressed speech (Aronson et al., 1971; Overman, 1971). Both are robust findings that have been replicated in older adults with age-normal hearing, albeit with older adults having a differentially greater recall decrement with time-compression relative to young adults (Gordon-Salant & Fitzgibbons, 1993; Konkle et al., 1977; Sticht & Gray, 1969), and a less complete recovery with time-restoration (Wingfield et al., 1999).

The use of noise-band vocoding added to these findings in showing that, although both 10-channel and 6-channel vocoding produced a V-shaped pattern of effects of time-compression and time-restoration, the detrimental effects of time-compression—most notably for 6-channel vocoding—were larger, and recovery with time-restoration smaller, than was seen for nonvocoded clear speech.

We can attribute the large effect of time compression, and the modest level of recovery with time-restoration in the 6-channel vocoding condition, to the combined effects of the severely reduced spectral richness of 6-channel vocoding, and a limit on the rate at which this degraded signal can be perceptually processed and encoded in memory by the listener. It is also likely that 6-channel vocoding combined with time compression may lead to an intelligibility decline at the lexical or sub-lexical level (Heiman et al., 1986). Within the framework of Norman and Bobrow (1975), we may thus characterize performance with time-compressed speech heard with 6-channel vocoding as approaching a state of data-limited processing, and hence beyond full recovery even with additional time and effort. That is, one may expect to see poor recall of perceptually challenging speech rescued by allowing the listener additional processing time, but only to the extent that there is sufficient input data on which to operate.

It has been shown that when hearing lists of digits for recall, pupil size progressively increases with the incremental arrival of list items, commonly interpreted as reflecting the build-up of the memory load prior to recall (Granholm et al., 1996; Kahneman & Beatty, 1966; Peavler, 1974). A similar progressive increase in pupil dilation has been shown over the course of hearing a sentence presented for recall (Amichetti et al., 2021; Piquado et al., 2010). As we have seen, however, this was not the case for the narrative passages. Rather, in the case of full narratives, as opposed to a single sentence, our data show that, following an initial increase in pupil size in the early part of a passage, for both normal speech rate and the time-restoration condition, there was a progressive *decrease* in adjusted pupil size as more and more of a narrative was heard.

One possible interpretation is that this decline was associated with a *facilitation effect*, in which there was a reduced need for effort-related processing resources as listeners developed a progressive understanding of the narrative structure, with concomitant ease of integrating new information.

We have taken the term, *facilitation effect*, from the reading-time literature, where the term is used in reference to the frequent finding that readers tend to spend shorter dwell times on the words of a prose passage toward the end of a passage than at the beginning of the passage (Ferreira & Henderson, 1990; Stine, 1990; Stine et al., 1996). This pattern has been interpreted as reflecting the reader's development of an understanding of the meaning of the text that allows the reader to more easily integrate the new information with what has been read to that point

(Gernsbacher, 1990; Haberlandt & Graesser, 1989; Rayner et al., 1989). We suggest that an analogous facilitation effect operated while participants listened to our spoken narratives. It is thus possible that, in this case, the ease of processing the incoming speech as a narrative unfolded in time was reflected in a reduced need for processing effort, indexed here by a progressive reduction in the size of the pupillary response over the course of the passage. The observation of a decline in pupil size over the course of a narrative passage that is less pronounced may be indicative of a failure to develop an over-arching understanding of the structure and meaning of the narrative, and/or a failure to take advantage of this knowledge.

It should be noted in this regard that it is not uncommon to see a general decline in pupil size over the course of an experiment and in tasks that do not involve narrative comprehension (cf. Ayasse & Wingfield, 2020; McGarrigle et al., 2021; Zhao et al., 2019), often interpreted as a decline in arousal. We cannot deny the possibility that this may underlie, at least in part, the decline we observe over the course of a single narrative. We suggest that further research is needed to dissociate potentially co-occurrent effects of declining levels of arousal and declining levels of task-related processing effort as contributors to such observed declines in pupil size.

Within these overall effects of time-compression and time-restoration on relative pupil size, there was a general trend toward 10-channel and 6-channel vocoding to be accompanied by a progressive increase in pupil diameter relative to clear nonvocoded speech. This pattern is consistent with the pupillary response as an indicator of the need to expend additional effort when the listener is confronted with perceptually challenging speech stimuli (Ayasse et al., 2017; Kuchinsky et al., 2014; Winn & Teece, 2021; Winn et al., 2015; Zekveld et al., 2011).

Experiment 1 offers findings preliminary to our exploration of the effects of time-compression on narrative memory by CI users in the form of an “existence proof” for the effects of time-compression and time-restoration, and the further effects when the spectral richness of the speech was reduced by noise-band vocoding.

In Experiment 2, we used the same materials and procedures as employed in Experiment 1, with the exception that the narratives would be presented without vocoding and that the participants would be users of cochlear implants rather than normal hearing listeners.

## Experiment 2

### Method

**Participants.** The participants in Experiment 2 were 15 CI users, four men and 11 women, ranging in age from 28 to 78 ( $M = 57.33$  years). Of the 15 participants, 13 had bilateral implants. Of the two participants who had unilateral implants, one had a profound hearing loss from 125–8,000 Hz in their

un-implanted ear as indexed by pure tone thresholds, and one had hearing thresholds in the moderate range at 125 Hz sloping to severe at 500 Hz and profound at 1,000–8,000 Hz.

The CI users had a mean Shipley vocabulary score of 15.64 ( $SD = 1.82$ ), which was as a group superior to that of the normal-hearing young adults in Experiment 1,  $t(39) = 3.38$ ,  $p = .001$ . There was no significant difference between the CI users and the normal-hearing young adults in Experiment 1 in either mean R-Span ( $M = 8.38$ ;  $SD = 3.30$ ;  $p = .273$ ), or mean Flanker score ( $M = 68.12$  ms;  $SD = 33.32$ ;  $p = .291$ ).

A summary of the CI users' demographic information is given in Table 3, ordered by participant age. The first two columns give the participants' age and sex, and the next three columns give their implant information. This is followed by the etiology of the hearing loss where known and the years of experience with their CIs. The final two columns give participants' percent correct CNC-30 scores and the percent correct scores for repetition of AzBio sentences in noise.

It should be noted that the CNC-30 scores, used here as an index of word recognition ability (Luxford, 2001), are on average 22% lower than the more commonly used CNC recordings (Skinner et al., 2006). We have used the CNC-30 version because it is less likely to result in ceiling effects, an advantage for clinical research. As is common among CI users, there was a wide range in the effectiveness of word recognition with CNC-30 word-recognition scores among our sample that ranged from 20% to 69% correct.

All participants reported themselves to be in good health with no known history of stroke, Parkinson's disease, or other neuropathology that might interfere with their ability to perform the experimental task. Written informed consent was obtained from all participants according to a protocol approved by the Institutional Review Boards of Brandeis University and the New York University School of Medicine.

## Stimuli and Procedures

The stimuli and procedures followed those described in Experiment 1, with two major exceptions. The first is that all narratives were presented without vocoding. This increased the number of narratives that would be heard in each condition. The second exception is that the stimuli were presented via a sound field loudspeaker. The loudspeaker was positioned at zero azimuth, 1 m away from the listener. The sound level was set at 65 dB (C scale), which was reported to be a comfortable listening level by all the participants. Testing was conducted with the participant using his or her everyday implant program settings.

## Results

**Narrative Recall.** Figure 5 shows the mean percentage of propositions recalled from the narratives when presented at the original normal speech rate ( $N$ ), when time-compressed to 60% of the original playing time (TC), and with time-restoration (TR), where silent pauses were inserted into the compressed passages at clause and sentence boundaries. (One CI user stopped after the first experimental block. For this participant, accuracy data was based on three trials in each condition.)

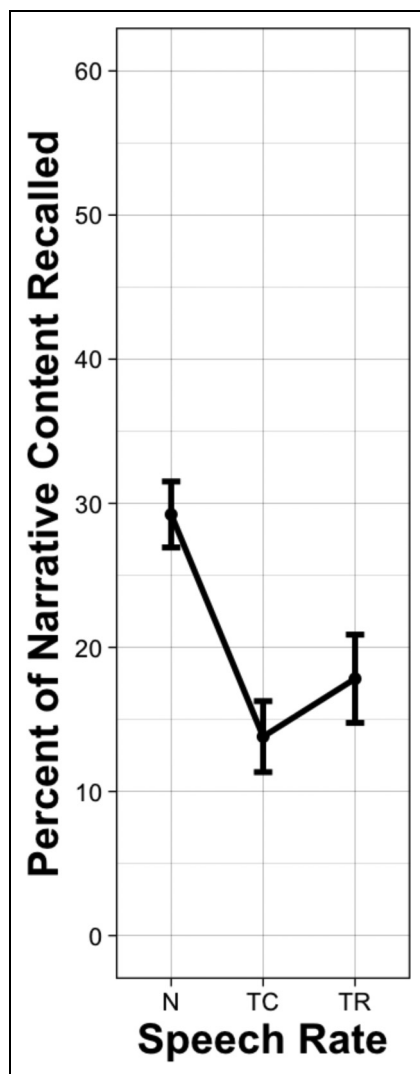
The data in Figure 5 were submitted to a linear mixed effects model, with the same model selection and comparison procedure as described in Experiment 1, with the exception that the model did not include a clarity (vocoding) factor. The outcome is given in Table 4. As expected, time compression significantly reduced recall performance relative to normal speech rate ( $p < .001$ ). A significant increase in recall accuracy was found when time-compressed passages were heard with time-restoration ( $p = .003$ ). However, a significant difference nevertheless remained between the recall

**Table 3.** CI User Demographic Information.

Age	Sex	Years of experience (L/R)	Electrode (L/R)	Processor (L/R)	Etiology	CNC30	AzBION
28	F	17.77/14.01	CI24R/CI24RE	N7/N7	Congenital	31%	64.24%
32	F	21.59/7.83	CI24 M/CI24RE	N6/N6	Congenital	61%	75.08%
33	F	1.61/16.54	CI632/CI24RE	N7/N7	Congenital	20%	32.25%
34	F	18.63/-	HiFocus 1J/-	Marvel/-	Waardenburg	33%	32.06%
38	M	16.66/10.35	HiFocus 1J/HiFocus 1J	Naida Q90/Naida Q90	Congenital	42%	65.41%
61	F	10.24/8.97	Flex24/Flex24	Rondo2/Rondo2	Unknown progressive	45%	58.20%
62	F	31.59/5.24	HiFocus/HiFocus MS	Harmony/Naida Q90	Unknown progressive	75%	91.70%
64	F	2.05/1.12	HiRes Ultra MS/HiRes Ultra MS	Marvel/Marvel	Unknown progressive	69%	87.04%
67	M	3.62/4.95	Flex24/Flex24	Sonnet/Sonnet	Unknown progressive	45%	77.69%
68	F	14.16/8.69	CI24RE/CI24RE	N7/N7	Unknown progressive	52%	73.92%
72	F	8.47/6.13	CI24RE (CA)/CI512	Kanso/Kanso	Unknown progressive	38%	56.53%
72	F	-/1.45	-/CI632	-/Kanso2	Unknown progressive	66%	81.94%
75	F	4.16/9.71	CI532/CI24RE	N7/N7	Unknown progressive	57%	75.72%
76	M	3.95/1.45	CI622/CI632	N7/N7	Unknown progressive	38%	19.24%
78	M	13.97/25.23	HiFocus 1J/HiFocus 1J	Nadia Q90/Nadia Q90	Unknown progressive	33%	46.21%

level for the normal speech rate and the recall level in the time-restoration condition ( $p < .001$ ).

**Predictors of Individual Differences in Recall.** For CI users, AzBio scores predicted overall recall performance for narratives ( $p = .001$ ), but CNC-30 scores did not ( $p = .562$ ). None of the ancillary tests (Shipley vocabulary, R-Span, Flanker), nor participant age had a systematic effect on overall recall accuracy, effects of time-compression, or effects of time-restoration. There was, however, considerable variability among the CI users in their degree of recovery with time-restoration that prompted further exploration.



**Figure 5.** Mean percentage of narrative content (propositions) recalled from narratives by CI users when the narratives were presented at their original normal speech rate (N), when time-compressed to 60% of their original playing time (TC) and in the time-restoration (TR) condition in which the lost processing time was restored by silent periods inserted at clause and sentence boundaries. Error bars are one standard error.

**Subgroup Differences in the Effect of Time-Restoration.** In an exploratory analysis, each individual's degree of recovery with time-restoration was calculated using the equation  $(TR - TC)/(N - TC)$ , where TR is the mean recall accuracy when time-compressed speech was heard with time-restored, TC is the mean recall when speech was heard with time-compression, and N is the mean recall accuracy when speech was heard at a normal rate. Five of the CI users showed what may be considered to be the meaningful recovery of 50% or more, and five CI users showed little or no recovery (10% or less).

The two subgroups are shown in Table 5 along with the three metrics that distinguished the two subgroups followed by the three metrics that did not. The group we defined as having meaningful recovery with time-restoration had higher CNC30 scores,  $t(8) = 5.57$ ,  $p < .001$ , AzBio scores,  $t(8) = 3.19$ ,  $p = .006$ , and R-Span scores,  $t(7) = 2.50$ ,  $p = .020$ , than the group that failed to show meaningful recovery, but were not significantly different in vocabulary score,  $t(8) = 1.73$ ,  $p = .061$ , Flanker score,  $t(8) = 1.17$ ,  $p = .137$ , or age,  $t(8) = 0.93$ ,  $p = .190$ .

**Pupillometry.** The pupillometry data for the CI users are shown in Figure 6 adjusted for presentence baselines and presented as a percentage of individuals' pupillary dynamic range. Figure 6 shows these mean adjusted pupil sizes as participants listened to narratives presented with the original normal speech rate, when time-compressed, and when time-compressed with time-restoration. These time series data are again plotted as a percentage of the passage durations to take into account the reduced duration of the time-compressed passages.

Although with moment-to-moment variability, one sees the same three features in the pupillary response for the CI users as was observed for the normal hearing participants in Experiment 1. First, time compression of the narratives resulted in larger pupil dilations relative to when the passages were heard at a normal speech rate. Second, pupil sizes were reduced when the compressed passages were heard with time-restoration. Third, a facilitation effect was observed in

**Table 4.** Log-Likelihood Model Comparisons for Linear Mixed Effects Analysis of Cochlear Implant Users' Narrative Recall.

Predictor	$\chi^2$ <sup>a</sup>	$df$ <sup>b</sup>	$p$ <sup>c</sup>
Rate	31.22	2	<b>&lt;0.001</b>
AzBio	10.71	1	<b>.001</b>
CNC30	0.34	1	.562
Age	0.17	1	.683

Notes:

<sup>a</sup> $\chi^2$  value for comparisons of each step of the model.

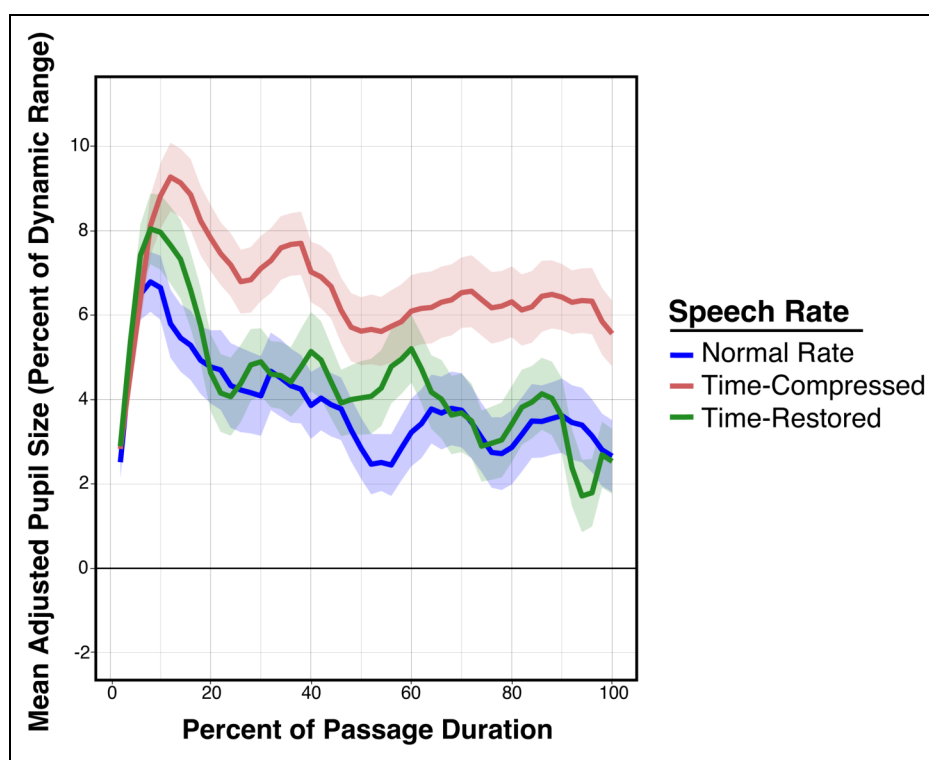
<sup>b</sup>Degrees of freedom for the  $\chi^2$  test.

<sup>c</sup>Value of  $p$  reflects the significance of the change in model fit at each step of the model.

Significant values of  $p$  indicated in bold.

**Table 5.** Subgroup Differences in Effect of Time-Restoration.

	Effect of restoration	CNC30	AzBio	Reading span	Shipley	Flanker	Age
Meaningful Recovery	228%	69	87.04	10	17	52.65	64
	100%	66	81.94	7	17	52.65	71
	61%	57	75.72	10	17	75.5	73
	59%	75	91.7	13	17	60.18	62
	50%	61	75.08	10	17	104.43	31
Little or no Recovery	9%	45	77.69	-	17	48.53	66
	6%	20	32.25	8	12	109.1	32
	6%	38	19.24	3	16	5.83	76
	0%	33	32.06	6	15	17.9	34
	-18%	42	65.41	8	17	44.45	38



**Figure 6.** Time course of mean pupil dilations adjusted to baseline and to each individual's dynamic range while CI users listened to narrative passages. Data are shown for narratives that were time-compressed (red), presented at a normal speech rate (blue), and heard with time-restoration (green). Data are transformed to represent changes in pupil size as a percentage of the passage duration. Error ribbons are one standard error.

the form of a reduced pupillary response as the passages proceeded.

The data shown in Figure 6 were analyzed using GAMMs with the final model containing fixed effects of Rate (normal rate, time-restored, time-compressed) and factor smooths (random effects) of both participants and narratives. Autocorrelation was calculated from an initial model and used as an autoregressive parameter in the final model. Inspection of the final model revealed that specifying this parameter appropriately factored out autocorrelation

between residuals. The results of the final model are summarized in Table 6.

As with the young adults, the CI users showed significantly higher pupil sizes when listening to time-compressed narratives relative to the narratives presented at a normal speech rate ( $p < .001$ ). The pupillary responses for time-compressed narratives heard with time-restoration were again not significantly different from the pupillary responses to passages heard at a normal rate ( $p = .250$ ).

**Time Regions of Significant Differences.** As in Experiment 1, we once again note that while Table 6 displays a convenient summary of significant effects, it is also important to examine the specific points in time when significant differences between conditions occurred. Figure 7 displays the difference curves for CI users estimated by the GAMM between the pupil traces for the passages heard with normal speech rate subtracted from those heard with time compression (Figure 7(A)), the pupil traces for passages heard at a normal speech rate subtracted from the those heard with time-restoration (Figure 7(B)), and the pupil traces for passages heard with time-restoration subtracted from those heard with time-compression (Figure 7(C)).

For CI users, as seen in Figure 7(A), which shows the full-time series, the pupil size was significantly larger when participants heard time-compressed passages compared to passages at a normal speech rate. This difference occurred for 94.04% of the passage, starting at 5.96% after stimulus onset and continuing until the offset of the passage. In Figure 7(B), when comparing the full-time series of pupil dilation, pupil size was significantly larger when participants

heard passages with time-restoration than when participants heard passages at a normal speech rate for 38.60% of the passages' durations. This is notably different than the effect of time-restoration seen in Table 6, where there is no significant difference when the full-time series is not considered. The significant differences occurred in three time windows from 4.97% to 17.84%, 36.67% to 41.60%, and 45.56% to 64.36% of the passages' durations. Finally, it can be seen in Figure 7(C) that pupil size was larger for passages heard with time-compression compared to passages heard with time-restoration for 79.19% of the passages' durations, in two time windows from 11.90% to 51.49% and 60.40% to 100% of the passages' durations.

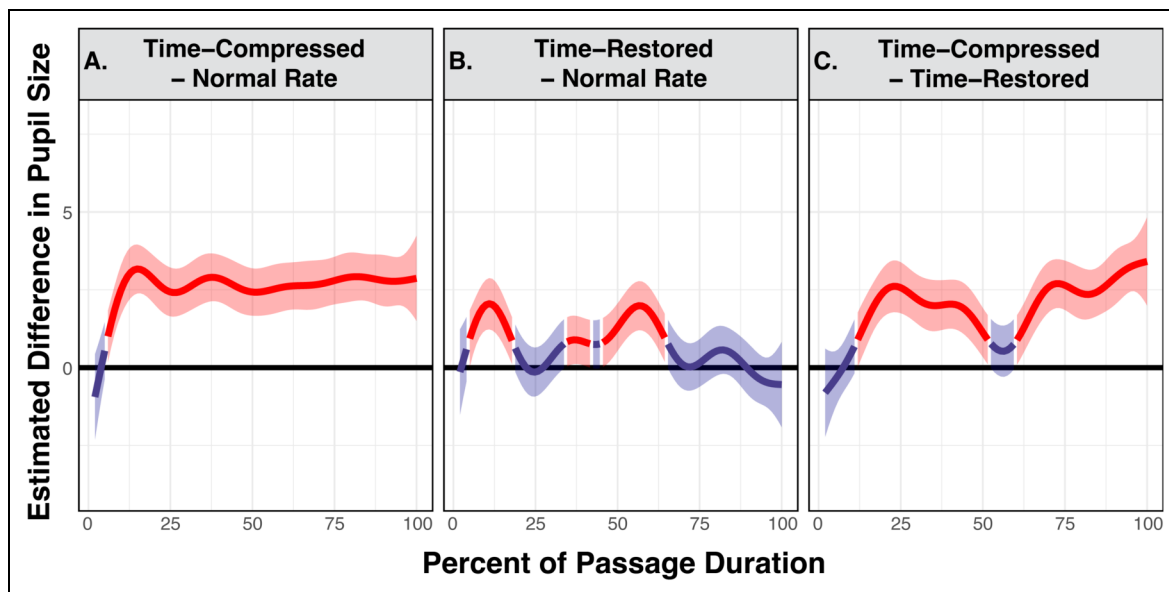
## General Discussion

For both normal-hearing young adults in Experiment 1, and CI users in Experiment 2, the detrimental effects of time-compression on recall were rescued, albeit to variable degrees, by the insertion of silent periods at linguistically salient points in the time-compressed narratives. Because in the time-restoration condition, the speech itself remained time-compressed, the improved recall can be attributed to allowing listeners' processing additional time to "catch up" with the input rather than affecting the clarity of the signal itself.

It has been shown in prior research that restoration of the processing time lost due to time compression is most effective when silent periods are added at points corresponding

**Table 6.** Generalized Additive Mixed Model Summary of Cochlear Implant Users' Pupillary Responses

Parametric coefficients	Estimate	SE	t-Value	p-Value
Intercept (normal rate)	3.79	1.25	3.03	.002
Time-restored	0.67	0.58	1.15	.250
Time-compressed	2.70	0.58	4.64	<.001



**Figure 7.** Generalized additive mixed model predicted difference curves of adjusted pupillary responses when CI users listened to narrative passages. The top panels show contrasts in predicted pupil size between conditions of time-compression versus normal rate (A), time-restored versus normal rate (B), and time-compressed versus time-restored (C). A higher estimated difference is an indication of the greater difference between compared conditions, with regions highlighted in red indicating time points where there is a statistically significant difference. Shaded ribbons are 95% confidence intervals.



to major linguistic constituents (Wingfield et al., 1999). This was done in the present study. The effectiveness of this placement would be expected to the extent that major linguistic constituents (major clause and sentence boundaries) are presumed to be natural processing points for connected discourse (Ferreira & Anes, 1994). By contrast, adding time by uniformly expanding the duration of the speech signal that includes stretching out the duration of individual words, has failed to show a significant improvement in speech intelligibility (Winn & Teece, 2021), while other studies have found a trend toward poorer recall performance relative to speech presented at a normal speech rate (cf., Ji et al., 2013; Schmitt, 1983; Schmitt & McCroskey, 1981; Schmitt & Moore, 1989). That is, how one offers the listener extra processing time is important to its effectiveness.

For normal-hearing young adults, pauses inserted strategically in rapid speech rescues the listener from the detrimental effect of rapid speaking rate, with that benefit diminishing as the signal gets poorer and poorer. The recall decrement resulting from time compression was fully reversed in the time-restoration condition for clear speech, and nearly so when the speech was heard with 10-channel vocoding. These cases reflect the expectations for a resource-limited process. That is, in both cases there was sufficient sensory information available to the listener to allow an effective allocation of processing resources, along with the presence of periodic silent intervals during which no new information was arriving, to yield a level of recall approximating that observed for noncompressed speech.

A different case was observed for the normal-hearing listeners with 6-channel vocoding and for the group effect for the CI users. Both showed roughly a 20% drop in recall accuracy from baseline with time-compression and a significant but modest 5–10% recovery when the lost processing time was strategically restored. In both cases, the shift from a resource-limited process to a data-limited process was approached but not completely reached. Conceptually, this would imply that the reduced sensory information due to the loss of acoustic richness when time-compression was combined with 6-channel vocoding created a situation approaching the data-limited end of the resource-limited—data limited continuum.

It is interesting to note the similar level of relative recovery with time-restoration for CI users and the normal-hearing listeners with 6-channel vocoding in view of the equivalent of four to eight channels available to many CI users (Karoui et al., 2019). It should also be noted, however, that the overall level of recall performance by CI users with their everyday implant program settings was lower than that observed for normal hearing listeners with a 6-channel vocoder simulation. Consistent with this finding is the possibility raised by O'Neill et al. (2019) that vocoder simulations with normal-hearing listeners may underestimate the degree of degradation in the CI signal. As these authors suggest, such degradation need not be uniform. This lack of

uniformity could occur with unevenness in spectral resolution along the length of the cochlea due to unevenness in neural survival, quality of the electrode-neural interface across the electrode array, or both. (O'Neill et al., 2019). In this regard, it can also be argued that valid acoustic models of sound perception by CI users should incorporate individual amounts of frequency mismatch between the analysis filters and the output noise bands or tones (Svirsky et al., 2021). This was not done in the present study.

It is well-recognized that, due to factors such as electrode placement, current spread, neural survival, duration of CI experience, and other factors, variability is a virtual hallmark of outcome performance among adult CI users (e.g., Bierer, 2010; Lenarz et al., 2012; Mahmoud & Ruckenstein, 2014).

This outcome variability may well have overwhelmed the potential effects that participant age, vocabulary, or executive function, as measured by flanker scores, might have had on the effectiveness of time restoration. This is especially notable given the wide age range among the CI users in our sample and the important role adult aging is known to play in a range of speech comprehension domains (see Gordon-Salant et al., 2020 for a review).

One aspect of this variability was illustrated by the appearance of a subset of CI users who showed meaningful recovery with time-restoration, and a subset that did not. As we saw, these two subgroups were distinguished by both perceptual and cognitive factors. Specifically, those CI users with better word recognition ability, as indexed by repetition accuracy for CNC-30 words and AzBio sentences, and better working memory capacity, as indexed by R-Span scores, showed better recovery from the detrimental effects of time compression when given added processing time in the time-restoration condition than those that did not.

As noted, pupillometry has seen wide use as a measure of relative processing effort in a range of speech and problem-solving tasks (e.g., Van der Wel & Steenbergen, 2018; Zekveld et al., 2018). In the present case, the TEPR offered objective physiological evidence of greater effort being allocated to the task when the speech input was very rapid due to time compression, and reduced processing effort when the task was made more manageable with the insertion of silent periods at salient points in the passages that allowed the listener additional processing time. The GAMM visualizations revealed that for CI users, complete recovery of processing effort did not appear until later in the passage than was the case for normal hearing young adults. This implies that the effortful demand associated with rapid speech, combined with the spectrally impoverished signal of the CI, delayed the point at which strategic pauses completely relieved CI users from the effort of listening to time-compressed speech.

The general decline in pupil size across the duration of a passage was seen for both the normal-hearing young adults and the CI users. For the normal hearing young adults in the normal rate and time-restored conditions within the

clear speech and the 10-channel vocoding conditions, this decline resulted in pupil size returning to baseline levels. It is interesting to note that within the 6-channel condition for the normal hearing young adults and for the CI users, the pupil size for the normal rate and time-restored condition does not entirely return to baseline by the end of the passage. It is possible that the perceptual challenge represented in these two cases resulted in participants continuing to show a level of effortful processing for a longer period.

## Conclusion

The compression-restoration paradigm could be seen as modeling communication in real-world discourse where speech is often rapid and presented without interruption, potentially overloading the listener. In this regard, there is increasing interest in the medical literature on the detrimental effect of rapid speech on healthcare providers' communication with patients, and especially for communication with patients unfamiliar with medical terms and procedures (Rainey et al., 2022). One would expect such effects to be exacerbated when dealing with individuals with hearing loss, or as in the present case, those who hear using a cochlear implant.

Extrapolating from the present findings to everyday communication would encourage healthcare providers, and others, to be mindful of the limitations of human processing rates and the further decrements to recall of information when processing the speech requires extra effort.

Numerous studies have shown that factors such as eye contact can be used by a listener to signal a speaker to pause in order for the listener to ask a question or to gain time to process what has been heard (see Degutyte & Astell, 2021, for a review). A more direct method for metering the rate of information, however, is for the speaker to intentionally pause at critical processing points as they deliver information. In the present case, pauses were inserted after clauses and sentences within a narrative. Translated into everyday listening, our present results would encourage conversation partners to pause at strategic processing points, such as extra time between statements or communicative topics, to allow processing and memory consolidation before proceeding to a subsequent point.

As we have seen, some CI users might not show a meaningful benefit from a speaker pausing within rapid speech. Although this list is not exhaustive, the present data suggest that CI users' ability to make use of this extra time is influenced by the postimplantation perceptual outcome, and also by participants' working memory capacity (see Pisoni, 2000, for a discussion of cognitive factors in CI outcomes). These results emphasize the importance of sensory-cognitive interactions in individual tailoring of amelioration tactics for communication beyond the single sentence level.

## Acknowledgments

We thank Matthew Winn for helpful and insightful comments on earlier versions of this article.

## Authors' Notes

Mario A. Svirsky and Arthur Wingfield are cosenior authors.

## Data Availability

Data and processing codes will be made available upon reasonable request.








## 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

This work was supported by the National Institutes of Health under award R01 DC016834 from the National Institute on Deafness and Other Communication Disorders. We also gratefully acknowledge support from the Stephen J. Cloobek Research Fund. The funders were not involved in the study design, data collection, analysis, or interpretation of these data.

## ORCID iDs

Jonathan Neukam  <https://orcid.org/0000-0002-9629-727X>  
 Thomas A. Hansen  <https://orcid.org/0000-0002-7950-8039>  
 Alexander J. Kinney  <https://orcid.org/0000-0003-2446-3133>  
 Nicole Capach  <https://orcid.org/0000-0001-7604-8375>  
 Mario A. Svirsky  <https://orcid.org/0000-0002-9238-0682>  
 Arthur Wingfield  <https://orcid.org/0000-0003-3887-3483>  
 Ryan M. O'Leary  <https://orcid.org/0000-0003-1042-0466>

## References

- Abdel-Latif, K. H. A., & Meister, H. (2022). Speech recognition and listening effort in cochlear implant recipients and normal-hearing listeners. *Frontiers in Neuroscience, 15*. <https://doi.org/10.3389/fnins.2021.725412>
- Amichetti, N. M., Neukam, J., Kinney, A. J., Capach, N., March, S. U., Svirsky, M. A., & Wingfield, A. (2021). Adults with cochlear implants can use prosody to determine the clausal structure of spoken sentences. *Journal of the Acoustical Society of America, 150*(6), 4325. <https://doi.org/10.1121/10.0008899>
- Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. *Comprehension and Teaching: Research Reviews, 77*–117.
- Aronson, D., Markowitz, M., & Shapiro, H. (1971). Perception and immediate recall of normal and compressed auditory sequences. *Perception and Psychophysics, 9*, 338–344. <https://doi.org/10.3758/BF03208691>
- Ayasse, N. D., Lash, A., & Wingfield, A. (2017). Effort not speed characterizes comprehension of spoken sentences by older adults with mild hearing impairment. *Frontiers in Aging Neuroscience, 8*, 1–12. <https://doi.org/10.3389/fpsyg.2019.02947>
- Ayasse, N. D., & Wingfield, A. (2018). A tipping point in listening effort: Effects of linguistic complexity and age-related hearing loss on sentence comprehension. *Trends in Hearing, 22*, 1–14. <https://doi.org/10.1177/2331216518790907>

- Ayasse, N. D., & Wingfield, A. (2020). Anticipatory baseline pupil diameter is sensitive to differences in hearing thresholds. *Frontiers in Psychology, 10*, 2947. <https://doi.org/10.3389/fpsyg.2019.02947>
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv, 1506*, 04967. <https://doi.org/10.48550/arXiv.1506.04967>
- Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 142–162). Cambridge University Press (pp. 142–162).
- Bierer, J. A. (2010). Probing the electrode-neuron interface with focused cochlear implant stimulation. *Trends in Amplification, 14*(2), 8495. <https://doi.org/10.1177/1084713810375249>
- Bitsios, P., Prettyman, R., & Szabad, E. (1996). Changes in autonomic function with age: A study of pupillary kinetics in healthy young and old people. *Age and Ageing, 25*(6), 432–438. <https://doi.org/10.1093/ageing/25.6.432>
- Blumstein, S. E., Katz, B., Goodglass, H., Shrier, R., & Dworetzky, B. (1985). The effects of slowed speech on auditory comprehension in aphasia. *Brain and Language, 24*(2), 246–265. [https://doi.org/10.1016/0093-934X\(85\)90134-8](https://doi.org/10.1016/0093-934X(85)90134-8)
- Boersma, P., & Weenink, D. (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.05, retrieved 5 January 2022 from <http://www.praat.org/>.
- Brown, C., Snodgrass, T., Kemper, S. J., Herman, R., & Covington, M. A. (2008). Automatic measurement of propositional idea density from part-of-speech tagging. *Behavior Research Methods, 40*(2), 540–545. <https://doi.org/10.3758/BRM.40.2.540>
- Chodorow, M. S. (1979). Time-compressed speech and the study of lexical and syntactic processing. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistics studies presented to Merrill Garrett* (pp. 87–111). Erlbaum.
- Coretta, S. (2022). tidyv: Tidy model visualization for generalized additive models [Computer software manual]. R package version 3.3.0. <https://CRAN.R-project.org/package=tidyv>.
- Croghan, N. B. H., Duran, S. I., & Smith, Z. M. (2017). Re-examining the relationship between number of cochlear implant channels and maximal speech intelligibility. *Journal of the Acoustical Society of America, 142*(6), EL537–EL543. <https://doi.org/10.1121/1.5016044>
- Cychoz, M., Mahr, T., Munson, B., Newman, R., & Edwards, J. R. (2023). Preschoolers rely on rich speech representations to process variable speech. *Child Development, 94*(4), e197–e214. <https://doi.org/10.1111/cdev.13922>
- 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](https://doi.org/10.1016/S0022-5371(80)90312-6)
- Degutyte, Z., & Astell, A. (2021). The role of eye gaze in regulating turn taking in conversations: A systematized review of methods and findings. *Frontiers in Psychology: Language Sciences, 12*. <https://doi.org/https://doi.org/10.3389/fpsyg.2021.616471>
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*, 143–149. <https://doi.org/10.3758/BF03203267>
- Faulkner, A., Rosen, S., & Wilkinson, L. (2001). Effects of the number of channels and speech-to-noise ratio on rate of connected discourse tracking through a simulated cochlear implant. *Ear and Hearing, 22*(5), 431–438. <https://doi.org/10.1097/00003446-200110000-00007>
- Ferreira, E., & Henderson, J. M. (1990). The use of verb information in syntactic parsing: A comparison of evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 16*(4), 555–568. <https://doi.org/10.1037/0278-7393.16.4.555>
- Ferreira, F., & Anes, M. D. (1994). Why study spoken language processing? In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 35–56). Academic Press.
- Foulke, E. (1971). The perception of time compressed speech. In D. L. Horton & J. Jenkins (Eds.), *The perception of language* (pp. 79–107). Merrill.
- Friese, 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. *Journal of the Acoustical Society of America, 110*(2), 1150–1163. <https://doi.org/10.1121/1.1381538>
- Frijters, J. (2004). Ikvm, an implementation of java for mono and the net framework [computer software and documentation].
- Fu, Q. J., & Nogaki, G. (2005). Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing. *Journal of the Association for Research in Otolaryngology, 6*(1), 19–27. <https://doi.org/10.1007/s10162-004-5024-3>
- Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Erlbaum.
- Gordon-Salant, S., & Fitzgibbons, P. J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *Journal of Speech and Hearing Research, 36*(6), 1276–1285. <https://doi.org/10.1044/jshr.3606.1276>
- Gordon-Salant, S., Shader, M. J., & Wingfield, A. (2020). Age-related changes in speech understanding: Peripheral versus cognitive influences. In K. S. Helfer, E. L. Bartlett, A. N. Popper, & R. R. Fays (Eds.), *Aging and hearing, springer handbook of auditory research* (Vol. 72, pp. 199–230). Springer Handbook of Auditory Research. [https://doi.org/10.1007/978-3-030-49367-7\\_9](https://doi.org/10.1007/978-3-030-49367-7_9).
- Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology, 33*(4), 457–461. <https://doi.org/10.1111/j.1469-8986.1996.tb01071.x>
- Guillon, M., Dumbleton, K., Theodoratos, P., Gobbe, M., Wooley, C. B., & Moody, K. (2016). The effects of age, refractive status, and luminance on pupil size. *Optometry and Vision Science, 93*(9), 1093–1100. <https://doi.org/10.1097/OPX.0000000000000893>
- Haberlandt, K., & Graesser, A. C. (1989). Buffering new information during reading. *Discourse Processes, 12*(4), 479–494. <https://doi.org/10.1080/01638538909544741>
- Hegarty, M., Shah, P., & Miyake, A. (2000). Constraints on using the dual-task methodology to specify the degree of central executive involvement in cognitive tasks. *Memory & Cognition, 28*(3), 376–385. <https://doi.org/10.3758/BF03198553>
- Heiman, G. W., Leo, R. J., Leighbody, G., & Bowler, K. (1986). Word intelligibility decrements and the comprehension of time-compressed speech. *Perception & Psychophysics, 40*(6), 407–411. <https://doi.org/10.3758/BF03208200>
- Hershman, R., Henik, A., & Cohen, N. (2018). A novel blink detection method based on pupillometry noise. *Behavioral Research*

- Methods*, 50(1), 107–114. <https://doi.org/10.3758/s13428-017-1008-1>
- Hoyle, K. J., Brownell, H., & Wingfield, A. (2009). Components of speech prosody and their use in detection of syntactic structure by older adults. *Experimental Aging Research*, 35(1), 129–151. <https://doi.org/10.1080/03610730802565091>
- Huang, Y., & Snedeker, J. (2020). Evidence from the visual world paradigm raises questions about unaccusativity and growth curve analyses. *Cognition*, 200, 104251. <https://doi.org/10.1016/j.cognition.2020.104251>
- Ji, C., Galvin, J. J., Xu, A., & Fu, Q. J. (2013). Effect of speaking rate on recognition of synthetic and natural speech by normal-hearing and cochlear implant listeners. *Ear & Hearing*, 34(3), 313–323. <https://doi.org/10.1097/AUD.0b013e31826fe79e>
- Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology*, 47(2), 310–339. <https://doi.org/10.1037/h0078820>
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs: Prentice-Hall.
- Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, 154(3756), 1583–1585. <https://doi.org/10.1126/https://doi.org/science.154.3756.1583>
- Karoui, C., James, C., Barone, P., Bakhos, D., Marx, M., & Macherey, O. (2019). Searching for the sound of a cochlear implant: Evaluation of different vocoder parameters by cochlear implant users with single-sided deafness. *Trends in Hearing*, 23, 2331216519866029. <https://doi.org/10.1177/2331216519866029>
- Katz, J. (2002). *Handbook of clinical audiology* (5th Ed). Lippincott Williams & Wilkins.
- Kim, M., Beversdorf, D. Q., & Heilman, K. M. (2000). Arousal response with aging: Pupillographic study. *Journal of the International Neuropsychological Society*, 6(3), 348–350. 00206099709071969. <https://doi.org/10.1017/S135561770000309X>
- Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363. <https://doi.org/10.1037/0033-295X.85.5.363>
- Kong, Y.-Y., Somarowthu, A., & Ding, N. (2015). Effects of spectral degradation on attentional modulation of cortical auditory responses to continuous speech. *JARO: Journal of the Association for Research in Otolaryngology*, 16(6), 783–796. <https://doi.org/10.1007/s10162-015-0540-x>
- Konkle, D. F., Beasley, D. S., & Bess, F. H. (1977). Intelligibility of time-altered speech in relation to chronological aging. *Journal of Speech and Hearing Research*, 20(1), 108–115. <https://doi.org/10.1044/jshr.2001.108>
- Kuchinsky, S. E., Ahlstrom, J. B., Cute, S. L., Humes, L. E., Dubno, J. R., & Eckert, M. A. (2014). Speech-perception training for older adults with hearing loss impacts word recognition and effort. *Psychophysiology*, 51(10), 1046–1057. <https://doi.org/10.1111/psyp.12242>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). Package ‘lmerTest’. *R Package Version*, 2(0), 734. 10.18637/jss.v082.i13
- Lenarz, M., Sönmez, H., Joseph, G., Büchner, A., & Lenarz, T. (2012). Long-term performance of cochlear implants in postlingually deafened adults. *Otolaryngology – Head and Neck Surgery*, 147(1), 112–118. <https://doi.org/10.1177/0194599812438041>
- Liu, H. (2004). MontyLingua. <http://web.media.mit.edu/hugo/montylingua/>
- Luxford, W. (2001). Minimum speech test battery for postlingually deafened adult cochlear implant patients. *Otolaryngology-Head and Neck Surgery*, 124(2), 125–126. <https://doi.org/10.1067/mhn.2001.113035>
- Mahmoud, A. F., & Ruckenstein, M. (2014). Speech perception performance as a function of age at implantation among postlingually deaf adult cochlear implant recipients. *Otology and Neurotology*, 35(10), e286–e291. <https://doi.org/10.1097/MAO.0000000000000581>
- Mandler, J. M. (1978). A code in the node: The use of a story schema in retrieval. *Discourse Processes*, 1(1), 14–35. <https://doi.org/10.1080/01638537809544426>
- McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, 24(2), 222–243. <https://doi.org/10.1037/a0017619>
- McGarrigle, R., Knight, S., Rakusen, L., Geller, J., & Mattys, S. (2021). Older adults show a more sustained pattern of effortful listening than young adults. *Psychology and Aging*, 36(4), 504–519. <https://doi.org/10.1037/pag0000587>
- McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., & Amitay, S. (2014). Listening effort and fatigue: What exactly are we measuring? A British society of audiology cognition in hearing special interest group ‘white paper’. *International Journal of Audiology*, 53(7), 433–440. <https://doi.org/10.3109/https://doi.org/14992027.2014.890296>
- McLaughlin, D. J., Zink, M. E., Gaunt, L., Spehar, B., Van Engen, K. J., Sommers, M. S., & Peelle, J. E. (2022). Pupillometry reveals cognitive demands of lexical competition during spoken word recognition in young and older adults. *Psychonomic Bulletin and Review*, 29(1), 268–280. <https://doi.org/10.3758/s13423-021-01991-0>
- Mirman, D. (2017). *Growth curve analysis and visualization using R*. CRC press.
- Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Kreuger, S. (2005). Divided attention in younger and older adults: Effects of strategy and relatedness on memory performance and secondary task costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3), 520–537. <https://doi.org/10.1037/0278-7393.31.3.520>
- Nooteboom, S. (1997). The prosody of speech: Melody and rhythm. In W. J. Hardcastle & J. Laver (Eds.), *The handbook of phonetic sciences* (pp. 640–673). Blackwell.
- Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, 7(1), 44–64. [https://doi.org/10.1016/0010-0285\(75\)90004-3](https://doi.org/10.1016/0010-0285(75)90004-3)
- O’Neill, E. R., Kreft, H. A., & Oxenham, A. J. (2019). Cognitive factors contribute to speech perception in cochlear-implant users and age-matched normal-hearing listeners under vocoded conditions. *Journal of the Acoustical Society of America*, 146(1), 195–2190. <https://doi.org/10.1121/1.5116009>
- Overman, R. A. (1971). Processing time as a variable in the comprehension of time-compressed speech. In E. Foulke (Ed.), *Proceedings of the 2<sup>nd</sup> Louisville conference on rate and/or frequency controlled speech* (pp. 103–118). University of Louisville.
- Pals, C., Sarampalis, A., & Başkent, D. (2013). Listening effort with cochlear implant simulations. *Journal of Speech, Language, and Hearing Research*, 56(4), 1075–1084. [https://doi.org/10.1044/1092-4388\(2012\)12-0074](https://doi.org/10.1044/1092-4388(2012)12-0074)

- Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Psychophysiology*, *11*(5), 559–566. <https://doi.org/10.1111/j.1469-8986.1974.tb01114.x>
- Peelle, J. E. (2018). Listening effort: How the cognitive consequences of acoustic challenge are related in brain and behavior. *Ear & Hearing*, *39*(2), 204–214. <https://doi.org/10.1097/AUD.0000000000000494>
- Perreau, A., Tyler, R. S., & Witt, S. A. (2010). The effect of reducing the number of electrodes on spatial hearing tasks for bilateral cochlear implant recipients. *Journal of the American Academy of Audiology*, *21*(2), 110–120. <https://doi.org/10.3766/jaaa.21.2.5>
- Peterson, G. E., & Lehiste, I. (1962). Revised CNC lists for auditory tests. *Journal of Speech and Hearing Disorders*, *27*(1), 62–70. <https://doi.org/10.1044/jshd.2701.62>
- Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Homsby, B. W. Y., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G., Phillips, N. A., Richter, M., Rudner, M., Sommers, M. S., Tremblay, K. L., & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear 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. <https://doi.org/10.1111/j.1469-8986.2009.00947.x>
- Pisoni, D. B. (2000). Cognitive factors and cochlear implants: Some thoughts on perception, learning, and memory in speech perception. *Ear and Hearing*, *21*(1), 70–708. <https://doi.org/10.1097/00003446-200002000-00010>
- Rainey, R., Theiss, L., Lopez, E., et al. (2022). Characterizing the impact of verbal communication and health literacy in the patient-surgeon encounter. *The American Journal of Surgery*, *224*(3), 943–948. <https://doi.org/doi.10.1016/j.amjsurg.2022.04.034>
- Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. Jr (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, *4*(3-4), S121–S149. <https://doi.org/10.1080/01690968908406362>
- Reilly, J., Kelly, A., Kim, S. H., Jett, S., & Zuckerman, B. (2019). The human task-evoked pupillary response function is linear: Implications for baseline response scaling in pupillometry. *Behavioral Research Methods*, *51*(2), 865–878. <https://doi.org/10.3758/s13428-018-1134-4>
- Schmitt, J. F. (1983). The effects of time compression and time expansion on passage comprehension by elderly listeners. *Journal of Speech and Hearing Research*, *26*(3), 373–377. <https://doi.org/10.1044/jshr.2603.373>
- Scmitt, J. F., & McCroskey, R. L. (1981). Sentence comprehension in elderly listeners: The factor of rate. *Journal of Gerontology*, *36*(4), 441–445. <https://doi.org/10.1093/geronj/36.4.441>
- Scmitt, J. F., & Moore, J. R. (1989). Natural alternation of speaking rate: The effect on passage comprehension by listeners over 75 years of age. *Journal of Speech and Hearing Research*, *32*(2), 445–450. <https://doi.org/10.1044/jshr.3202.445>
- Seropian, L., Ferschneider, M., Cholvy, F., Michey, C., Bidet-Caulet, A., & Moulin, A. (2022). Comparing methods of analysis in pupillometry: Application to the assessment of listening effort in hearing-impaired patients. *Heliyon*, *8*(6), e09631. <https://doi.org/10.1016/j.heliyon.2022.e09631>
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303–304. <https://doi.org/10.1126/science.270.5234.303>
- Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistics Research*, *25*(2), 193–247. <https://doi.org/10.1007/BF01708572>
- Shipley, W., Gruber, C., Martin, T., & Klein, M. (2009). *Shipley-2 professional manual*. Western Psychological Services.
- Skinner, M. W., Holden, L. K., Fourakis, M. S., Hawks, J. W., Holden, T., Arcaroli, J., & Hyde, J. (2006). Evaluation of equivalency in two recordings of monosyllabic words. *Journal of the American Academy of Audiology*, *17*(5), 350–366. <https://doi.org/10.3766/jaaa.17.5.5>
- Spahr, A. J., Dorman, M. F., Litvak, L. M., Van Wie, S., Gifford, R. H., Loizou, P. C., Loiselle, L. M., Oakes, T., & Cook, S. (2012). Development and validation of the AzBio sentence lists. *Ear and Hearing*, *33*(1), 112–117. <https://doi.org/10.1097/AUD.0b013e31822c2549>
- Sticht, T. G., & Gray, B. B. (1969). The intelligibility of time compressed words as a function of age and hearing loss. *Journal of Speech and Hearing Research*, *12*(2), 443–448. <https://doi.org/10.1044/jshr.1202.443>
- Stine, E. A. L. (1990). On-line processing of written text by younger and older adults. *Psychology and Aging*, *5*(1), 68–78. <https://doi.org/10.1037/0882-7974.5.1.68>
- Stine, E. A. L., Loveless, M. K., & Soederberg, L. M. (1996). Resource allocation in on-line reading by younger and older adults. *Psychology and Aging*, *11*(3), 475–486. <https://doi.org/10.1037/0882-7974.11.3.475>
- Stine, E. A. L., & Hindman, J. (1994). Age differences in reading time allocation for propositionally dense sentences. *Aging, Neuropsychology, and Cognition*, *1*(1), 2–16. <http://dx.doi.org/10.1080/09289919408251446>
- Svirsky, M. A., Capach, N. H., Neukam, J. D., Azadpour, M., Sagi, E., Hight, A. E., Glassman, E. K., Lavender, A., Seward, K. P., Miller, M. K., Ding, N., Tan, C.-T., & Fitzgerald, M. B. (2021). Valid acoustic models of cochlear implants: One size does not fit all. *Otology & Neurotology*, *42*(10), S2–S10. <https://doi.org/10.1097/MAO.0000000000003373>
- Tun, P. A., McCoy, S., & Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and Aging*, *24*(3), 761–766. <https://doi.org/10.1037/a0014802>
- Turner, A., & Greene, F. (1978). The construction and use of a propositional text base. *JSAS: Catalog of Selected Documents in Psychology*, *8*, 1–81.
- Van der Wel, P., & Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin and Review*, *25*(6), 2005–2015. <https://doi.org/10.3758/s13423-018-1432-y>
- 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. <https://doi.org/10.1177/2331216519832483>
- van Rij, J., Wieling, M., Baayen, R. H., & van Rijn, H. (2020). itsadug: Interpreting time series and autocorrelated data using gamms (R package version 2.4).
- Wilson, B. S. (2004). Engineering design of cochlear implants. In F.-G. Zeng, A. N. Popper, & R. R. Fay (Eds.), *Cochlear implants: Auditory prostheses and electric hearing* (Vol. 20, pp. 14–52). Springer.

- Wingfield, A., McCoy, S. L., Peelle, J. E., Tun, P. A., & Cox, L. C. (2006). Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity. *Journal of the American Academy of Audiology*, *17*(7), 487–497. <https://doi.org/10.3766/jaaa.17.7.4>
- Wingfield, A., Peelle, J. E., & Grossman, M. (2003). Speech rate and syntactic complexity as multiplicative factors in speech comprehension by young and older adults. *Journal of Aging, Neuropsychology and Cognition*, *10*(4), 310–322. <https://doi.org/10.1076/anec.10.4.310.28974>
- Wingfield, A., Tun, P. A., Koh, C. K., & Rosen, M. J. (1999). Regaining lost time: Adult aging and the effect of time restoration on recall of time-compressed speech. *Psychology and Aging*, *14*(3), 380–389. <https://doi.org/10.1037/0882-7974.14.3.380>
- 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), 1–13. <https://doi.org/10.1097/AUD.0000000000000145>
- Winn, M. B., & Teece, J. H. (2021). Slower speaking rate reduces listening effort among listeners with cochlear implants. *Ear and Hearing*, *42*(3), 584–595. <https://doi.org/10.1097/AUD.0000000000000958>
- 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*, 1–32. [10.1177/2331216518800869](https://doi.org/10.1177/2331216518800869)
- Wood, S. (2017). *Generalized additive models: An introduction with R* (2nd ed.). CRC Press.
- Zekveld, A. A., Koelewijn, T., & Kramer, S. E. (2018). The pupil dilation response to auditory stimuli: Current state of knowledge. *Trend in Hearing*, *22*, 2331216518777174. <https://doi.org/10.1177/2331216518777174>
- 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. <https://doi.org/10.1097/AUD.0b013e31820512bb>
- Zhao, S., Bury, G., Milne, A., & Chait, M. (2019). Pupillometry as an objective measure of sustained attention in young and older

listeners. *Trends in Hearing*, *23*, 2331216519887815. <https://doi.org/10.1177/2331216519887815>

## Appendix A

### Examples of Levels of Recall Accuracy

#### Original Narrative:

Joseph was a 19-year-old student from the University of Southern California who had never been to Europe. He was very excited because he had family in Scotland. He was offered a full-time summer job looking after the displays of jewels and crowns at Edinburgh Castle. He took history books out of the library but found he learned more from conversations with his coworkers than he did from reading the books.

#### Participant Responses

##### 64.1% of propositions recalled:

Joseph was a 19-year-old student at the University of Southern California, and he had never been to Europe but he was excited to go because he had family in Scotland and he got a summer job working with jewels and crowns in the castle in Edinburgh and he got history books and he felt like he learned more in his conversations with coworkers than the books.

##### 41.0% of propositions recalled:

Peter was a 19-year-old student from the University of Southern California who had never been to Europe. He was excited to go because he had family in Scotland. He went to Scotland, and he talked with his coworkers.

##### 15.4% of propositions recalled:

A college student is going to Europe for the first time. His family's in Scotland. He was offered a summer job involving tombs and a castle. And then I did not catch the rest of it.