



Published in final edited form as:

*Ear Hear.* 2020 ; 41(5): 1144–1157. doi:10.1097/AUD.0000000000000856.

## Tracking Cognitive Spare Capacity During Speech Perception with EEG/ERP: Effects of Cognitive Load and Sentence Predictability

Cynthia R. Hunter<sup>1,2</sup>

<sup>1</sup>Speech Perception, Cognition, and Hearing Laboratory, Department of Speech-Language-Hearing: Sciences and Disorders, University of Kansas

<sup>2</sup>Indiana Clinical and Translational Sciences Institute – TL1 Program

### Abstract

**Objectives.**—Listening to speech in adverse listening conditions is effortful. Objective assessment of cognitive spare capacity during listening can serve as an index of the effort needed to understand speech. Cognitive spare capacity is influenced both by signal-driven demands posed by listening conditions and top-down demands intrinsic to spoken language processing, such as memory use and semantic processing. Previous research indicates that electrophysiological responses, particularly alpha oscillatory power, may index listening effort. However, it is not known how these indices respond to memory and semantic processing demands during spoken language processing in adverse listening conditions. The aim of the current study was twofold: first, to assess the impact of memory demands on electrophysiological responses during recognition of degraded, spoken sentences, and second, to examine whether predictable sentence contexts increase or decrease cognitive spare capacity during listening.

**Design.**—Cognitive demand was varied in a memory load task in which young adult participants ( $n = 20$ ) viewed either low-load (one digit) or high-load (seven digits) sequences of digits, then listened to noise-vocoded spoken sentences that were either predictable or unpredictable, and then reported the final word of the sentence and the digits. Alpha oscillations in the frequency domain and event-related potentials (ERP) in the time domain of the electrophysiological data were analyzed, as was behavioral accuracy for both words and digits.

**Results.**—Measured during sentence processing, event-related desynchronization (ERD) of alpha power was greater (more negative) under high load than low load and was also greater for unpredictable than predictable sentences. A complementary pattern was observed for the P300/late positive complex (LPC) to sentence-final words, such that P300/LPC amplitude was reduced under high load compared to low load and for unpredictable compared to predictable sentences. Both words and digits were identified more quickly and accurately on trials in which spoken sentences were predictable.

**Conclusions.**—Results indicate that during a sentence recognition task, both cognitive load and sentence predictability modulate electrophysiological indices of cognitive spare capacity, namely

alpha oscillatory power and P300/LPC amplitude. Both electrophysiological and behavioral results indicate that a predictive sentence context reduces cognitive demands during listening. Findings contribute to a growing literature on objective measures of cognitive demand during listening, and indicate predictable sentence context as a top-down factor that can support ease of listening.

## Introduction

For listeners with normal hearing in quiet environments, speech perception seems effortless. However, speech perception becomes effortful when sensory representations of speech are degraded either by environmental factors, such as a noisy environment, or factors internal to the listener, such as hearing loss or cochlear implants (Mattys et al., 2012). Listening effort (LE) draws from a capacity-limited pool of general cognitive resources in order to support speech processing in adverse listening conditions (J. Rönnerberg et al., 2013; Pichora-Fuller et al., 2016). Speech perception is one of many everyday activities that draw on general cognitive capacity; other examples include participating in a spoken or typed conversation, remembering the name of a new acquaintance, or navigating to a destination. The amount of one's cognitive capacity that is available for the performance of a given task is not constant but varies due to factors such as fatigue, motivation, and the extent to which capacity is allocated to other concurrent tasks (Pichora-Fuller et al., 2016). The concept of a capacity-limited, general cognitive resource available for allocation to a wide variety of mental tasks was introduced by Kahneman (1973), and overlaps with the concepts of working memory (Baddeley & Hitch, 1974; Engle, 2002) and fluid cognitive ability (Conway et al., 2003). Behaviorally, a trade-off between performance on concurrently performed tasks has long been used as a measure of cognitive spare capacity, that is, the capacity remaining after an individual's total capacity has been temporarily depleted by task demands. For example, in the cognitive load paradigm, having to remember sequences of digits while performing another task is known to reduce performance on the other task, with greater decrements in performance with more digits to be remembered (Baddeley & Hitch, 1974).

For the purposes of the current study, cognitive spare capacity refers to the amount of general cognitive resources that remain available during listening to spoken language for allocation to concurrent tasks, such as processing the informational content of spoken language, storing or retrieving memories, or multitasking (Pichora-Fuller et al., 2016; Rönnerberg et al., 2013; Rudner, 2016). The impact of LE on cognitive spare capacity can be measured behaviorally as a decrement in performance on cognitive tasks done concurrently with speech processing that occurs as speech becomes degraded (for a review, see Gagne, Besser, & Lemke, 2017). Physiological measures of arousal and cognitive demand also reflect increased effort in adverse listening conditions (Obleser et al., 2012; Koelewijn et al., 2014; Winn, Edwards, & Litovsky, 2015). Interestingly, the tradeoff between LE and performance on concurrent cognitive tasks appears to some extent bidirectional, such that the cognitive load created by a concurrent attentional or memory task can impair speech processing (Francis & Nusbaum, 2009; Mattys, Brooks, & Cooke, 2009; Mattys & Wiget, 2011; Mattys, Barden, & Samuel, 2014; Hunter & Pisoni, 2018). Individuals who have enough overall cognitive capacity appear able to improve speech perception in adverse listening conditions by allocating those resources to listening (for reviews, see Akeroyd, 2008; Dryden et al., 2017; however, also see Füllgrabe & Rosen, 2016 for evidence that this

may not be the case among young adults with normal hearing). In the longer term, cognitive demands of listening may impact fatigue and quality of life among people with hearing loss (Hornsby, 2013; Bess & Hornsby, 2014; Hornsby, Naylor, & Bess, 2016).

Multiple electrophysiological measures are known to be generally sensitive to attentional allocation and cognitive demand, including event-related potentials (ERPs) in the time domain as well as neural oscillations in the frequency domain (Luck, Woodman, & Vogel, 2000; Klimesch, 2012). Notably among recent studies, oscillations in the alpha frequency band (8–13 Hz) have been observed to track with the effort involved in speech perception. Alpha oscillations are a prominent feature of the frequency domain of EEG that are known to reflect cognitive resource availability (Pfurtscheller & Da Silva, 1999; Pfurtscheller, Stancak Jr, & Neuper, 1996; Ray & Cole, 1985), and alpha has been used in other domains for purposes ranging from controlling brain-computer interfaces (Wolpaw & McFarland, 1994; Yuan & He, 2014) to assessing mental fatigue in machine operators (Borghini et al., 2014). Recent studies have shown that power fluctuations in the alpha band track variations in speech signal quality, consistent with the idea that alpha tracks LE during speech perception (Obleser & Weisz, 2011; Obleser et al., 2012; Bernarding et al., 2014; Wöstmann et al., 2015; McMahon et al., 2016; Miles et al., 2017). Further, alpha power has been observed to track the combined impact of adverse listening conditions and cognitive (memory) load (Obleser & Weisz, 2011; Obleser et al., 2012), indicating that alpha can serve as a general index of cognitive spare capacity that captures the impact of both acoustic and cognitive factors.

There is growing interest in developing objective measures of LE and cognitive spare capacity for research and clinical settings (N. Rönnerberg et al., 2014; Rudner, 2016; Smith, Pichora-Fuller, & Alexander, 2016). Such measures would have a range of clinical uses, for example indexing the cognitive benefit of a hearing aid, or improving prediction of how well an individual with hearing loss will function in everyday spoken communication (McGarrigle et al., 2014). Subjective measures of LE such as rating scales are simple to administer and indicate an individual's self-assessment of listening difficulty, but may tap different aspects of effort than objective measures (Feuerstein, 1992; Gosselin & Gagne, 2011; Pals, Sarampalis, & Ba kent, 2013). Objective dual-task or memory behavioral tests administered during a speech perception task can index the cognitive demand of listening by tapping into downstream cognitive performance decrements resultant from reduced availability of cognitive resources (Rudner, 2016; Smith et al., 2016). Even more direct indices of cognitive demand may be obtained from physiological measures, such as pupil dilation (for a review, see Zekveld, Koelewijn, & Kramer, 2018) or electrophysiology, which track responses of the central nervous system to changing cognitive demands with high temporal precision.

Currently, clinical prediction of how well an individual with hearing loss will be able to use spoken language in their everyday life often relies on behavioral accuracy in a sentence recognition task. Due to relatively high face validity, sentence recognition tasks are often preferred to single word or nonsense word tests for assessing everyday functioning with spoken language. Sentence recognition engages cognitive and language understanding processes that are used in everyday language spoken communication, including syntactic,

semantic, and coarticulatory cues that can scaffold accurate speech perception (Theunissen, Swanepoel, & Hanekom, 2009). For example, it is well known that when words are embedded in a predictive sentence context, adults with substantial hearing loss often identify spoken words as accurately as those without hearing loss, a finding that has motivated the development of clinical assessments that include predictable sentences (Kalikow, Stevens, & Elliott, 1977; Bilger et al., 1984). Importantly however, such behavioral tests cannot accurately measure the amount of effort needed to reach a given level of performance, nor do they provide an index of the cognitive spare capacity available during speech recognition.

The first aim of the current study was to examine electrophysiological measures of cognitive demand, taken as indices of reduction in cognitive spare capacity, during a sentence recognition task. The broader goal was to begin to develop an electrophysiological measure of cognitive spare capacity during a task that could provide an index of the cognitive demand of listening with high face validity as a model of everyday listening. Task design is of particular importance in measuring EEG oscillations in the alpha band because the directionality of the alpha response can switch depending on whether the brain areas most affected by the task are relevant to task performance and hence become activated by the task, or are irrelevant to task performance and hence are inhibited during task performance (Klimesch, 2012). Generally, alpha oscillations reflect widespread, phase-synchronized neural firing in the alpha band across broad regions of cortex (Klimesch, 1999; Pfurtscheller et al., 1996). This synchronized firing in the alpha frequency range tends to be strongest when external task demands are minimal. As particular brain regions are recruited to process external events, global alpha power drops due to desynchronization of the recruited brain regions in order to form local, functional networks for task-relevant processing (Klimesch, 1999; Palva & Palva, 2007; Pfurtscheller & Da Silva, 1999). The amount of desynchronization typically tracks with the difficulty of a task, being greater for more demanding tasks (Klimesch, 1999). For these reasons, event-related desynchronization (ERD) in the alpha band, reflecting processing in task-relevant brain areas, has long been regarded as an index of cortical activation or engagement.

However, recent work in the visual modality has shown that synchronization in the alpha band, as opposed to desynchronization, plays a role in the functional inhibition of task-irrelevant processing (Klimesch, Sauseng, & Hanslmayr, 2007; Jensen & Mazaheri, 2010; Händel, Haarmeier, & Jensen, 2011). This inhibitory function of alpha may be observed when the inhibition of task-irrelevant brain areas is central to task performance, such as when participants must withhold or control the execution of a response, selectively attend to one signal while ignoring other signals (Kerlin, Shahin, & Miller, 2010; Wilsch et al., 2014), or anticipate a stimulus in a cued location (Banerjee et al., 2011). In fact, the classic observation of alpha ERD as an index of cortical activation may reflect a release from inhibition of task-relevant areas. Obleser and colleagues have provided substantial evidence for the inhibitory function of alpha in the auditory modality in a series of studies on LE. Using experimental designs in which alpha is measured during a post-stimulus delay period that follows presentation of a speech signal and precedes a response cue, these researchers showed that alpha power increases relative to a pre-trial baseline as speech signal quality declines (Obleser, Wöstmann, Hellbernd, et al., 2012; Petersen, Wöstmann, Obleser, et al., 2015; Strauß, Wöstmann, & Obleser, 2014; Wilsch, Henry, Herrmann, et al., 2014), or as

cognitive load increases (Obleser et al., 2012). This pattern of event-related synchronization (ERS) with greater cognitive demand is consistent with the functional inhibition account, wherein high alpha power inhibits processing in task-irrelevant areas when task demands are high. In such a design, alpha power has been observed to track the combined impact of adverse listening conditions and cognitive demand (Obleser et al., 2012).

Fewer studies of the cognitive demand of listening have focused on alpha oscillations measured in designs for which activation in task-relevant brain areas is central to task performance. However, alpha oscillations from task-relevant areas have potential for use as objective indices of cognitive spare capacity during listening. For example, during a sentence recognition task, alpha ERD in task-relevant brain areas may determine the overall pattern observed in the scalp-recorded EEG response because listening to spoken language requires active processing of the incoming speech signal. Given a task such as sentence recognition that would be expected to engage active processing, it is reasonable to expect that, compared to task-irrelevant sources, task-relevant sources of alpha power modulation may be more numerous, moderated more strongly, or both, and thus have a greater influence on the summed activity measured at scalp electrodes. Multiple studies have observed alpha suppression during speech perception, and have interpreted the alpha suppression as an indicator of active, attentive processing of spoken language (Becker, Pefkou, Michel, et al., 2013; Bowers, Saltuklaroglu, Harkrider, et al., 2014; Edwards et al., 2009; Jenson, Harkrider, Thornton, et al., 2015; Krause, Pörn, Lang, et al., 1997; Strauß, Kotz, Scharinger, et al., 2014). Moreover, in a sentence recognition task, response inhibition and/or selective attention are not necessarily central to performance (Lavie, 2005) (although note that selective attention, posited to require alpha ERS to inhibit processing of task-irrelevant information, may become engaged, particularly if sentences are masked by background noise (Strauß et al., 2014)).

In the current study, electrophysiological indices of cognitive demand including both alpha oscillatory activity and event-related potentials were examined in a design in which participants listened attentively to the speech signal in a sentence recognition task. In order to manipulate cognitive demand, a concurrent memory task was deployed during attentive listening to spoken sentences. Specifically, a cognitive load design was used in which either several digits (high load) or a single digit (low load) were presented visually at the beginning of each trial. Listeners were asked to remember the digits as they listened to a spoken sentence in order to identify the final word of each sentence. Given that attending to speech was central to task performance, it was hypothesized that alpha power during sentence processing would track the engagement of task-relevant brain areas used in speech processing rather than the functional inhibition of task-irrelevant processing. Therefore, it was expected that alpha would desynchronize during speech processing relative to a pre-trial baseline. Further, it was predicted that the event-related desynchronization (ERD) of alpha power in task-relevant areas would be greater under high cognitive load than low cognitive load, tracking a decrease in cognitive spare capacity under high load. All spoken sentences were degraded in order to make the listening task effortful and thereby model a listening situation that presents both speech processing and cognitive demands. Speech degradation was accomplished by spectrally vocoding the sentence stimuli with noise-band vocoding. Vocoding, rather than background noise, was chosen in order to maximize the opportunity

to observe alpha dynamics of task-relevant brain areas, given that a noisy background could elicit selective attention to the speech signal and trigger alpha-band functional inhibition of the background noise (Strauß et al., 2014). Although few studies to date have examined alpha oscillatory indices of LE during speech recognition, existing studies using noise degradation have observed alpha ERS as a function of intelligibility (Dimitrijevic, Smith, Kadis, et al., 2017; McMahan et al., 2016). In contrast, studies using vocoding have observed decreased alpha power with more severe spectral degradation, consistent with alpha ERD (McMahan et al., 2016; Miles et al., 2017).

The second aim of the current study was to examine how the use of sentence predictability during speech processing affects electrophysiological measures of cognitive spare capacity. Prediction is ubiquitous in perceptual and linguistic processing, and is especially important for processing degraded signals (Morton, 1969; Bar, 2007; Huettig, 2015). As discussed above, a well established finding is that if spoken words are embedded in predictable sentences, adults with substantial hearing loss can match the performance of adults without hearing loss, whereas lower levels of performance are observed for isolated words (Kalikow et al., 1977; Bilger et al., 1984). Yet, it is not known whether the benefit to speech recognition accuracy from a predictive context is effortless and results in increased ease of listening, or conversely, requires cognitive resources and thereby reduces cognitive spare capacity. Both effortful and effortless routes to prediction are possible, and may be accomplished via distinct neural systems (Huettig, 2015; Kahnehman, 2011). Prediction that is fast, automatic, and effortless may take place via simple associative, spreading activation among the neural representations of related words (Bar, 2007; Huettig, 2015). Prediction that is effortful may require attentional allocation and draw on central cognitive capacity subserved largely by the frontal cortices of the brain (Bonhage, Mueller, Friederici, et al., 2015; Dikker & Pykkänen, 2013; Friederici, Fiebach, Schlesewsky, et al., 2006).

Given that working memory is used to maintain earlier parts of a sentence for processing and integrate those with later words to form a higher-order understanding, it has been suggested that listening effort may increase when degraded speech is held in memory awaiting disambiguating context (Pichora-Fuller, Schneider, & Daneman, 1995; Zekveld et al., 2012). Further, several studies have linked individual differences in the recognition benefit from a sentence context or semantic prime to overall working memory capacity (for a review, see Besser et al., 2013). These studies have found that individuals with greater working memory capacity get a greater word recognition benefit from sentence predictability, which suggests that working memory resources may be used to benefit from sentence predictability (Zekveld et al., 2011; Zekveld, Rudner, Johnsrude, et al., 2013). However, more recent evidence that LE is reduced when sentences are predictable suggests the opposite conclusion. For example, in a behavioral study with a task similar to the current study, Hunter and Pisoni (2018) observed downstream benefits of a predictable context on memory for visually-presented digits such that more digits were remembered on trials in which sentences were predictable. In another study using a subjective measure, ratings by hearing-aid users indicated decreased listening effort for sentences that matched versus mismatched the topic of a previous sentence (Holmes et al., 2018). Few physiological studies to date have examined the cognitive demand of processing a semantic context; however, in a recent pupillometry study, Winn (2016) observed reduced cognitive demand

during and after listening to predictable versus unpredictable sentences. In sum, it is not yet clear whether the use of context to support speech recognition requires active processing that reduces cognitive spare capacity, or alternatively, whether context makes speech recognition both more accurate and less effortful. This question is addressed in the current study by comparing electrophysiological and behavioral measures of cognitive demand in predictable and unpredictable sentences.

In addition to oscillatory electrophysiological activity, event-related potential (ERP) indices of cognitive load and sentence predictability time-locked to sentence-final spoken words were examined in the current study. Analysis focused on the N400 semantic context effect and the P300 or late positive potential (LPC), both of which reflect cognitive-level processing (Kok, 2001; Kutas & Federmeier, 2011; Polich, 2007; Polich & Kok, 1995). The N400 is a negative deflection appearing at approximately 400 ms after the onset of a visual, spoken, or pictured word and is an established index of lexical-semantic processing (Kutas & Federmeier, 2000, 2011; Laszlo & Federmeier, 2009). The N400 semantic context effect refers to larger amplitude of the N400 for sentence-final words that appear in unpredictable compared to predictable sentences (Kutas & Hillyard, 1980, 1984). Therefore, an N400 semantic context effect was expected to differentiate the predictable and unpredictable sentences in the current study. The high-amplitude N400 to words in unpredictable sentences reflects an increased difficulty of word recognition (Lau, Phillips, & Poeppel, 2008) and/or contextual integration (Hagoort, Hald, Bastiaansen, et al., 2004) when the final word of a sentence is not predictable. It was hypothesized that the N400 might be sensitive to cognitive load because it has been shown in prior work to be sensitive to attentional allocation (Kutas & Federmeier, 2011). An effect of load on the N400 would suggest that context use is at least to some extent effortful, whereas if the N400 reflects effortless prediction, then it should not be affected by cognitive load.

In addition, a P300/LPC was expected to be time-locked to sentence-final words and to show effects of cognitive load and sentence predictability. The P300/LPC is associated with attentional and memory processes and is often observed in tasks that require evaluation of a stimulus (for reviews, see Kok, 2001; Polich, 2007; Polich & Kok, 1995; Verleger, 1997). Both the amplitude and latency of the P300/LPC vary with the attentional resources elicited by a perceptual task, such that stimulus evaluations that elicit greater attention induce later latency and larger amplitude (Donchin & Coles, 1988; Polich & Kok, 1995). For example, adding a memory load increases P300 latency and reduces amplitude (Kok, 1997, 2001; Verleger, 1997). In addition, the P300/LPC tends to covary with alpha spectral power (Intriligator & Polich, 1994; Polich, 2007; Polich & Kok, 1995; Yordanova, Kolev, & Polich, 2001). Therefore, similar to alpha oscillations, it was expected that P300/LPC amplitude would index effects of both cognitive load and sentence predictability on cognitive spare capacity, such that amplitude would be larger and latency would be later under high than low cognitive load and for unpredictable than predictable sentences.

Finally, with respect to behavioral responses, based on a prior behavioral study with a similar design to the current study, it was expected that pre-load digits would be recalled more accurately following predictable than unpredictable sentences, consistent with downstream benefits of a release of cognitive resources in high-predictability sentences.

Based on the same study, it was also expected that behavioral accuracy for the recognition of sentence-final words would decrease under high cognitive load, reflecting a syphoning off by the digit processing task of attentional resources needed for word recognition (Hunter & Pisoni, 2018).

In summary, the current study focused on higher-level cognitive and linguistic demands on cognitive spare capacity in a design in which electrophysiological responses were measured during a sentence recognition task. Predictable and unpredictable spoken sentences were presented in a memory load design, and spectral degradation of the sentences was employed to make listening effortful. This task was used to examine how the factors of memory load and sentence predictability would modulate electrophysiological and behavioral indices of cognitive demand, or, inversely, of reductions in cognitive spare capacity. The main aims of the current study were twofold: first, to identify and track the directionality at scalp electrodes of electrophysiological indices of cognitive demand, operationalized as cognitive (memory) load, in a sentence recognition task, and second, to determine whether and how sentence predictability would affect those measures of cognitive demand.

## Materials and Methods

### Participants

Twenty-two young adults recruited from the Indiana University campus participated in this study (12 females, age range 19–26). Data from two participants (1 female) were excluded from further analysis due to an unacceptably low number of EEG trials (< 20 per condition) remaining after data preprocessing. All participants were native English speakers who reported no history of hearing or speech disorders. Participants all gave informed consent and were paid \$10 for each hour of participation, in accordance with procedures approved by the Institutional Review Board at Indiana University at Bloomington. All participants provided their written informed consent for the study

### Stimuli

The speech stimuli were a subset of the sentences from the revised version of the Speech Perception in Noise (SPIN-R) test (Bilger et al., 1984). The target stimuli for this test are the final words of sentences, which are either predictable from the preceding context (e.g., “Stir your coffee with a spoon”) or are not easily predicted from the preceding context (e.g., “John discussed the spoon.”). For each word there is a predictable and an unpredictable sentence version. A total of 148 sentence-final words were selected. A male talker with a Midwestern accent recorded predictable and unpredictable sentences for each sentence-final word. Stimuli were recorded in a sound-attenuating booth using a free-field microphone. Each sentence was spliced into a separate .wav file and normalized to a root-mean-square amplitude of 68 dB SPL. Mean sentence duration was 1.70 sec (range = 1.14 – 2.58,  $SD = 0.223$ ).

Spectral degradation was accomplished by noiseband vocoding using Tiger CIS (<http://www.tigerspeech.com>). Noise vocoding involved an analysis phase, which divides the signal into frequency bands and derives the amplitude envelope from each band, and a synthesis



phase, which replaces the frequency content of each band with noise that is modulated with the appropriate amplitude envelope. Stimuli were bandpass filtered into eight spectral channels between 200 and 7000 Hz using Greenwood's filter function (24 dB/octave slope). The temporal envelope of each channel was then derived using a low pass filter with an upper cutoff at 160 Hz with a 24 dB/octave slope. In the synthesis phase, the spectral information in each channel was replaced with band-pass noise that was modulated by the corresponding temporal envelope.

The visual stimuli for the digit pre-load task were strings of either one (low-load) or seven (high-load) digits in 36-point font. Digit strings were randomly selected on each trial from a set of digit strings. The set contained all possible combinations of the digits 1 through 9 with a set size of one (low-load) or seven (high-load), with no repetitions and no forward consecutive sequences (e.g., "1 2" did not occur in any of the digit sequences, although "1 3" and "2 1" did occur). On low load trials, the single digit was flanked by zeros (three on each side) in order to approximate the visual processing of high load trials.

## Procedure

Stimulus presentation and behavioral data collection was accomplished with Eprime 2.0. Audio signals were presented binaurally through Etymotic ER-3A insert earphones. Participants performed the task while continuous EEG was recorded from the scalp. Prior to the experimental trials, participants received a pre-practice familiarization with vocoded sentences as well as a block of practice trials. In the pre-practice block, ten SPIN-R sentences that were degraded to the same number of vocoded channels as the stimuli in the experimental block were presented. Following each spoken sentence, its written version was displayed on the computer screen for one second and was immediately followed by an on-screen response box in which participants were asked to type the final word of each sentence. The practice block had the same structure as the experimental block and used a set of ten additional SPIN sentences. None of the items for the experimental trials were used in the practice or pre-practice. On-screen instructions preceded each block to orient participants to the materials and requirements of the upcoming task.

Each trial began with visual presentation of the digit pre-load stimuli on the computer screen. The digits were displayed for one second and followed by a blank screen for the duration of an inter-stimulus interval (ISI). Following the ISI, the spoken sentence was presented. A second ISI followed the offset of the spoken sentence, after which a response box appeared on the computer screen to prompt participants to type the final word of the sentence. Immediately after the participant entered their response, a second response box appeared, in which participants typed the digits. The inter-trial interval (ITI) began immediately after the keypress response to the second response box and was jittered randomly from values of 1.75, 2.00, and 2.25 seconds in order to prevent alpha phase-locking to stimulus presentation rate (Woodman, 2010). Similarly, each ISI was jittered randomly from values of 1.00, 1.25, and 1.5 seconds.

A total of 148 trials were presented to each participant, consisting of 37 trials for each combination of sentence type (predictable or unpredictable) and level of digit load (high- or low-load). A set of four counterbalanced lists were used such that across participants, each

word was presented in a predictable and unpredictable sentence, and within each level of predictability, each word was presented with both a high and low cognitive load. Order of presentation of items within a list was randomized. The experiment lasted on average 1.5 hours, including EEG cap fitting.

### EEG Recording and Preprocessing

Electroencephalogram (EEG) was recorded with a 64-channel Geodesic Sensor Net (Electrical Geodesics Inc.) using a high-impedance EGI NetAmps 400 amplifier and EGI Netstation software. Data was recorded with a vertex reference at a sampling rate of 1000 Hz and a band-pass filter of 0.1 to 200 Hz. Electrode impedances were kept below 50 kOhm as per the manufacturer's recommended guidelines. Impedances were tested at the beginning of the experiment session and then approximately every 15 minutes, allowing any high-impedance electrode contacts to be corrected if necessary.

Post-acquisition, all cortical recordings were analyzed using the EEGLAB toolbox (Delorme & Makeig, 2004), an analysis toolbox for Matlab, including in-house routines written to run in EEGLAB. The data was digitally high-pass filtered 0.1 Hz and low-pass filtered at 100 Hz. The continuous data was initially segmented into epochs beginning 1 sec before the onset of the digit stimuli and extending for 7 sec. This period included the onset and offset of the spoken sentences. These epochs were visually inspected to identify bad channels, which were removed – the most channels removed for any participant was four, and the median removed was zero. Epochs with gross electro-ocular and/or electromyographic artifacts ( $>500 \mu\text{v}$ ) were removed using visual inspection (a mean of 5.24 percent of trials were removed ( $SD = 4.00$ , range: 0.00 – 12.84)). Independent component analysis (ICA) was then used to remove remaining eye and muscle movement artifacts (Bell & Sejnowski, 1995; Delorme & Makeig, 2004). After pre-processing, data were re-referenced to an average mastoids reference. Note that trials with incorrect responses for either the word recognition task and the digit recall task, or both, were included in analysis. Given large effects of the predictability factor on word recognition and of the load factor on digit recall, excluding incorrect trials for either behavioral response would have resulted in a highly unbalanced number of trials across conditions, with concomitant effects on signal-to-noise ratio in the EEG data (Cohen, 2014), and would also have created fewer trials per condition than the cutoff value for participant exclusion ( $< 20$  trials).

**Event-related potentials (ERPs).**—The data were re-epoched into 1.6 s epochs (0.1 s before and 1.5 s after the onset of the sentence-final word) and baseline corrected for 100 ms before target onset. All ERPs were generated by averaging epochs. The number of trials in the averaged ERP was similar across conditions, as follows: predictable, high load  $M = 35.35$ ,  $SD = 1.76$ , range: 32 – 38; predictable, low load,  $M = 35.35$ ,  $SD = 1.93$ , range: 30 – 38; unpredictable, high load  $M = 34.80$ ,  $SD = 1.40$ , range: 32 – 37; unpredictable, low load,  $M = 34.45$ ,  $SD = 2.70$ , range: 27 – 38.

**Event-related spectral perturbations (ERSPs).**—ERSPs are event-related changes in spectral power from baseline across a range of frequencies. Using the EEGLAB function *newtimef()*, time-frequency analysis was conducted with Hanning-windowed sinusoidal

wavelets, for which the cycle number linearly increases with frequency, from a minimum of 2 cycles for 3 Hz to 13.3 cycles for 100 Hz. The wavelets were 743 ms in length and overlapped approximately every 20 ms. To compute the ERSPs, log-transformed spectral power in a baseline period of the second half of the ITI (–3 to –2.5 sec relative to sentence onset) was subtracted from the log power during the trial, an interval (–2.5 to 2 sec relative to sentence onset (Makeig, 1993)) that included the following: a portion of the ITI (approximately 500 ms, depending on trial-to-trial jitter), digit onset and offset, and sentence onset and offset. For statistical analysis, mean ERSPs were then extracted in the alpha frequency range (8–13 Hz) for an interval including sentence presentation (0 to 2 sec, relative to sentence onset) for each electrode, participant, and condition.

**Spectral power.**—For epochs time-locked to sentence-final words, it was not feasible to subtract power in the baseline period to calculate an ERSP due to the short length of single-word epochs and the combined factors of jitter in ISIs and differences across trials in the relative timing of sentence-final word onsets and sentence onsets. However, in order to provide an index of spectral power time-locked to sentence-final word processing, absolute spectral power was examined within short epochs time-locked to sentence-final words. First, new 500 ms epochs were created (0 s before and 0.5 s after the onset of the sentence-final word). Power across the entire epoch was then calculated in 0.69-Hz frequency bins using the EEGLab function *spectopo()*. For statistical analysis, mean spectral power was then extracted across the alpha frequency range (8 to 13 Hz) for each electrode, participant, and condition.

## Statistical Analyses

Statistical analysis for sentence-final word recognition and digit recall used generalized linear mixed-effects models (GLMM) (Jaeger, 2008). All analyses were conducted using R open-source statistical software (R Development Core Team, 2013), and analysis of behavioral accuracy used the lme4 package for linear mixed-effects models (Bates, 2005). Electrophysiological data (ERSPs and ERPs) were analyzed with repeated-measures ANOVAs including within-subjects factors of cognitive load (low, high), sentence predictability (predictable, unpredictable), and scalp region (midline, left, right).

## Results

### Behavior

Mean accuracy for word and digit responses are shown in Table 1. Accuracy for words and digits was analyzed using GLMM with binomial link function using the lme4 package in R. Fixed factors were load and predictability. The model included the random effects structure that was justified by the data (Baayen, Davidson, & Bates, 2008). For the model of word accuracy, this structure included by-subjects and by-item random intercepts and by-subject random slopes for predicability. For the model of digit accuracy, this structure included by-subjects random intercepts and by-subject random slopes for load.

As shown in Table 1, sentence-final words were identified more accurately in predictable than unpredictable sentences ( $\beta = -2.26$ ,  $SE = 0.15$ ,  $p < .0001$ ), with a mean

context benefit of approximately 25 percent more accurate word recognition in predictable sentences. No other effects were significant for word identification accuracy (all  $z < 1$ ). Digits were identified more accurately under low load, that is, on trials for which there were fewer digits to remember ( $\beta = 3.14$ ,  $SE = 0.27$ ,  $p < .0001$ ). Importantly, digits were also identified more accurately on trials in which the spoken sentence was predictable ( $\beta = -0.33$ ,  $SE = 0.11$ ,  $p < .01$ ). However, in a re-analysis of digit recall that included only trials for which the word recognition response was correct, the effect of predictability was no longer significant ( $\beta = -0.16$ ,  $SE = 0.12$ ,  $p = 0.19$ ). The effect of load remained significant ( $\beta = 3.33$ ,  $SE = 0.30$ ,  $p < .0001$ ).

## ERSPs

The scalp distribution of alpha was strongest at centro-posterior sites, as is typical (Obleser & Weisz, 2011; see Supplementary Figure 3). Thus, centro-posterior sites distributed across the midline, left, and right lateral regions were chosen for analysis of the ERD (Cz, Pz, P0z, Oz, CP3, P3, P03, 01, CP4, P4, P04, O2 (Luu & Ferree, 2005)). Figure 1 shows the event-related change in alpha activity within a trial for each condition, averaged across electrodes. Event-related desynchronization (ERD) of alpha power began with the onset of the visual digits and continued through the presentation of the spoken sentence.

Figure 2 shows the mean alpha ERD during presentation of the spoken sentences. During sentence presentation, alpha ERD appears greatest in the condition with the greatest hypothesized cognitive demand, that is, low predictability sentences under high cognitive load. Similarly, alpha ERD is least evident in the condition with the least hypothesized cognitive demand, that is, high predictability sentences under low cognitive load. Statistical analysis confirmed that alpha ERD was significantly smaller (less negative) under low load compared to high load [ $F(1,19) = 7.39$ ,  $MSE = 4.51$ ,  $p < .02$ ,  $ges = .043$ ], consistent with greater cognitive spare capacity under low cognitive load. Alpha ERD was also significantly smaller (less negative) during listening to predictable sentences than unpredictable sentences [ $F(1,19) = 12.81$ ,  $MSE = 1.68$ ,  $p < .003$ ,  $ges = .028$ ], indicating greater cognitive spare capacity when sentences were predictable. The interaction between the load and predictability factors was not significant [ $F(1,19) = 3.10$ ,  $p = .095$ ].

In order to explore the consistency of the main effects of load and predictability on the alpha ERSP among individual participants, difference scores for each factor were calculated for each participant. As can be seen in the figures of Supplementary Digital Content 1 and 2, the difference scores for both main effects were in the expected direction for the majority of participants (15 out of 20 for the load factor; 17 out of 20 for the predictability factor).

## Spectral Power

Figure 3 shows absolute spectral power in the alpha band during epochs time-locked to the onset of sentence-final words. Statistical analysis was consistent with the ERSP analysis across presentation of the entire sentence (above), in that alpha power during presentation of sentence-final words was reduced under high cognitive load [ $F(1,19) = 18.29$ ,  $MSE = 0.97$ ,  $p < .001$ ,  $ges = .005$ ]. That is, during a time window restricted to the processing of the sentence-final words, raw alpha power was lower under high load than low load. However, a

the effect of predictability was not significant [ $F(1,19) = 3.42$ ,  $MSE = 0.44$ ,  $p = 0.08$ ,  $ges = .0005$ ].

## ERPs

Figure 4 shows mean ERPs time-locked to the onset of sentence-final words at fronto-central sites (Fz, FCz, Cz, F3, FC3, C3, F4, FC4, C4 (Luu & Ferree, 2005)) and at centro-posterior sites (see above for a list of electrodes). At fronto-central sites, a negative potential centered at approximately 500 ms (N400) is evident. At centro-posterior sites, a positive potential centered at approximately 800–900 ms (P300 or late positive complex (LPC)) is evident. Mean amplitude and 50 percent area latency of the N400 and P300/LPC potentials were analyzed within time windows and across electrodes determined based on visual inspection and the typical time course and scalp distribution of each potential (Luck, 2014). The N400 time window was set as 400 – 600 ms, as is typical for this potential (Kutas & Federmeier, 2011). The P300/LPC time window was set as 700 – 1,000 ms as in prior research on spoken word recognition (Hunter, 2016; Woodward, Owens, & Thompson, 1990).

The amplitude of the N400 appears to differ across conditions, with unpredictable sentences having a greater negative deflection than predictable sentences. The N400 also appears to be slightly larger and/or to have an earlier latency under low memory load, particularly for low-predictability sentences. For N400 amplitude, there was no statistical support for either visual impression ( $F_s < 2$ ). However, N400 latency was significantly shorter under low load than high load ( $F(1,19) = 11.53$ ,  $MSE = 286.68$ ,  $p < .003$ ,  $ges = .061$ ). A three-way interaction of load, predictability, and region ( $F(1,38) = 3.84$ ,  $MSE = 56.00$ ,  $p < .04$ ,  $ges = .008$ ), followed up at each level of predictability, showed that the effect of load on N400 latency was significant in low-predictability sentences ( $F(1,19) = 13.98$ ,  $MSE = 300.47$ ,  $p < .003$ ,  $ges = .123$ ) but not in high-predictability sentences ( $F < 1$ ). Among low predictability sentences, load interacted with region, and when followed up at each region separately, yielded a significant effect of load at left hemisphere ( $F(1,19) = 17.26$ ,  $MSE = 208.05$ ,  $p < .001$ ,  $ges = .326$ ) and midline sites ( $F(1,19) = 6.03$ ,  $MSE = 154.88$ ,  $p < .03$ ,  $ges = .080$ ) that was marginal at right hemisphere sites ( $F = 4.18$ ,  $p < .06$ ).

The amplitude of the P300/LPC appears to differ across conditions, with greater amplitude under low than high cognitive load, and greater amplitude for predictable than unpredictable sentences. These observations were confirmed statistically by main effects of load ( $F(1,19) = 6.10$ ,  $MSE = 5.41$ ,  $p < .03$ ,  $ges = .037$ ) and sentence predictability ( $F(1,19) = 4.91$ ,  $MSE = 7.69$ ,  $p < .04$ ,  $ges = .042$ ) on P300/LPC amplitude. For P300/LPC latency, an interaction of predictability with region ( $F(2,38) = 3.54$ ,  $MSE = 41.86$ ,  $p < .04$ ,  $ges = .003$ ) was followed up at each region, yielding a significant effect of predictability at left hemisphere sites ( $F(1,19) = 4.57$ ,  $MSE = 317.64$ ,  $p < .05$ ,  $ges = .045$ ), where latency was shorter for high-predictability ( $M = 851.44$ ,  $SD = 17.88$ ) than low-predictability ( $M = 858.41$ ,  $SD = 16.07$ ) sentences.

## Discussion

The current study examined the impact of cognitive (memory) load and sentence predictability on electrophysiological and behavioral indices of cognitive demand in order

to track cognitive spare capacity during effortful listening in a sentence recognition task. The two main aims were to (a) use a memory load manipulation to identify and track the directionality of electrophysiological and behavioral effects of cognitive demand in a sentence recognition task, and (b) determine whether the use of sentence predictability during speech processing would affect the same measures, and if so to compare the effects of load and predictability to infer whether predictable or unpredictable sentences were more cognitively demanding. Both cognitive load and sentence predictability impacted alpha oscillatory power, confirming that both factors impact an electrophysiological measure of cognitive demand during spoken language processing. Both factors also modulated event-related potentials time-locked to sentence-final words, most notably the P300/LPC response. Behavioral responses were also sensitive to cognitive load and sentence predictability.

### Alpha oscillations

With respect to oscillatory EEG indices of cognitive spare capacity, the initial aim of the current study was to confirm that the cognitive load manipulation would modulate alpha oscillatory power measured at scalp electrodes during speech processing in a sentence recognition task, and if so, to determine in what direction cognitive load modulated alpha power. Alpha power was of particular interest because it has emerged in recent literature as a potential index of cognitive spare capacity, having been shown to reflect both cognitive load and changes in listening effort as a function of speech signal degradation (Obleser et al., 2012). Given that the design used in the current study measured alpha during attentive listening to speech in the absence of background noise, it was predicted that a decrease in cognitive spare capacity under high cognitive load would be indexed by an ERD of alpha power in task-relevant neural areas, rather than an ERS in task-irrelevant areas. Indeed, alpha power during sentence processing was desynchronized relative to baseline, and greater cognitive load during spoken sentence processing was accompanied by greater alpha ERD. The results indicate that alpha measured during attentive listening to degraded speech in the absence of background noise reflects task-relevant neural processing, such that a decrease in alpha power reflects a decrease in available capacity (Klimesch, 1999).

Notably, the observed ERD in alpha power as a function of cognitive load was a change relative to baseline in the opposite direction as the ERS that has been observed when alpha is measured during a stimulus-free delay period during which spoken items are held in memory (see Obleser et al., 2012). This difference is likely due to the extent to which the different task demands engage active processing in task-relevant brain areas versus inhibition of processing in task-irrelevant areas. Specifically, holding items in working memory during a delay in which no stimuli are presented is a task that requires inhibition of task-irrelevant processing to maintain items in memory, and during this task cognitive spare capacity appears to be indexed by alpha ERS (Freunberger, Werkle-Bergner, Griesmayr, et al., 2011; Jensen, Gelfand, Kounios, et al., 2002; Obleser et al., 2012). In contrast, attentive listening to spoken language requires activation of task-relevant neural areas, and the observed ERD as a function of cognitive load suggests that during this task changes in cognitive spare capacity are reflected by alpha ERD, as in other tasks that require active, attentive processing (Edwards et al., 2009; Klimesch, 1999; Krause et al., 1997; Palva & Palva, 2007; Pfurtscheller & Da Silva, 1999). However, even during sentence processing,

top-down factors can also trigger functional inhibition. For example, in a recent study in which to-be-ignored spoken stimuli were presented in quiet, it was found that the need to ignore the speech as task-irrelevant triggered the functional inhibition mechanism of alpha, causing listening effort to be indexed by alpha synchronization (Wöstmann, Lim, & Obleser, 2017). The current observation of alpha power desynchronization, as opposed to synchronization, may also reflect the choice to degrade speech with noise vocoding rather than background maskers. In designs that use background noise to degrade speech signals, the functional inhibition mechanism of alpha may be triggered by the maskers (Strauß et al., 2014). In order to fully develop paradigms for research and clinical measurement of alpha oscillations as an index of cognitive spare capacity, further work is needed to systematically examine the impacts of top-down cognitive-linguistic and task-relevance factors, as well as bottom-up factors such as type and degree of stimulus degradation. An ideal task design for an electrophysiological measure of cognitive spare capacity will likely be one in which both top-down, cognitive-linguistic and bottom-up, signal-related factors modulate electrophysiological indices in the same direction.

The second aim of the current study was to compare the effects of cognitive load and sentence predictability, and to infer whether electrophysiological measures indicated a decrease or an increase in cognitive spare capacity when sentences were predictable. Although it has long been known that speech perception accuracy in adverse listening conditions is boosted by sentence predictability, an open question has been whether that accuracy boost comes at the cost of using limited cognitive resources, for example via active processing of sentence meanings in working memory, or whether instead, predictable contexts reduce the effort needed to process spoken sentences. Previous studies have yielded mixed results. Recent studies of subjective ratings of listening effort (Holmes et al., 2018), objective downstream behavioral measures of cognitive spare capacity (Hunter & Pisoni, 2018), and physiological pupil size measures (Winn, 2016) have indicated that cognitive spare capacity is greater during listening to predictable than unpredictable sentences, indicating that predictable contexts increase available cognitive capacity. However, other studies have observed correlations of working memory scores with the extent of the accuracy boost in the presence of semantic context or semantic cues, such that individuals with higher working memory scores benefited more from meaningful sentence context or cues (Zekveld et al., 2011, 2013). These results suggest that cognitive capacity is used in order to get a context benefit in speech perception (for a review, see Besser et al., 2013).

In the current study, electrophysiology was used to track measures of cognitive demand during listening to predictable and unpredictable sentences. It was observed that alpha ERD was greater for unpredictable than predictable sentences. To interpret this effect in terms of cognitive spare capacity a linking assumption is needed as to whether the effect was generated by an alpha ERD that was greater for unpredictable sentences, or instead an ERS that was greater for predictable sentences. The former would indicate that unpredictable sentences were more cognitive demanding than predictable sentences, and the latter would indicate the opposite. There are multiple reasons to infer that the predictability effect in the current study is driven by an ERD. First, alpha ERD reflects task-relevant processing and sentence contexts were highly relevant to the sentence-final word recognition task, as shown by the large behavioral effect of predictability. Also consistent with task-relevant

neural processing, the overall spectral power change during the experimental trials was a desynchronization relative to baseline (see Figure 1). Further, during sentence processing in the same time window, the effect of cognitive load on alpha power was an ERD that was greater under high than low cognitive load, such that a larger ERD corresponded to greater cognitive load and hence lower cognitive spare capacity (see Figure 2). Finally, converging findings from the P300/LPC potential and the behavioral data (see below) indicate reduced cognitive demand when sentences were predictable. Thus, it seems reasonable to infer that the effect of predictability most likely reflects an alpha ERD that was reduced during listening to predictable sentences. As such, the findings indicate that sentence predictability, similar to a reduction in memory load, led to a decrease in cognitive demand, and hence an increase in cognitive spare capacity, during sentence processing. This suggests that the well-known boost in word recognition accuracy for predictable sentences is accompanied by a boost in cognitive spare capacity during listening.

Finally, absolute spectral power was used to quantify alpha oscillatory power time-locked to sentence-final words, complementary to the ERSP analysis that focused on the sentence as a whole. The analysis of absolute alpha power showed a significant reduction under high load compared to low load. This confirmed that the effects of cognitive load on sentence processing persisted through the final word of the sentence. The main effect of predictability on absolute alpha power was not significant, although the pattern of means was in the expected direction. That is, the main effect of predictability was significant when examined across the full sentence in the ERSP analysis, but was not significant when examined as absolute alpha power in a time interval restricted to the sentence-final word. These results are consistent with the idea that sentence predictability influences cognitive demand throughout a sentence, rather than uniquely during processing of the final, most predictable word. Predictable sentences used in the current study generally contained semantically-related content words prior to the sentence-final word (e.g., “The girl swept the floor with a broom”), whereas unpredictable sentences did not (e.g., “Ruth’s grandmother discussed the broom”). Thus, the sentence stimuli in the current study would have enabled an influence of predictability on sentence processing prior to the sentence-final word. This is in line with prior evidence that lexical-level predictions build incrementally as evidence accumulates during language processing (Berkum, Hagoort, & Brown, 1999; Federmeier, 2007; Payne, Lee, & Federmeier, 2015).

### **Event-related potentials**

Analysis of event-related potentials focused on the N400 semantic context effect and the P300/LPC. With respect to N400 amplitude, an unexpected null result was that this potential was not significantly modulated by sentence predictability, although a visible trend in the expected direction was noted. Similarly, N400 amplitude was not modulated by cognitive load. However, cognitive load did have an effect on the latency of the N400 for unpredictable sentences, such that latency was shorter under low than high load. This finding suggests that the timing of cascading activation of semantic information during sentence processing may be sensitive to cognitive resource availability. This would be consistent with prior studies in which a processing load or dual-task was used to manipulate cognitive



resource availability during word processing (Hohlfeld, Mierke, & Sommer, 2004; Hohlfeld, Sangals, & Sommer, 2004; D'Arcy, Connolly, & Hawco, 2005).

With respect to the P300/LPC potential, given the well-established sensitivity of this potential to attention and working memory resources (cf. Polich & Kok, 1995), it was expected that the P300/LPC amplitude would index effects of both cognitive load and sentence predictability on cognitive spare capacity during processing of sentence-final words. Indeed, P300/LPC amplitude was reduced under high cognitive load and was also reduced for unpredictable sentences compared to predictable sentences. It is unlikely that either effect could be attributed to overlapping effects in the N400 time window, given the difference in scalp topography of the two potentials, the late time window of the P300/LPC (700 – 1,000 ms), and the absence of significant effects of either load or predictability on amplitude during the N400 time window. Thus, the modulation of P300/LPC amplitude by both cognitive load and predictability indicates that attentional resource availability was reduced during processing of the final word of spoken sentences both under cognitive load and when sentences were not predictable. In line with the observed alpha oscillatory dynamics, these findings with respect to the P300/LPC support the conclusion that predictive context increases ease of listening in adverse listening conditions.

### Behavioral responses

In addition to basic findings that words would be recognized more accurately in predictable than unpredictable sentences, and that a single digit would be recalled more accurately than several digits, it was predicted that pre-load digits would be recalled more accurately following predictable than unpredictable sentences, consistent with a downstream benefit of release of cognitive resources due to sentence context. As predicted, the behavioral data indicated that digits were more accurately remembered on trials in which the spoken sentence was predictable, replicating a finding of Hunter & Pisoni (2018). However, it was also observed that when trials with incorrect responses on the sentence-recognition task were excluded from analysis of digit recall, the effect of predictability on digit recall was no longer significant. A possible explanation for these results is that it was not unpredictable sentences but rather the failure to report the sentence-final word that disrupted digit recall. On the other hand, removing trials on which sentence-final words were not recognized arguably removes from analysis the most difficult trials, for which the most effort would have been needed, and this could also account for the null result once these trials were removed.

Also based on the prior study, it was expected that behavioral accuracy for the recognition of sentence-final words would decrease under high cognitive load, reflecting a syphoning off by the digit processing task of attentional resources needed for word recognition. However, no effect of cognitive load on word recognition accuracy was observed in the current study. The reason for this discrepancy between the studies is not clear, although it is worth noting that in the previous study the modulation of word recognition accuracy at the level of vocoding that matches that of the current study (i.e., eight spectral channels) was rather small, approximately a three percent difference across high and low load conditions. Also,

given that words were always reported prior to digits in the current study, listeners may have prioritized the word recognition task.

### **Indices of Cognitive Spare Capacity in a Sentence Recognition Task**

Adverse listening conditions, such as background noise or having a hearing loss or cochlear implant, make listening effortful. Standard clinical measures of speech recognition accuracy do not capture the effort needed to reach any given level of performance, and thus fail to index an important impact of hearing loss on quality of life and the ability to function in everyday listening situations. Multiple measures of listening effort and cognitive spare capacity, including subjective, objective behavioral, and objective physiological measures, are currently being developed across multiple laboratories to assess the cognitive demands of listening. For the purpose of assessing the ability to function in complex, everyday listening environments, the most useful measures of the cognitive demand of listening may be those that encompass not only bottom-up demands due to adverse listening conditions but also cognitive and linguistic demands intrinsic to spoken language understanding, such as memory and semantic processing demands.

To this end, the current study examined electrophysiological and behavioral indices of external cognitive and linguistic demand during a sentence recognition task in which spoken sentences were degraded in order to induce listening effort. It was observed that alpha oscillations, the P300/LPC ERP, and behavioral responses in the secondary digit memory task (unless incorrect trials in the sentence recognition task were removed) were all sensitive to cognitive (memory) load and to semantic processing demands (sentence predictability). Alpha oscillations and the P300/LPC are both closely associated with cognitive resource allocation and are related measures that are known to covary (Intriligator & Polich, 1994; Kok, 1997; Polich, 2007; Polich & Kok, 1995; Yordanova et al., 2001). For these reasons, the strongest hypotheses of the current study for the electrophysiological measures were that alpha and the P300/LPC would reflect cognitive load and sentence predictability, and these hypotheses were confirmed.

In contrast, the N400 is not closely associated with cognitive resource allocation. Although the N400 can be modulated by attentional allocation (Deacon & Shelley-Tremblay, 2000), it is also elicited when attention to meaning is minimized (Holcomb, 1988). Insofar as the N400 does reflect the allocation of attention, it may do only insofar as attention heightens the processing of information relevant to meaning (Kutas & Federmeier, 2011). In the current study, N400 amplitude was not modulated by either predictability or load. However, given that absence of a statistically significant effect of predictability on N400 amplitude represents a failure to obtain the well-known N400 context effect, the absence of load effects may reflect a lack of power in the current study. It was found that N400 latency was sensitive to load such that latency for sentence-final words of unpredictable sentences was shorter under low load than high load. This appears to reflect sensitivity to cognitive load in the expected direction, that is, such that the less cognitively demanding condition exhibited a facilitated N400 latency.

The N400 effect that was obtained did not have the same pattern as the alpha and P300/LPC effects, wherein both load and predictability had significant main effects. However, even

if the N400 were to follow the same pattern of modulation by these factors as alpha and the P300/LPC, this would not necessarily indicate that these electrophysiological measures reflect the same underlying components of cognitive demand. As noted above, the N400 may reflect attentional demand only insofar as attention has a downstream influence on the processing of information that feeds into the N400 processing stage (Kutas & Federmeier, 2011). In contrast, alpha oscillations and the P300/LPC appear to be more direct indices of cognitive demand as it is allocated in real time (Klimesch, 2012; Polich, 2007).

A related question is whether any of the electrophysiological effects reflect the same underlying processes as the behavioral measures. As opposed to electrophysiological measures which were measured during processing of the spoken sentences, behavioral responses are known to reflect the sum total of neurocognitive operations leading up to a response, which may include post-perceptual guessing or other response strategies (Balota & Chumbley, 1984; Ratcliff, McKoon, & Verwoerd, 1989). In addition, extensive prior research has dissociated P300 amplitude and latency from response-related processing (see Kok, 2001). In short, behavioral measures are unlikely to reflect exactly the same underlying processes as the electrophysiological measures. Nevertheless, there was broad agreement across the behavioral, event-related potential, and alpha oscillatory data on the basic finding that cognitive demand was reduced under low cognitive load and for sentences that were predictable.

Future work could systematically assess the sensitivity of electrophysiological and behavioral measures to cognitive-linguistic demand across a variety of listening conditions in order to understand the potential for the various measures to capture the combined influence of bottom-up processing demands posed by degraded speech signals and higher-level cognitive and linguistic demands during spoken sentence recognition. Ultimately, the development of tools to track overall level of cognitive demand during spoken language listening in ecologically valid conditions could improve prediction of the capacity of listeners to cope in everyday, adverse listening environments.

### **Predictable Sentences and Cognitive Spare Capacity**

It has long been recognized that people with hearing loss are able to recognize spoken words that are embedded in predictable contexts more accurately than words in isolation or in unpredictable contexts (Kalikow et al., 1977; Bilger et al., 1984). The current findings indicate that sentence predictability also influences listening effort. Sentence predictability modulated alpha oscillations, the P300/LPC, and behavioral responses, such that each measure indicated that cognitive spare capacity was greater when sentence were predictable. These results indicate that listening effort is reduced by predictable sentence contexts. However, it should be noted that word recognition accuracy was also affected by sentence predictability. Given that word recognition accuracy was not equivalent for predictable and unpredictable sentences and that analysis of electrophysiological data included trials with both correct and incorrect sentence recognition responses, the current study has not shown that cognitive spare capacity increased in predictable sentences independently of word recognition accuracy. Further, when trials with incorrect responses on the sentence-

recognition task were excluded from the analysis of digit recall, the effect of predictability on digit recall accuracy was no longer significant.

Importantly, unlike prior studies that relied on behavioral measures alone, the electrophysiological measures used in the current study tracked cognitive spare capacity during speech processing with high temporal precision. Results indicate that when the cognitive demands of listening are measured with brain electrophysiology, predictable contexts can be shown to support ease of listening. Interestingly, in a recent pupillometry study, a similar result was found, in that reduced cognitive demand during and after listening to predictable sentences was observed in young listeners with normal hearing for sentences presented in quiet. However, in that study when stimuli were noise-vocoded as in the current study, pupil size differences were not observed until after sentence offset (Winn, 2016). It may be that the electrophysiological measures used in the current study are more sensitive than pupillometry to real-time changes in cognitive demand when listeners needed to cope with both top-down linguistic demands and bottom-up demands from stimulus degradation. To date, little is known of the relative sensitivity of pupillometry and electrophysiological measures of cognitive demand during listening (although see McMahon et al., 2016; Miles et al., 2017). Finally, a clinical implication of the present findings is that predictable sentence contexts could prove useful for improving ease of listening for patients with hearing loss.

### Limitations

In the current study, accuracy in the primary speech recognition task varied across conditions, specifically across predictable and unpredictable sentences. Thus, from the current results it cannot be determined whether the increase in cognitive spare capacity during listening to predictable as compared to unpredictable sentences that was indicated by the electrophysiological measures is a function of sentence predictability *per se*, or rather the covarying influence of intelligibility, or a combination of both. In many prior studies examining listening effort as a function of speech degradation, intelligibility has also covaried with speech degradation (Dimitrijevic et al., 2017; e.g., McMahon et al., 2016; Obleser & Kotz, 2011; Obleser & Weisz, 2011; Obleser, Wöstmann, Hellbernd et al., 2012). However, unlike these prior studies, speech intelligibility in the current study was not a function of *signal* intelligibility, given that sentences were presented at the same level of degradation. That is, acoustic-phonetically, all conditions were equally intelligible. Moreover, because sentence-final words were counterbalanced across predictable and unpredictable conditions, intelligibility was also not a function of lexical characteristics such as word frequency or neighborhood density (Luce & Pisoni, 1998). In other words, intelligibility was a function of sentence predictability. Given the substantial top-down, linguistic influence of sentence predictability on word intelligibility, it is not possible to equate intelligibility without introducing additional confounds either by removing incorrect trials from analysis (introducing a confound of very different numbers of trials across conditions) or by presenting predictable and unpredictable sentences at different signal-to-noise ratios (introducing a confound of acoustic-phonetic intelligibility across predictability conditions).

A related point is that the current study has provided evidence that both cognitive load and sentence predictability impact cognitive spare capacity during listening to spoken sentences. However, the cognitive demand assessed in this study may have been created by a variety of underlying processes, including but not limited to attentional allocation, arousal, perceptual or response uncertainty, lexical access and selection, and contextual integration. It is generally agreed that the construct of LE is imprecisely defined and that the various measures of LE in the current literature may index differing underlying cognitive processes (see McGarrigle et al., 2014; Ohlenforst et al., 2017; Pichora-Fuller et al., 2016 for reviews on this issue). Given that cognitive spare capacity during listening to speech is defined as those central, fluid cognitive resources that are not currently allocated to effortful listening (see Rudner, 2016), it was not the aim of the current study to delineate the cognitive processes underlying cognitive spare capacity at a mechanistic level. Further, no attempt was made to localize the underlying neural generators of the electrophysiological effects. Particularly given that the underlying sources were not localized, it is likely that multiple cortical generators contributed to the alpha ERD observed at scalp electrodes as a function of cognitive load and sentence predictability. Also, some of the contributing sources may have been modulated in amplitude in the opposite direction as the effect measured at the scalp (i.e. an ERS, presumably from task-irrelevant areas). If so, these contributions were not observable from the scalp measures due to combination with the task-relevant sources that were evidently more numerous, located closer to the scalp, or modulated more strongly as a function of the independent variables. Finally, as discussed above, the alpha oscillatory power, P300/LPC, N400, and behavioral measures used in the current study may reflect somewhat differing underlying sources of cognitive demand.

Finally, the young adults in the current study likely varied to some degree in working memory capacity, oral language skills, and hearing loss. Neither working memory nor oral language skills were assessed independently of the experiment itself. Also, audiological testing was not done to confirm that young adults in the current study indeed had normal hearing as they reported. Although young adult college students are a relatively homogeneous group in these areas, nevertheless individual differences certainly exist in this population. There was no attempt made in the current study to measure or statistically account for such individual differences.

## Conclusions

The current study used electrophysiology to track the cognitive demand of listening to spoken language in adverse listening conditions. Specifically, electrophysiological and behavioral measures of cognitive spare capacity were examined as a function of cognitive (memory) load and sentence predictability in a sentence recognition task in which sentences were spectrally degraded. Modulations of the amplitude of EEG alpha oscillations during listening to spoken sentences as well as the P300/LPC time-locked to sentence-final words both indicated greater cognitive spare capacity under low cognitive load and during listening to predictable sentences. Together with downstream behavioral effects of sentence predictability on digit recall, these findings from the electrophysiological tracking of cognitive spare capacity during listening indicate that predictable sentences increase cognitive spare capacity. Results highlight the potential of highly temporally

precise electrophysiological measures to capture top-down cognitive and linguistic demands intrinsic to spoken language processing in everyday environments. The current results also indicate that predictable linguistic contexts support ease of listening in adverse listening conditions.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

Financial Disclosures / Conflicts of Interest

This research was supported in part by a TL1 postdoctoral fellowship from the National Institutes of Health, Grant TL1TR001107 (A. Shekhar, PI), National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award

## References

- Akeroyd MA (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology*, 47(sup2), S53–S71. [PubMed: 19012113]
- Baayen RH, Davidson DJ, & Bates DM (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.
- Baddeley AD, & Hitch G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47–89). Elsevier.
- Balota DA, & Chumbley JI (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, 10(3), 340. [PubMed: 6242411]
- Banerjee S, Snyder AC, Molholm S, et al. (2011). Oscillatory alpha-band mechanisms and the deployment of spatial attention to anticipated auditory and visual target locations: supramodal or sensory-specific control mechanisms? *Journal of Neuroscience*, 31(27), 9923–9932. [PubMed: 21734284]
- Bar M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11(7), 280–289. [PubMed: 17548232]
- Becker R, Pefkou M, Michel CM, et al. (2013). Left temporal alpha-band activity reflects single word intelligibility. *Frontiers in Systems Neuroscience*, 7, 121. [PubMed: 24416001]
- Berkum JJ van Hagoort P, & Brown CM (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, 11(6), 657–671. [PubMed: 10601747]
- Bernarding C, Strauss DJ, Hannemann R, et al. (2014). Objective assessment of listening effort in the oscillatory eeg: Comparison of different hearing aid configurations. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE* (pp. 2653–2656). IEEE.
- Bess FH, & Hornsby BW (2014). Commentary: Listening can be exhausting—Fatigue in children and adults with hearing loss. *Ear and Hearing*, 35(6), 592. [PubMed: 25255399]
- Besser J, Koelewijn T, Zekveld AA, et al. (2013). How linguistic closure and verbal working memory relate to speech recognition in noise—a review. *Trends in Amplification*, 17(2), 75–93. [PubMed: 23945955]
- Bilger RC, Nuetzel JM, Rabinowitz WM et al. (1984). Standardization of a test of speech perception in noise. *Journal of Speech, Language, and Hearing Research*, 27(1), 32–48.
- Bonhage CE, Mueller JL, Friederici AD, et al. (2015). Combined eye tracking and fMRI reveals neural basis of linguistic predictions during sentence comprehension. *Cortex*, 68, 33–47. [PubMed: 26003489]

- Borghini G, Astolfi L, Vecchiato G, et al. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience & Biobehavioral Reviews*, 44, 58–75. [PubMed: 23116991]
- Bowers AL, Saltuklaroglu T, Harkrider A, et al. (2014). Dynamic modulation of shared sensory and motor cortical rhythms mediates speech and non-speech discrimination performance. *Frontiers in Psychology*, 5, 366. [PubMed: 24847290]
- Cohen MX (2014). *Analyzing neural time series data: theory and practice*. MIT press.
- Conway AR, Kane MJ, & Engle RW (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, 7(12), 547–552. [PubMed: 14643371]
- D’Arcy RC, Connolly JF, & Hawco CS (2005). The influence of increased working memory load on semantic neural systems: a high-resolution event-related brain potential study. *Cognitive Brain Research*, 22(2), 177–191. [PubMed: 15653292]
- Deacon D, & Shelley-Tremblay J. (2000). How automatically is meaning accessed: a review of the effects of attention on semantic processing. *Frontiers in Bioscience*, 5(Part E), 82–94.
- Dikker S, & Pykkänen L. (2013). Predicting language: MEG evidence for lexical preactivation. *Brain and Language*, 127(1), 55–64. [PubMed: 23040469]
- Dimitrijevic A, Smith ML, Kadis DS, et al. (2017). Cortical alpha oscillations predict speech intelligibility. *Frontiers in Human Neuroscience*, 11, 88. [PubMed: 28286478]
- Donchin E, & Coles MG (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, 11(3), 357–374.
- Dryden A, Allen HA, Henshaw H, et al. (2017). The Association Between Cognitive Performance and Speech-in-Noise Perception for Adult Listeners: A Systematic Literature Review and Meta-Analysis. *Trends in Hearing*, 21, 2331216517744675.
- Edwards E, Soltani M, Kim W, et al. (2009). Comparison of time–frequency responses and the event-related potential to auditory speech stimuli in human cortex. *Journal of Neurophysiology*, 102(1), 377–386. [PubMed: 19439673]
- Engle RW (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11(1), 19–23.
- Federmeier KD (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491–505. [PubMed: 17521377]
- Feuerstein JF (1992). Monaural versus binaural hearing: ease of listening, word recognition, and attentional effort. *Ear and Hearing*, 13(2), 80–86. [PubMed: 1601196]
- Francis AL, & Nusbaum HC (2009). Effects of intelligibility on working memory demand for speech perception. *Attention, Perception, & Psychophysics*, 71(6), 1360–1374.
- Friederici AD, Fiebach CJ, Schlesewsky M, et al. (2006). Processing linguistic complexity and grammaticality in the left frontal cortex. *Cerebral Cortex*, 16(12), 1709–1717. [PubMed: 16400163]
- Freunberger R, Werkle-Bergner M, Griesmayr B, Lindenberger U, & Klimesch W. (2011). Brain oscillatory correlates of working memory constraints. *Brain Research*, 1375, 93–102. [PubMed: 21172316]
- Füllgrabe C, & Rosen S. (2016). On the (un) importance of working memory in speech-in-noise processing for listeners with normal hearing thresholds. *Frontiers in Psychology*, 7, 1268. [PubMed: 27625615]
- Gagne J-P, Besser J, & Lemke U. (2017). Behavioral assessment of listening effort using a dual-task paradigm: A review. *Trends in Hearing*, 21, 2331216516687287.
- Gosselin PA, & Gagne J-P (2011). Older adults expend more listening effort than young adults recognizing speech in noise. *Journal of Speech, Language, and Hearing Research*, 54(3), 944–958.
- Hagoort P, Hald L, Bastiaansen M, et al. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, 304(5669), 438–441. [PubMed: 15031438]
- Händel BF, Haarmeier T, & Jensen O. (2011). Alpha oscillations correlate with the successful inhibition of unattended stimuli. *Journal of Cognitive Neuroscience*, 23(9), 2494–2502. [PubMed: 20681750]

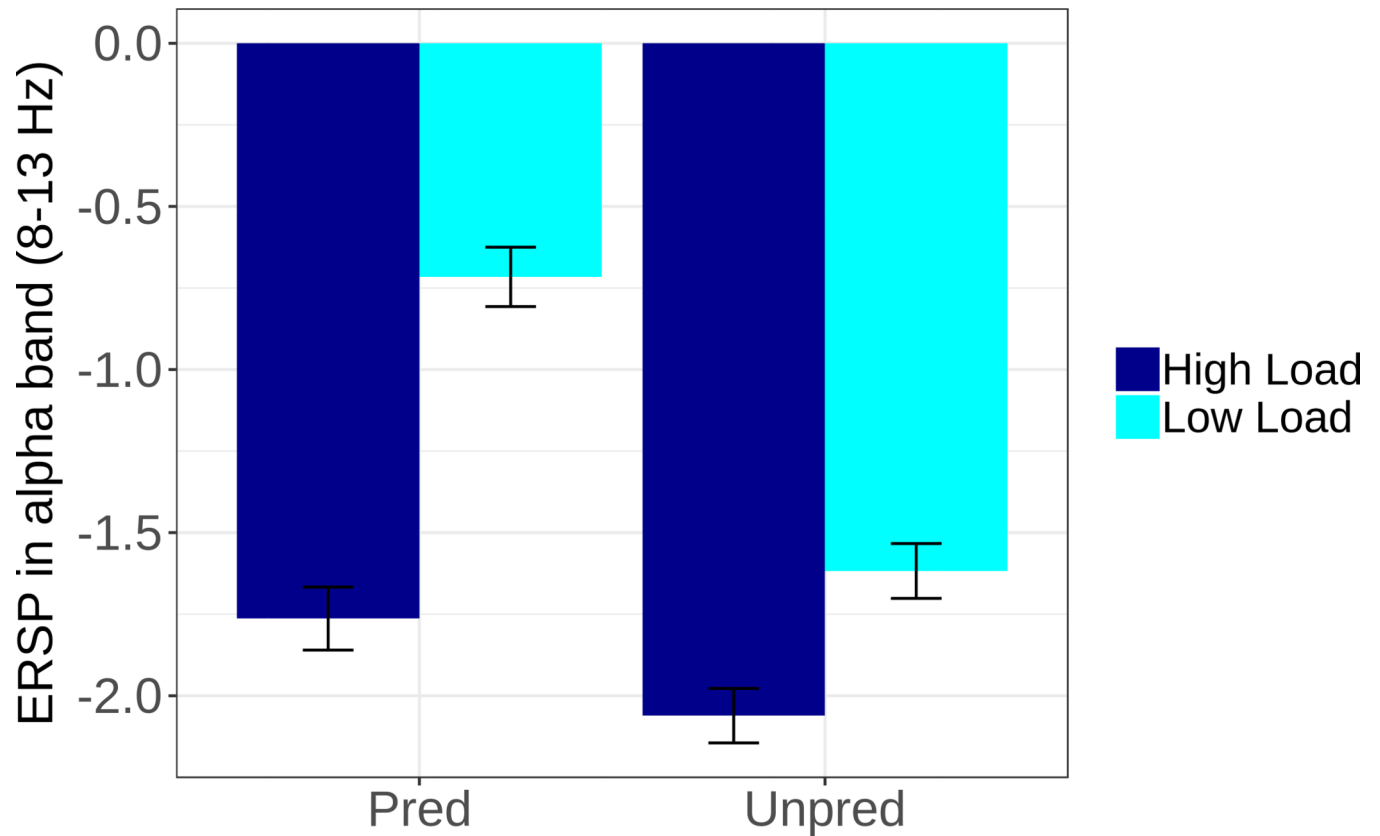
- Hohlfeld A, Mierke K, & Sommer W. (2004). Is word perception in a second language more vulnerable than in one's native language? Evidence from brain potentials in a dual task setting. *Brain and Language*, 89(3), 569–579. [PubMed: 15120547]
- Hohlfeld A, Sangals J, & Sommer W. (2004). Effects of additional tasks on language perception: an event-related brain potential investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(5), 1012.
- Holcomb PJ (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language*, 35(1), 66–85. [PubMed: 3179703]
- Holmes E, Folkeard P, Johnsrude IS et al. (2018). Semantic context improves speech intelligibility and reduces listening effort for listeners with hearing impairment. *International Journal of Audiology*, 1–10. 10.1080/14992027.2018.1432901
- Hornsby BW (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear and Hearing*, 34(5), 523–534. [PubMed: 23426091]
- Hornsby BW, Naylor G, & Bess FH (2016). A taxonomy of fatigue concepts and their relation to hearing loss. *Ear and Hearing*, 37(Suppl 1), 136S. [PubMed: 27355763]
- Huetig F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118–135. [PubMed: 25708148]
- Hunter CR (2016). Is the time course of lexical activation and competition in spoken word recognition affected by adult aging? An event-related potential (ERP) study. *Neuropsychologia*, 91, 451–464. [PubMed: 27616158]
- Hunter CR, & Pisoni DB (2018). Extrinsic Cognitive Load Impairs Spoken Word Recognition in High-and Low-Predictability Sentences. *Ear and Hearing*, 39(2), 378–389. [PubMed: 28945658]
- Intriligator J, & Polich J. (1994). On the relationship between background EEG and the P300 event-related potential. *Biological Psychology*, 37(3), 207–218. [PubMed: 7948466]
- Jenson D, Harkrider AW, Thornton D, et al. (2015). Auditory cortical deactivation during speech production and following speech perception: an EEG investigation of the temporal dynamics of the auditory alpha rhythm. *Frontiers in Human Neuroscience*, 9, 534. [PubMed: 26500519]
- Jensen O, Gelfand J, Kounios J, & Lisman JE (2002). Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cerebral Cortex*, 12(8), 877–882. [PubMed: 12122036]
- Jensen O, & Mazaheri A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, 4, 186. [PubMed: 21119777]
- Kahneman D. (1973). *Attention and effort*. Englewood Cliffs, N.J.: Prentice-Hall.
- Kalikow DN, Stevens KN, & Elliott LL (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the Acoustical Society of America*, 61(5), 1337–1351. [PubMed: 881487]
- Kerlin JR, Shahin AJ, & Miller LM (2010). Attentional gain control of ongoing cortical speech representations in a “cocktail party.” *Journal of Neuroscience*, 30(2), 620–628. [PubMed: 20071526]
- Klimesch W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews*, 29(2–3), 169–195. [PubMed: 10209231]
- Klimesch W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, 16(12), 606–617. [PubMed: 23141428]
- Klimesch W, Sauseng P, & Hanslmayr S. (2007). EEG alpha oscillations: the inhibition–timing hypothesis. *Brain Research Reviews*, 53(1), 63–88. [PubMed: 16887192]
- Koelewijn T, Zekveld AA, Festen JM et al. , (2014). The influence of informational masking on speech perception and pupil response in adults with hearing impairment. *The Journal of the Acoustical Society of America*, 135(3), 1596–1606. [PubMed: 24606294]
- Kok A. (1997). Event-related-potential (ERP) reflections of mental resources: a review and synthesis. *Biological Psychology*, 45(1–3), 19–56. [PubMed: 9083643]
- Kok A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, 38(3), 557–577. [PubMed: 11352145]



- Krause CM, Pörn B, Lang AH, et al. (1997). Relative alpha desynchronization and synchronization during speech perception. *Cognitive Brain Research*, 5(4), 295–299. [PubMed: 9197516]
- Kutas M, & Federmeier KD (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, 4(12), 463–470. [PubMed: 11115760]
- Kutas M, & Federmeier KD (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621–647.
- Kutas M, & Hillyard SA (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203–205. [PubMed: 7350657]
- Kutas M, & Hillyard SA (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307(5947), 161. [PubMed: 6690995]
- Laszlo S, & Federmeier KD (2009). A beautiful day in the neighborhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory and Language*, 61(3), 326–338. [PubMed: 20161064]
- Lau EF, Phillips C, & Poeppel D. (2008). A cortical network for semantics:(de) constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920. [PubMed: 19020511]
- Lavie N. (2005). Distracted and confused?: Selective attention under load. *TRENDS in Cognitive Sciences*, (9)2: 75–82. [PubMed: 15668100]
- Luce PA, & Pisoni DB (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19(1), 1. [PubMed: 9504270]
- Luck SJ (2014). *An introduction to the event-related potential technique*. MIT press.
- Luu P, & Ferree T. (2005). Determination of the HydroCel Geodesic Sensor Nets' Average Electrode Positions and Their 10–10 International Equivalents. Inc, Technical Note.
- Mattys SL, Barden K, & Samuel AG (2014). Extrinsic cognitive load impairs low-level speech perception. *Psychonomic Bulletin & Review*, 21(3), 748–754. [PubMed: 24189992]
- Mattys SL, Brooks J, & Cooke M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology*, 59(3), 203–243. [PubMed: 19423089]
- Mattys SL, Davis MH, Bradlow AR et al. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, 27(7–8), 953–978.
- Mattys SL, & Wiget L. (2011). Effects of cognitive load on speech recognition. *Journal of Memory and Language*, 65(2), 145–160.
- McGarrigle R, Munro KJ, Dawes P, et al. (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*.
- McMahon CM, Boisvert I, de Lissa P, et al. (2016). Monitoring Alpha Oscillations and Pupil Dilation across a Performance-Intensity Function. *Frontiers in Psychology*, 7. 10.3389/fpsyg.2016.00745
- Miles K, McMahon C, Boisvert I, et al. (2017). Objective Assessment of Listening Effort: Coregistration of Pupillometry and EEG. *Trends in Hearing*, 21, 2331216517706396.
- Morton J. (1969). Interaction of information in word recognition. *Psychological Review*, 76(2), 165–178.
- Obleser J, & Kotz SA (2011). Multiple brain signatures of integration in the comprehension of degraded speech. *Neuroimage*, 55(2), 713–723. [PubMed: 21172443]
- Obleser J, & Weisz N. (2011). Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cerebral Cortex*, 22(11), 2466–2477. [PubMed: 22100354]
- Obleser J, Wöstmann M, Hellbernd N, et al. (2012). Adverse listening conditions and memory load drive a common alpha oscillatory network. *Journal of Neuroscience*, 32(36), 12376–12383.
- Ohlenforst B, Zekveld AA, Jansma EP, et al. (2017). Effects of hearing impairment and hearing aid amplification on listening effort: A systematic review. *Ear and Hearing*, 38(3), 267. [PubMed: 28234670]
- Pals C, Sarampalis A, & Ba kent D. (2013). Listening effort with cochlear implant simulations. *Journal of Speech, Language, and Hearing Research*, 56(4), 1075–1084.
- Palva S, & Palva JM (2007). New vistas for  $\alpha$ -frequency band oscillations. *Trends in Neurosciences*, 30(4), 150–158. [PubMed: 17307258]

- Payne BR, Lee C-L, & Federmeier KD (2015). Revisiting the incremental effects of context on word processing: Evidence from single-word event-related brain potentials. *Psychophysiology*, 52(11), 1456–1469. [PubMed: 26311477]
- Peterson EB, Wöstmann M, Obleser J, et al. (2015). Hearing loss impacts neural alpha oscillations under adverse listening conditions. *Frontiers in Psychology*, 6, 177. [PubMed: 25745410]
- Pfurtscheller G, & Da Silva FL (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, 110(11), 1842–1857. [PubMed: 10576479]
- Pfurtscheller G, Stancak A Jr, & Neuper C. (1996). Event-related synchronization (ERS) in the alpha band—an electrophysiological correlate of cortical idling: a review. *International Journal of Psychophysiology*, 24(1–2), 39–46. [PubMed: 8978434]
- Pichora-Fuller MK, Kramer SE, Eckert MA et al. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and Hearing*, 37, 5S–27S. [PubMed: 27355771]
- Pichora-Fuller MK, Schneider BA, et al. (1995). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America*, 97(1), 593–608. [PubMed: 7860836]
- Polich J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128–2148. [PubMed: 17573239]
- Polich J, & Kok A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biological Psychology*, 41(2), 103–146. [PubMed: 8534788]
- Ratcliff R, McKoon G, & Verwoerd M. (1989). A bias interpretation of facilitation in perceptual identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(3), 378.
- Ray WJ, & Cole HW (1985). EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science*, 228(4700), 750–752. [PubMed: 3992243]
- Rönnerberg J, Lunner T, Zekveld A, et al. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, 7.
- Rönnerberg N, Rudner M, Lunner T, et al. (2014). Assessing listening effort by measuring short-term memory storage and processing of speech in noise. *Speech, Language and Hearing*, 17(3), 123–132.
- Rudner M. (2016). Cognitive spare capacity as an index of listening effort. *Ear and Hearing*, 37, 69S–76S. [PubMed: 27355773]
- Smith SL, & Pichora-Fuller MK (2015). Associations between speech understanding and auditory and visual tests of verbal working memory: effects of linguistic complexity, task, age, and hearing loss. *Frontiers in Psychology*, 6, 1394. [PubMed: 26441769]
- Smith SL, Pichora-Fuller MK, & Alexander G. (2016). Development of the Word Auditory Recognition and Recall Measure: A working memory test for use in rehabilitative audiology. *Ear and Hearing*, 37(6), e360–e376. [PubMed: 27438869]
- Strauß A, Kotz SA, Scharinger M, et al. (2014). Alpha and theta brain oscillations index dissociable processes in spoken word recognition. *Neuroimage*, 97, 387–395. [PubMed: 24747736]
- Strauß A, Wöstmann M, & Obleser J. (2014). Cortical alpha oscillations as a tool for auditory selective inhibition. *Frontiers in Human Neuroscience*, 8.
- Theunissen M, Swanepoel DW, & Hanekom J. (2009). Sentence recognition in noise: Variables in compilation and interpretation of tests. *International Journal of Audiology*, 48(11), 743–757. [PubMed: 19951143]
- Verleger R. (1997). On the utility of P3 latency as an index of mental chronometry. *Psychophysiology*, 34(2), 131–156. [PubMed: 9090263]
- Wilsch A, Henry MJ, Herrmann B, et al. (2014). Alpha oscillatory dynamics index temporal expectation benefits in working memory. *Cerebral Cortex*, 25(7), 1938–1946. [PubMed: 24488943]
- Winn MB (2016). Rapid release from listening effort resulting from semantic context, and effects of spectral degradation and cochlear implants. *Trends in Hearing*, 20, 2331216516669723.
- Winn MB, Edwards JR, & Litovsky RY (2015). The impact of auditory spectral resolution on listening effort revealed by pupil dilation. *Ear and Hearing*, 36(4), e153. [PubMed: 25654299]

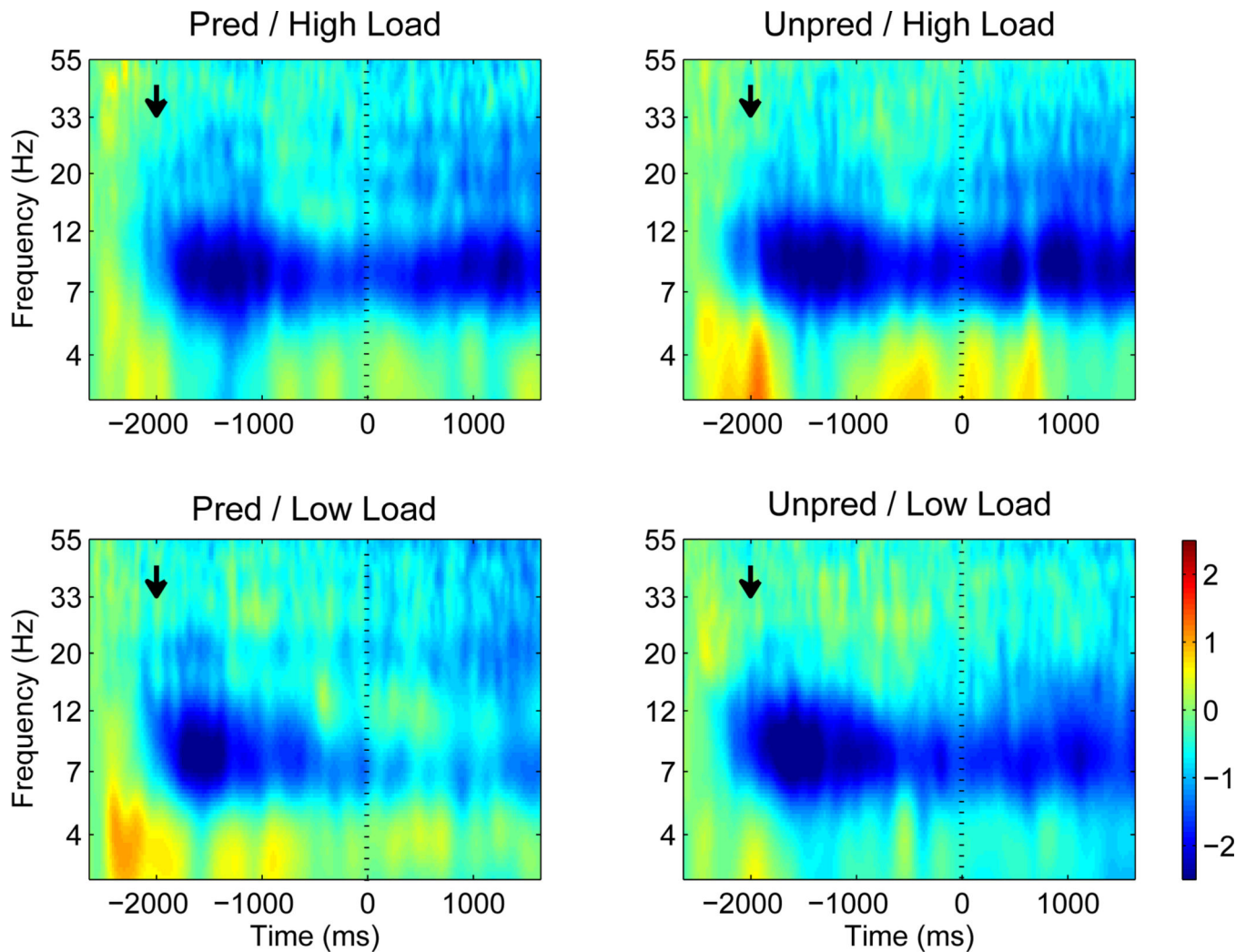
- Wolpaw JR, & McFarland DJ (1994). Multichannel EEG-based brain-computer communication. *Electroencephalography and Clinical Neurophysiology*, 90(6), 444–449. [PubMed: 7515787]
- Woodward SH, Owens J, & Thompson LW (1990). Word-to-word variation in ERP component latencies: Spoken words. *Brain and Language*, 38(4), 488–503. [PubMed: 2375978]
- Wöstmann M, Herrmann B, Wilsch A, et al. (2015). Neural alpha dynamics in younger and older listeners reflect acoustic challenges and predictive benefits. *Journal of Neuroscience*, 35(4), 1458–1467. [PubMed: 25632123]
- Wöstmann M, Lim S-J, & Obleser J. (2017). The Human Neural Alpha Response to Speech is a Proxy of Attentional Control. *Cerebral Cortex*, 27(6), 3307–3317. [PubMed: 28334352]
- Yordanova J, Kolev V, & Polich J. (2001). P300 and alpha event-related desynchronization (ERD). *Psychophysiology*, 38(1), 143–152. [PubMed: 11321615]
- Yuan H, & He B. (2014). Brain–computer interfaces using sensorimotor rhythms: current state and future perspectives. *IEEE Transactions on Biomedical Engineering*, 61(5), 1425–1435. [PubMed: 24759276]
- Zekveld AA, Koelewijn T, & Kramer SE (2018). The Pupil Dilation Response to Auditory Stimuli: Current State of Knowledge. *Trends in Hearing*.
- Zekveld AA, Rudner M, Johnsrude IS, et al. (2011). The influence of semantically related and unrelated text cues on the intelligibility of sentences in noise. *Ear and Hearing*, 32(6), e16–e25. [PubMed: 21826004]
- Zekveld AA, Rudner M, Johnsrude IS, et al. (2012). Behavioral and fMRI evidence that cognitive ability modulates the effect of semantic context on speech intelligibility. *Brain and Language*, 122(2), 103–113. [PubMed: 22728131]
- Zekveld AA, Rudner M, Johnsrude IS, et al. (2013). The effects of working memory capacity and semantic cues on the intelligibility of speech in noise. *The Journal of the Acoustical Society of America*, 134(3), 2225–2234. [PubMed: 23967952]



**Figure 1.**

ERSP response during a trial

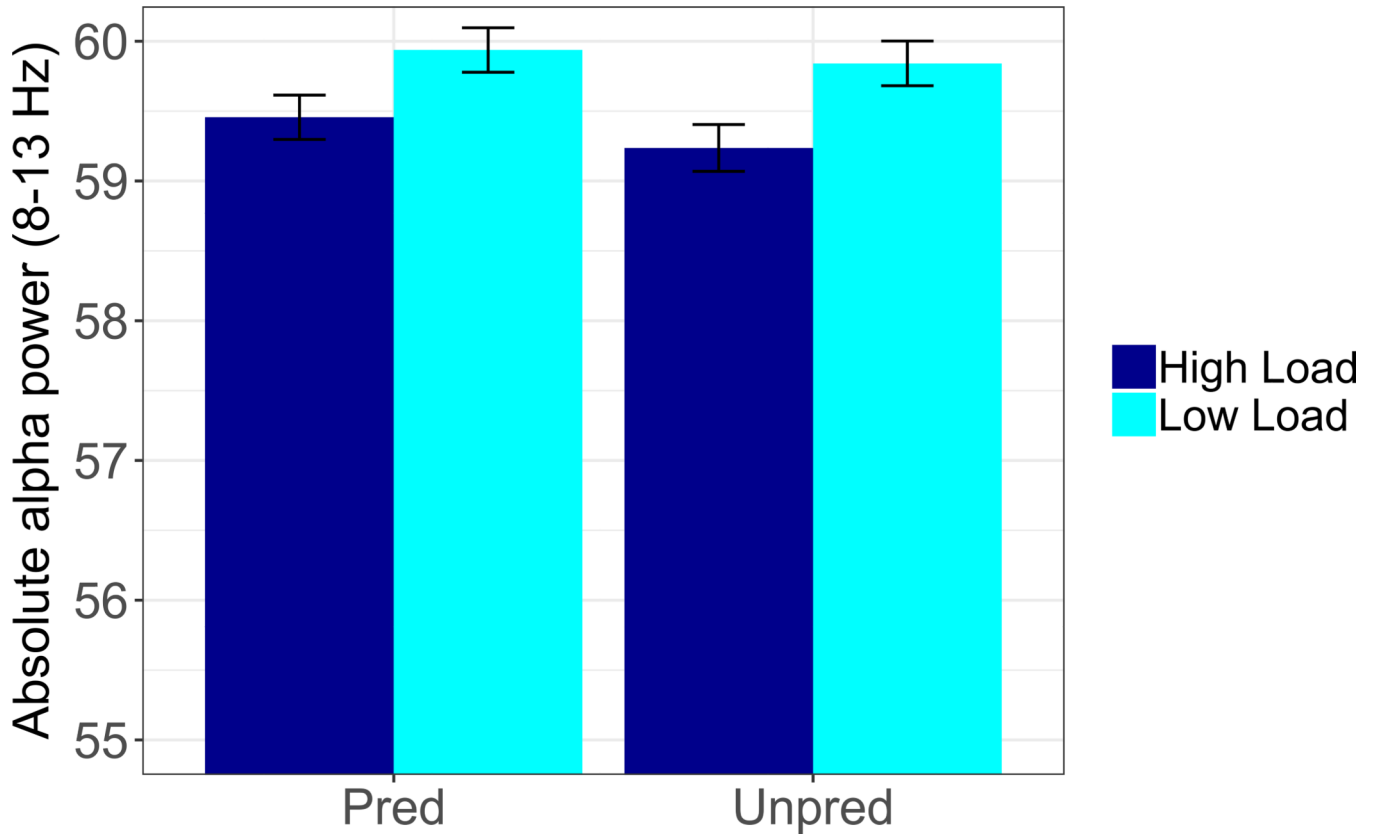
*Note.* Shown is the mean ERSP response across a trial for each level of predictability and memory load. Time range shown on the x-axis is 2.5 sec before and 2 sec after sentence onset. Sentence onset is marked as time zero. Arrow marks approximate onset of visual digits, which appeared approximately 2 sec before sentence onset and remained for 1 sec, followed by an ISI of approximately 1 sec (see text for details). The mean ERSP is collapsed across participants, regions, and electrodes. Pred, predictable; Unpred, unpredictable.



**Figure 2.**

Alpha ERSP during sentence processing

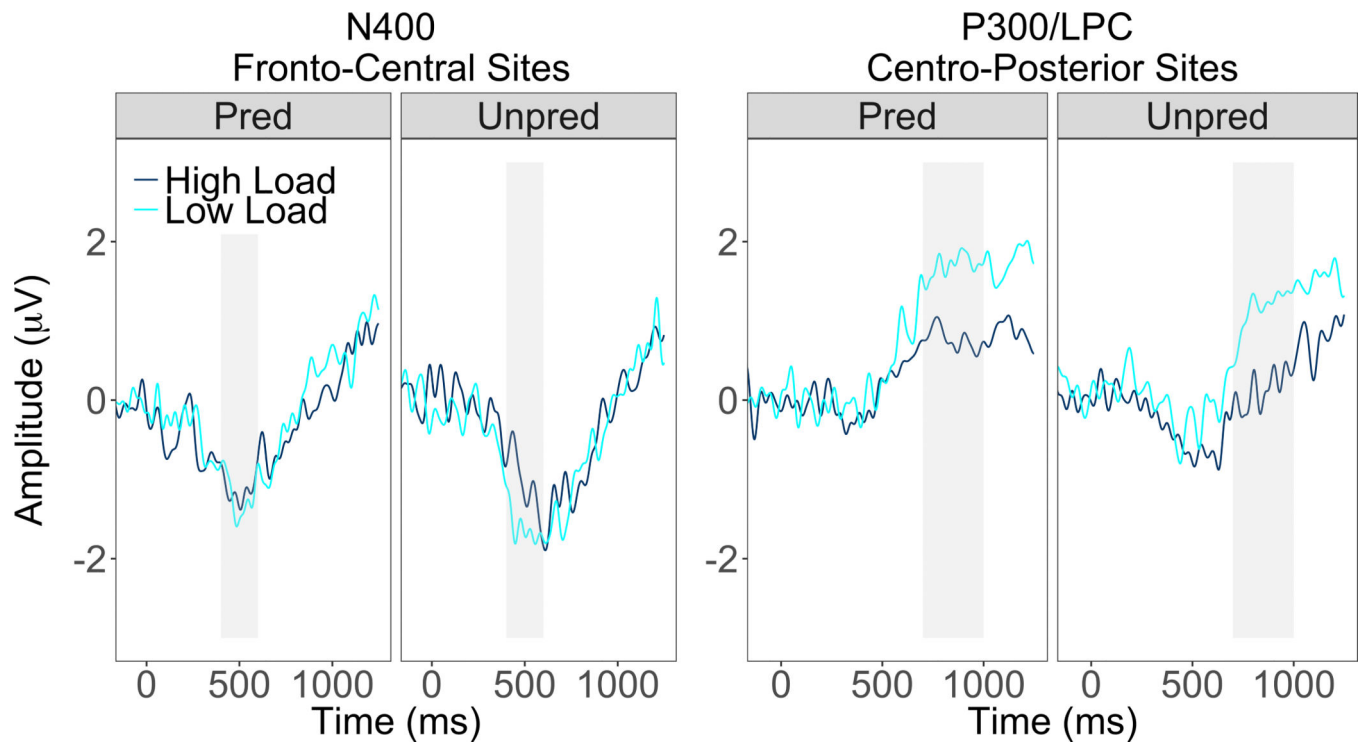
*Note.* Mean alpha ERSP during spoken sentence presentation (0 – 2 sec following sentence onset). Shown is mean ERSP in the alpha band collapsed across participant, region, and electrode. Error bars show  $\pm 1$  SE, where SE is scaled to represent within-subjects variance for the repeated-measures design (Cousineau, 2005). Pred, predictable; Unpred, unpredictable.



**Figure 3.**

Absolute alpha spectral power during sentence processing

*Note.* Mean absolute alpha power time-locked to sentence-final words (0 – 0.5 sec following word onset). Shown is mean power in the alpha band collapsed across participant, region, and electrode. Error bars show  $\pm 1$  SE, where SE is scaled to represent within-subjects variance for the repeated-measures design (Cousineau, 2005). Pred, predictable; Unpred, unpredictable.



**Figure 4.**  
ERP response to sentence-final words  
*Note.* Mean ERP to sentence-final words averaged across participant, region, and electrode. Left, N400 averaged across frontocentral sites; right, P300/LPC averaged across posterior sites.

**Table 1.**

Behavioral accuracy for words and digits.

	Predictable		Unpredictable	
	High Load	Low Load	High Load	Low Load
Words	97.46 (.01)	98.14 (.01)	75.67 (.02)	74.89 (.02)
Digits	37.59 (.05)	91.26 (.02)	34.26 (.04)	86.81 (.02)

*Note.* Shown is accuracy with standard error in parentheses for sentence-final words and pre-load digits at each level of sentence predictability and memory load.