Age-Related Differences in the Intra-Individual Neural Timing Variability of Speech-Evoked Brain Potentials

Carryl L Baldwin*, Daniel M Roberts and Brian A Taylor
Department of Psychology, George Mason University, USA
Submission: July 28, 2017; Published: August 30, 2017
*Corresponding author: Carryl L Baldwin, Department of Psychology, George Mason University, USA Fax: 7039931384, 7039931359; Email: cbaldwi4@gmu.edu

Abstract
Event-Related Potentials (ERP) often indicates differences in language-processing between age groups. However, trial-to-trial variability in ERP component latency (latency jitter) may contribute to observed amplitude reductions in the average ERP. We examined the role of latency jitter in the observation of age-related changes in speech processing as indexed by the N400. Older and younger adults listened to sentences ending in either expected or unexpected final words. While a reduction in N400 amplitude was observed in older vs. younger adults, analysis on a single-trial basis revealed older adults had significantly greater intra-subject variability in N400 latency for unexpected speech stimuli than young listeners. Increased variability contributes to the smaller N400 amplitudes in the average ERP of older subjects, and indicates less precise timing of neurophysiological processes, or conversely greater specificity in the processing of individual sentence stimuli.

Descriptors: Event-related potentials, Aging, N400, Latency Jitter

Abbreviations: ERP Event-Related Potentials; ITC: inter-trial coherence; ERSP Event-Related Spectral Perturbation; LAN: Lateral Anterior Negativity; SPL: Sound Pressure Level; ISI: Inter-Stimulus Intervals; LAN: Lateral Anterior Negativity; MMSE: Mini-Mental State Examination

Introduction
The processing of speech involves the interplay of a highly complex cascade of processes, and research in this domain has only begun to uncover its dynamic interaction with aging. The elderly are particularly susceptible to challenges associated with difficulties in speech processing, as a considerable segment of this group experience accelerated declines in sensory, perceptual, and cognitive faculties that can interact negatively with speech comprehension Morse [1] McCoy [2]. Aging appears to bring about challenges in language processing when online processing is required Light [3] Sommers [4]. Yet in many aspects of language processing, older adults continue to perform at levels comparable to younger adults, even outperforming them in certain tasks. Specifically, word-knowledge is retained, and may even continue to improve with advancing age Horn [5] Wingfield & Tun [6]. Fluid language processing abilities have also been shown to be stable with age, such as when using context to facilitate word recall Burke [7] Light [8] Salthouse [9].

Age-Related Changes in Speech Processing
How and why speech processing changes with age has been the subject of a large number of theories and associated investigations. Some posit general changes while others suggest process specific explanations. Classically, it has been proposed that a general slowing of cognitive and perceptual processing rates may underpin most age-related changes Cerella [10] Salthouse [11], with concomitant implications for speech processing. Speech processing requires rapid perception, comprehension, working memory maintenance, and both access and retrieval of semantic information from long-term memory. Additionally, evidence for age-related deficits in the ability to inhibit both exogenously and endogenously-generated irrelevant stimuli have been suggested to contribute to age-related speech processing difficulties Baldwin [12] Hashe [13] Burke [14].

Other evidence suggests that deficits in working memory capacity lead to comprehension difficulties Carpenter [15] or that it is the interaction between age-related changes in sensory processes and cognitive changes in attention and working memory that lead to speech processing challenges in the elderly Baldwin & Ash [12] Wingfield [16]. Age-related declines in sensory functioning may contribute to a “noisier” neural signal, leading to greater variability in the timing and accuracy of language processing in older adults Baldwin [17] Olsho [18]. Successful interpretation of this degraded stimulus comes at an attentional cost Baldwin [12] McCoy, et al. [2] Murphy [19];
Pichora-Fuller [20]. Referred to as the effortfulness hypothesis, this increased effort leaves fewer resources available for the processing of speech Wingfield, et al. [21] Rabbitt [22].

These various hypotheses continue to garner support as explanations for age-related changes in speech and language processing. Furthermore, their non-orthogonality suggests that it is possible-if not likely-that reality lies in some combination Benichov [23]. What is for certain, however, is that older adults frequently report difficulties with comprehending spoken language Gordon-Salant, Fitzgibbons [24]. With age, a significant proportion of adults will experience declines in sensory and perceptual systems, as well as cognitive processing, which can all contribute to problems in speech processing, particularly those reliant on fluid intelligence processes.

ERPs, Language, and Cognitive Aging

Event-Related Potentials (ERPs) have frequently been utilized in investigations of the neural mechanisms of speech processing and the influence of age on these mechanisms. The high temporal resolution of ERPs makes them particularly attractive for the study of the timing of speech-comprehension processes. Components of the ERP related to syntactic processing include the Lateral Anterior Negativity (LAN), the “early” LAN (ELAN), and the syntactic positive shift (SPS, or P600). A component known to be associated with automatic semantic processing of language stimuli is the N400, a negative-going deflection peaking at about 400 ms post-stimulus Kutas, Hillyard [25]. In a common paradigm, the N400 is elicited through manipulating the semantic agreement of a sentence-final word with the preceding sentence (see Table 1). Context facilitates semantic processing-decreasing the amplitude of the N400 (Van Petten, Luka [26].

Table 1: Example sentence stimuli Sentences ending in a semantically incongruent final word elicit larger N400 ERP components than stimuli ending in congruent final words. The effect is stronger in younger adults than in older adults.

<table>
<thead>
<tr>
<th>Sentence Type</th>
<th>Example Stimulus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congruent</td>
<td>The man swerved and drove his car into the ditch.</td>
</tr>
<tr>
<td>Incongruent</td>
<td>The man swerved and drove his car into the sky.</td>
</tr>
</tbody>
</table>

A larger (more negative) amplitude N400 component is observed with greater semantic incongruence. Evidence from studies leveraging the N400 ERP component to assess semantic processing deficits indicates that older adults may have difficulties using semantic context efficiently during online language processing Federmeier, McLennan [27]. It has been regularly observed that older adults exhibit reduced N400 component amplitudes to semantic violations compared with younger adults, commonly interpreted as evidence of older adults using context less efficiently than younger adults [28-31] (Table 1).

Seemingly in contrast, behavioral evidence finds that older adults make strong use of semantic context, and more so than younger adults, particularly when stimuli are speeded in time or degraded Madden [32] Pichora-Fuller, Schneider [33] Rogers, Jacoby [34] Sommers [35] Sommers & Danielson [4] Stine- Morrow, Loveless [36]. The challenge, therefore, has been in integrating the age-related differences in semantic context facilitation apparent at the fine temporal level made observable by ERPs with the gross level behavioral performance measures that indicate context is being efficiently processed and utilized by older adults. One possibility is that the information necessary for thoroughly interpreting the ERPs is being obscured during the ERP averaging process.

Single-Trial ERP Measures and Latency Jitter

Age-related N400 amplitude reductions have been observed in a number of previous investigations [e.g., Faustmann, et al. [28]; Federmeier, et al. [27] Federmeier, et al. [30] Kutas, Iragui [31]. However, a known caveat within the ERP literature is that if the ERPs of trials within a given condition have high timing variability, also termed latency jitter, ERP peak measurements are not appropriate and can lead to misinterpretation of results Rosenblith [37] Coles, Rugg [38] Luck [39] Spencer [40]. When within subject latency jitter occurs, a given subject’s averaged ERP is less representative of the underlying component, as the offsets in time of each trial serve to flatten and widen the average (see Figure 1). Alternative methods of evaluating the size of the N400 component, such as mean amplitude, offer protection against reductive effects of variability, but only when the measured latency range captures the entire component.
of interest without overlap from neighboring components. Furthermore, mean amplitude estimation on the averaged ERP component fails to capture potentially informative single-trial variability information (Figure 1).

Describes several methods for alleviating the effect of timing variability on ERP amplitude measurements, including using a 50-percent area, rather than peak magnitude, using event-related averages, using an adaptive template such as the Woody filter Woody [41] for identifying the component of interest on single trials, or by investigating ERPs effects in the frequency, rather than time domain by estimating phase and power changes of EEG frequency components via inter-trial coherence (ITC) Rappelsberger, Pfurtscheller [42] and Event-Related Spectral Perturbation (ERSP) Makeig [43] respectively.

The majority of previous investigations on single-trial ERP variability have used the P300 component Polich [44] for a recent review of the P300. For example, Kutas, McCarthy [45] used a single-trial measurement technique on data from a visual oddball task to clarify the functional significance of the P300. Using the aforementioned Woody adaptive filtering technique on single-trial epochs that had been low-pass filtered at approximately 6 Hz, Kutas et al. found that amplitude differences between task conditions that were present with a standard ERP analysis were eliminated after correcting for within-subject latency jitter.

Degree of single-trial ERP variability has further been studied as a possible explanation for aging effects on P300. It has frequently been observed that aging is associated with lower amplitude and longer latency P300 responses Polich [46] Polich [47] Walhovd, Rosquist [48] used a maximum-likelihood estimation method Jaskowski, Verlger [49] to estimate single-trial P300 latency jitter as a possible explanation for age-related differences in P300 amplitude. However, they concluded that differences in latency jitter across age groups could not account for correlations between age and P300 amplitude, suggesting that for the P300 response, age-related amplitude reductions are legitimate. In contrast, Schmiedt-Fehr, Basar-Eroglu [50] used ERSP and ITC measures to quantify the degree of event-related power and phase-locking of delta (0.5-3 Hz) and theta (4-7 Hz) activity during a Go/NoGo task. Younger subjects had both greater theta power and ITC during the P300 eliciting NoGo condition in comparison to older subjects. This result suggests that while reduced P300 responses in older adults aren't entirely due to latency jitter, an increase in timing variability may occur in tandem with real differences in ERP magnitude.

Measurement of ERP response on a trial-by-trial basis is valuable not just in terms of clarifying existing averaged-ERP results, but as an additional source of information on how the time-course of information processing in the brain changes with age. For example, in the semantic processing task traditionally used to study the N400, greater latency jitter may indicate greater processing variability arising from the individual sentence stimuli. A general increase in variability with age independent of semantic congruency would likely be consistent with age-related exogenous changes such as hearing loss as well as deficits in the ability to inhibit irrelevant stimuli. Alternatively, if age-related variability differences interact with congruency, this may indicate endogenous differences in the processing of semantic information. Although the sets of sentence stimuli used in these types of investigations share categorical membership within a condition (i.e., semantically congruent or semantically incongruent), there are still important differences between each sentence. It is plausible to think that there may be age-related differences due to the treatment of these sentences on a superordinate or subordinate categorical level. Older adults, with their well-developed lexical and semantic networks, may have more sentence-to-sentence variability of processing when encountering a word that is not apparently congruent with the sentence-level meaning. This would in turn contribute to the age-related reduction in amplitude of the N400 in the averaged ERPs.

**Current Investigation**

The apparent incongruence between behavioral studies showing older adults use semantic context efficiently and neurophysiologic investigations that seemingly indicate deficits in this domain may be explained, at least in part, by latency jitter. We tested the contribution of latency jitter by conducting a peak-pulling single-trial analysis of the N400. Further, we sought to rule out the possibility that age-group effects are driven by greater noise at the single-trial level in one age group more than another. Therefore, we used a bootstrapping procedure to extract Noise-to-Signal Transition Functions for both of our age groups. These functions permit the examination of the variability as single-trial ERPs are averaged into increasingly greater subset sizes, and approach the grand average ERP.

We hypothesized that older adults would show greater processing variability for speech stimuli compared with younger adults. We predicted that this increase in processing variability would be revealed through greater trial-by-trial variability of N400 ERP component latency in older participants, specifically for incongruent sentence-final words, at the single-trial level. Further, we predicted that the Noise-to-Signal Transition Functions would reveal significantly different intercepts for younger and older adults, with older adults showing consistently greater ERP variability. Greater latency jitter in the older adults compared with the young would support the hypothesis that latency jitter is contributing to the observed N400 amplitude reduction in older adults processing of incongruent sentence stimuli and help to bridge the findings of behavioral and neurophysiological investigations of aging and speech processing.
Methods

a) Participants

Thirty participants, 15 “young” (10 male, 18-27 years of age, M = 21) and 15 “old” (11 male, 61-79 years of age, M = 67) participated in this investigation following voluntary informed consent. “Young” participants were recruited from the undergraduate participant pool for credit in a Psychology course. “Old” participants were recruited from the local area and compensated with a modest cash payment. All participants reported being in good health, reported no known hearing problems, and were native English speakers.

b) Cognitive And Auditory Screening

A battery of cognitive tests was given to screen for cognitive deficits. The test battery consisted of the Mini-Mental State Examination (MMSE) Folstein, Folstein [51] Folstein, Robins [52], the Trail Making Test Part B (TMT-B) Corrigan & Hinkedly [53] Gaudino, Geisler, Squires [54] Lezak, Howieson, Loring [55] Reitan [56], and the Digit Symbol Substitution Task (DSST) subtest of the Wechsler Adult Intelligence Scale III Wechsler [57]. The MMSE is a 30-item questionnaire used to screen for cognitive impairment, such as is typically associated with dementia. Any participant getting less than 26 items correct (indicating some degree of cognitive impairment) was excluded from the study.

The TMT-B is a measure of speed of information processing, involving speed, attention, directed visual search ability, and mental flexibility. The test consists of completing a trail by linking alternating letters and numbers as quickly as possible. Completion of the task in more than 273 seconds is indicative of deficiency. All participants included in this study were able to complete the task in less than 273 seconds, with less than 4 errors.

The DSST is a test where participants are provided with a key including 9 different symbols each with a corresponding number (1-9), and are then given a set of single-digit numbers with corresponding empty boxes. Participants must fill in the empty boxes with as many symbols corresponding to each given number within 120 seconds. The DSST has been shown to be sensitive to cognitive impairment due to age, dementia, and brain injury [58]. An auditory pure tone threshold test was presented monaurally at 250, 500, 1000, 2000, 4000, and 8000 Hz following the revised assessment guidelines adopted by the American Speech and Hearing Association Arlinger [59]. Based on this exam, all participants had age-normal hearing or better.

Materials

The auditory stimuli consisted of 200 sentences (100 congruent and 100 incongruent) taken from the Block and Baldwin [60] set. Each sentence met the criterion of having a cloze probability of 67% or higher. Semantically incongruent sentences were constructed by replacing the high cloze probability sentence final word with an anomalous word. Each sentence was presented via two consecutive .wav files, the first containing everything but the sentence final word, and the second containing the sentence final word. The partitioning of the continuous audio recordings into pre and post sentence final word was performed manually with Adobe Audition software. A marker was sent to the continuous EEG data stream at the start of each .wav file to allow for accurate time-locking to the start of the sentence final word. Although presented from two files, the sentence final word followed seamlessly from the initial sentence portion, with no audible clicks or gap. The digital recordings were normalized and presented through Sony MDR-NC60 noise cancelling headphones at 62.5 dB Sound Pressure Level (SPL).

Procedure

a) Screening: Participants were tested individually in a single session lasting approximately 1/2 hours. Following completion of informed consent and the battery of audiometric and cognitive screening tests described above, participants performed a sentence verification task.

b) Sentence verification task: Participants were seated comfortably in a padded office chair in front of a computer monitor. The participants were asked to keep their eyes open and maintain fixation on centralized cross displayed on the monitor in order to reduce eye movement artifact. Each participant listened to a set of 200 sentences and was asked to indicate via key press whether the sentence final word was congruent or incongruent. Sentence sets contained a randomized order of 100 congruent and 100 incongruent sentences taken from the aforementioned Block and Baldwin [60] high cloze probability sentence set.

As illustrated in Table 1, incongruent sentences ended in an anomalous final word while congruent sentences ended in a more expected high probability cloze word. Participants were asked to restrict their eye-blinks to the periods of silence during the Inter-Stimulus Intervals (ISIs). Participants were instructed to take their time in responding and to be as accurate as possible. The ISI between sentences was jittered between 3 and 8 seconds. Each set of 200 sentences took approximately 27 minutes to complete.

c) EEG Recording: EEG was recorded continuously from 12 Ag/AgCl electrodes. Electrode placement followed the International 10-20 system (Jasper, 1958). Data were recorded from the sites Fz, F3, F4, Cz, C3, C4, Pz, and left and right mastoid points. Eye blinks were monitored using electrodes placed above and below the orbit of the left eye to record vertical electrooculogram (VEOG). The ground electrode was placed at site FPz. All channels were referenced to the left mastoid during recording. Data was recorded at 500 Hz using a Neuroscan NuAmps 40-channel amplifier and Neuroscan Scan 4.4 software. Impedances were maintained below 5 kΩ. Online filtering was set to a bandpass of 0.1 to 70 Hz.
d) ERP Single-Trial Analysis Procedure: The EEGLAB toolbox Delorme, Makeig [61] in conjunction with MATLAB v.2007b (The Math Works, Natick, MA) was used for analysis of the EEG recordings. EEG was re-referenced to the average of the left and right mastoid electrodes and band-pass filtered from 0.1 Hz to 6 Hz. An upper bound of 6 Hz was used as a basic technique to permit peak-pulling of single trial ERPs on the low frequency N400 component. Epochs were time-locked to the onset of the sentence-final word, and ranged from 150 ms pre-stimulus onset to 800 ms post-stimulus onset. The mean of the 150 ms pre-stimulus period was removed from each epoch. Any epochs containing eye blink artifacts, as defined as activity exceeding +/- 50μV on the upper VEOG electrode, were rejected from further analysis.

Based upon visual inspection and reference to previous investigations of the N400 in the auditory modality (for example, Federmeier, et al. [27] a latency window of 300-500 ms was selected for the N400. For each participant, the most negative amplitude during this window was identified for each electrode site at each trial, and both the latency and amplitude of these peak values were recorded. While other investigations have utilized template matching, rather than peak-pulling as a method of estimating single-trial values (for example, Ford, et al. [62], work directly comparing template matching using a half-period sine wave to peak-pulling with low-pass filtered data for estimation of a single-trial P300 component have suggested the two methods generate similar conclusions Smulders, Kenemans [63]. In this case, peak-pulling on low-pass filtered data was utilized as it involves fewer assumptions about the shape of the component on individual trials.

After identifying the latency and amplitude of the peak within the 300-500ms window of each trial, the standard deviation (SD score) of the latency of these peaks across the single-trials were calculated for each participant, at each channel. This value provided an estimate of the variability of the latency of the N400 component for each participant, at each channel, and functioned as the primary dependent measure of interest. Further, the peak amplitude values obtained from the single-trial peaks were averaged for each participant at each channel, providing an estimate of corrected N400 component amplitude. This average single-trial peak amplitude approximates the peak amplitude of an ERP component obtained through traditional averaging, if the underlying components on each trial were aligned with respect to peak component latency.

e) Noise-to-Signal Transition Function procedure. We used a bootstrapping procedure to assess variability scores (SD scores) as a function of ERP subset size. While the averaging across trials undertaken in traditional ERP analysis improves the signal-to-noise ratio, information regarding the distribution of component amplitude and latency across trials is lost. Conversely, single-trial analysis allows for the investigation of variability across trials, but at the cost of a diminished signal-to-noise ratio. To systematically investigate this trade-off with respect to information on component variability, ERPs were generated using variable numbers of trials in each bin average. Participants were first equated by selecting a random subset of 37 single-trial epochs (the minimum number of remaining incongruent epochs for any given participant, out of 100 original).

Next, average ERPs composed of varying numbers of individual trials were created by averaging together a randomly selected group of epochs ranging in size from 1 to 37 epochs per bin. The peak latency in the 300 to 500 ms time window was then pulled from each averaged ERP. This process was repeated 1000 times for every subset size (1 to 37), in order to estimate the trade-off between component variance information and signal-to-noise ratio as the number of epochs included in the ERP average increases. The number of epochs in each bin, or subset size, was plotted against the standard deviation of peak values for that bin, to produce "Noise-to-Signal Transition Functions." A line was fit to the data for each age group and compared.

Results

a) Sentence Verification Performance

The accuracy of congruent/incongruent judgments was assessed to ensure participants were accurately processing the stimuli. As anticipated, overall accuracy was found to be high (97.5%), indicating that participants were likely able to evaluate the sentence stimuli correctly. Response Time was also examined for effects of Age (Young vs. Old) and Sentence Type (Congruent vs. Incongruent) using a 2x2 mixed ANOVA was conducted. There was a significant effect of Sentence Type (F(1,28) = 12.32, p = .002, η2 = .31), with participants responding faster (in milliseconds) to congruent sentences (M = 761, SD = 335) than incongruent sentences (M = 831, SD = 338). There was not a significant effect of Age (F(1,28) = .302, p = .587, η2 = .01), nor a significant interaction effect of Sentence Type and Age (F(1,28) = .835, p = .369, η2 = .029).

b) N400-effect Verification Analysis

![Figure 2: Difference waves for Old versus Young participants recorded at site CZ.](image)
To confirm the presence of the traditional age-related reduction in the grand average N400, we calculated difference waves (Incongruent minus Congruent) for the averaged ERPs for each participant. The effect was assessed across central midline and parietal sites where the N400 is typically most prominent. Age (Young and Old) and Site (C3, CZ, C4, and PZ) were the independent measures. Grand average difference waves for young and old participants at site CZ are displayed in Figure 2. As expected, the N400 component was greater in amplitude (more negative) in younger adults compared with older adults.

c) N400 amplitude

A 2 x 4 mixed design ANOVA was conducted to examine the effects of Age (Young vs. Old) and electrode Site (C3, CZ, C4, and PZ) on peak amplitude. There was a significant main effect of Age on the peak amplitude of the N400 (F(1,28) = 4.456, p = .044, η2 = .14). Older adults had lower magnitude (less negative) peak amplitudes on average than younger adults, M = -4.59 μV (SE = .60) and M = -6.05 μV (SE = .60), respectively. There was also a significant main effect of electrode site (Greenhouse-Geisser corrected due to significant violation of Mauchly’s Test of Sphericity (p < .001); F(2.08,58.27) = 3.38, p = .039, η2 = .11). Bonferroni corrected post-hoc pairwise comparisons revealed that amplitudes at site C3 were significantly less negative (M = -4.53 μV, SE = .44) than either CZ (M = -5.57 μV, SE = .49, p = .043) or PZ (M = -5.70 μV, SE = .57, p = .033). No other site differences were significant (all ps > .1). The interaction between Age and Site was not significant (F(2.08,58.27) = 3.38, p = .039, η2 = .11).

d) N400 latency

A 2 x 4 mixed design ANOVA was conducted to examine the effects of Age (Young vs. Old) and electrode Site (C3, CZ, C4, and PZ) on peak latency. There was not a significant effect of Age (F(1,28) = .11) or Site (F(3,84) = .73), both ps > .1. There was also no significant interaction of Age and Site (F (3,84)=.788, p>.1) (Figure 2).

Single Trial N400 Component Measures

a) N400 Single Trial Peak Amplitude

The amplitude of single trial N400 responses was estimated by comparing the average of the single trial peak amplitudes. A 2 x 2 x 4 mixed design ANOVA was conducted to examine the effects of Age (Young vs. Old), Sentence Type (Congruent vs. Incongruent), and electrode Site (C3, CZ, C4, and PZ) on single-trial N400 Peak Amplitude. There was a significant effect of Age on the N400 single trial peak amplitude, (F (1,28) = 10.519, p=.003, η2 = .273). Older adults had lower magnitude (less negative) single trial peak amplitudes on average than younger adults (M = -2.96 μV (SE = .49) and M = -5.20 μV (SE = .49), respectively. There was additionally a main effect of Sentence Type on single-trial peak amplitude, (F(1,28) = 30.055, p < .001, η2 = .563).

Table 2: Main effect of electrode site on mean single-trial peak amplitude.

<table>
<thead>
<tr>
<th>Electrode Site</th>
<th>Mean (SE) Single-Trial Peak Amplitude (μV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C3</td>
<td>-4.37 (.37)</td>
</tr>
<tr>
<td>CZ</td>
<td>-4.55 (.43)</td>
</tr>
<tr>
<td>C4</td>
<td></td>
</tr>
<tr>
<td>PZ</td>
<td>-3.74 (.32)</td>
</tr>
</tbody>
</table>

The single-trial peaks from trials with congruent sentence-final words were of lower magnitude (M = -3.22 μV (SE = .36)) relative to trials with incongruent sentence-final words (M = -4.95 μV (SE = .38)). Further there was a main effect of electrode Site (F(3,28) = 7.918, p < .001, η2 = .220), with a greater magnitude of N400 peak at sites C3 and CZ, relative to sites C4 and PZ. Finally, there was a significant interaction between Age and Sentence Type, (F(1,28) = 8.572, p = .007, η2 = .234). This interaction of Age with Sentence Type is plotted in (Table 2), (Figure 3).

b) Intra-Individual N400 Peak Variability

The variability of single trial N400 responses to final words was investigated by assessing the standard deviation of the peak latency within each individual for both congruent and incongruent sentence types. The effect was assessed across central midline and parietal sites where the N400 is typically most prominent. A 2 x 2 x 4 mixed design ANOVA was conducted...
to examine the effects of Age (Young vs. Old), Sentence Type (Concurrent vs. Incongruent), and electrode Site (C3, CZ, C4, and PZ) on single-trial N400 Peak Variability (SD scores). There was a significant main effect of Age on the variability of the N400 on a trial-by-trial basis \( F(1,28) = 4.945, p = .034, \eta^2 = .15 \). Older adults had higher SD scores on average than younger adults (75.8 ms and 73.2 ms, respectively). There was also a main effect of Sentence Type on participants’ SD scores \( F(1,28) = 5.726, p = .024, \eta^2 = .17 \). Participants had more variable N400 components in response to congruent sentence-final words than incongruent sentence-final words (75.6 ms and 73.3 ms, respectively). There was not a significant main effect of electrode site \( F(3,28) = .241, p = .867, \eta^2 = .009 \).

There was a marginally significant interaction effect between Age and Sentence Type \( F(1,28) = 3.765, p = .062, \eta^2 = .12 \). This interaction was predicted by our hypothesis, so we examined this effect further. An analysis of simple effects revealed a significant effect of Age on SD scores for incongruent sentence-final words \( F(1,28) = 11.61, p = .002, r = .54 \), but no significant effect of Age for congruent sentence-final words \( F(1,28) = .173, p = .681, r = .08 \). Specifically, younger adults showed reduced variability \( (M = 71.1 \text{ms}, SE = .92) \) while processing the incongruent words compared with older adults \( (M = 75.6 \text{ms}, SE = .92) \). The interaction effect of Age and Sentence Type on SD scores is plotted in Figure 4.

![Figure 4: Effect of Age and Sentence Type on intra-individual processing variability. Higher bars indicate greater average variability of peak amplitude during the 300 – 500ms post-stimulus onset window. (Error bars: +/- 1 SE).](image)

As predicted, older adults exhibited significantly greater variability than younger adults overall; specifically, this difference is driven by young adults’ more consistent neural response timing in the semantically incongruent sentence condition. The filtering procedures used in the single-trial analysis rely on the assumption that a sufficient portion of the peaks pulled will reflect neural activity associated with semantic integration, as indexed by the N400 at the averaged-ERP level. Although almost certainly a proportion of these peaks will still reflect noise from background EEG, any reliable differences found between age groups ought to be indicative of true variability differences in underlying single-trial ERPs that have a veridical effect on the amplitude of the averaged ERP.

Although the effect found here is statistically reliable, there are two potential causes. The first is that, as we hypothesized, older adults’ N400s are more variable on a single-trial basis compared with younger adults. Alternatively, older adults may simply have very weak N400s on a single-trial basis. Our variability analysis may be victim to a noise ceiling effect where the peaks being pulled for older participants overwhelmingly represent noise, whereas the younger adults with stronger single-trial N400s allow us to resolve variability above the noise threshold.

To address this issue and to verify that older adults have reliably greater intra-individual N400 variability, we used a bootstrapping procedure to examine the variability of the contributing trial ERPs to the grand ERP in a linear manner, as described in the methods above. These Noise-to-Signal Transition Functions plot variability against increasingly larger ERP subset sizes with improving signal-to-noise ratios, permitting the comparison of the older and younger participant groups. If variability between the groups is equivalent, then it would have expected that the functions for the young and older groups would converge asset size increases and the older adults’ signal-to-noise ratio improves. Consistent with our hypothesis that older adults have greater intra-individual variability the two functions do not converge.

### Noise-to-Signal Transition Function Analysis

Noise-to-Signal Transition Functions (NSTFs) were created for both semantically incongruent and congruent types at each of the midline and parietal electrode sites (C3, CZ, C4, and PZ), and examined independently. Figure 5 presents the NSTFs for sites C3, CZ, C4, and PZ for the incongruent sentence types. Visually, at each site the two functions appear independent between the two groups, and do not intersect prior to complete averaging at ERP subset size 37. In each case, the older age group has higher SD scores across all ERP subset sizes. For each NSTF, lines of best fit were calculated using a least squares method. Two-tailed independent t-tests were conducted to explore differences in line intercepts between Age groups.

For the incongruent sentence types, there was a significant effect of Age on the intercept of the line of best fit for the NSTF at sites C3, CZ, t(28) = -2.24, p = .034, and C4 t(28) = -2.17, p = .039. In both cases, younger adults had lower intercepts (indicating less variability) than older adults (CZ: young M = 69.3, SE = 3.11, old M = 78.1, SE = 2.42; C4: young M = 68.8, SE = 2.49, old M = 76.9, SE = 2.75). There were not significant differences at sites C3 or
processing incongruent sentences, but rather a marked decrease in adults' increased variability relative to younger adults while the effect. Furthermore, our results indicate that the observed variation to our hypothesis that increased intra-individual variability of the incongruent stimuli would exhibit significantly more variability in the timing of the processing of this word may be more varied as word. Since the exogenous noise stemming from the unexpected context in a predictive manner. Conversely, older adults may be relying more heavily on context when younger adults encounter an unexpected semantically incongruent word, they may exhibit a broader categorial mismatch neural response. Therefore it may not matter precisely what the individual final word stimuli is, only that the final word is semantically incongruent with the word predicted by the preceding sentential context. Additionally, older adults with a more developed lexicon may treat the varying incongruent word stimuli with greater specificity, reflected in the greater trial-by-trial variability compared with younger adults. This interpretation would conform with behavioral findings that indicate older adults have more elaborate and well-established lexicons, leading to improved ability to access related vocabulary items like antonyms and synonyms. Potentially the preceding sentential context is activating a more widely distributed network of candidate predicted words. Unfortunately, we did not measure word recognition accuracy in the present investigation. It is recommended that in future investigations that verbal fluency measures and word recognition scores be assessed to directly test this explanation.

Another possible contribution to timing variability may arise from a reduced tendency for older adults to rely on the exogenous stimulus and a greater reliance on context due to perceptual deficits observed that older, relative to young adults, were more prone to incorrectly thinking they had heard a word that was not presented when it was presented in conjunction with an initial word from a pair they had been trained on. That is, they were more likely to think the presented word was the completion of the correct word pair previously learned when it was actually different but phonetically similar. This greater reliance of the older adults on the learned context rather than the exogenous stimulus was evident even after correcting for age-related differences in the signal to noise ratios necessary to identify neutral words.

Older adults might be relying more heavily on context in a strategically different manner than younger adults. Our observation that relative to older adults, young adults exhibited a more consistent N400 latency in response to incongruent sentence final words may be indicative of their greater use of context in a predictive manner. Conversely, older adults may be using context to disambiguate a noisier exogenous stimulus word. Since the exogenous noise stemming from the unexpected word could be expected to be more variable from trial to trial, the timing of the processing of this word may be more varied as a function of the strength of context given by each independent sentence stimulus. Ultimately, the exogenously presented incongruent word is processed correctly, as evidenced by the statistically identical and near-perfect performance accuracy for both incongruent and congruent stimulus types.

However, the use of context to disambiguate a noisy exogenous signal would place differential dependence on context.
for older adults. This, combined with sentences that vary in regards to cloze probabilities (although only moderately) may contribute to greater trial-by-trial variability of processing for older adults. This variability provides an alternative explanation for the lack of an N400 amplitude gradation related to semantic constraint variation observed in some older adults in previous studies Federmeier, Kutas [27] Federmeier, et al. [29].

In line with this explanation, Wlotko, Federmeier [64] propose that older adults may use context in a strategically different way than younger adults. They propose that younger adults may use sentential level context in a more predictive way, while older adults use context in a more integrative way. They based this hypothesis on evidence obtained from examination of the N400 in young and older adults obtained to sentence final words encountered while reading sentences varying systematically in cloze probability. Specifically, they observed that sentence level context effects were stronger in young adults-resulting in N400 amplitudes that were closely associated with the cloze probability of the sentence. Conversely, age-related N400 reduction appears to be due to increased item-by-item latency variability. Older adults show increased latency variability for unexpected sentence final words compared to younger adults. Compared to younger adults context effects were more variable in older adults and were subject to greater intra individual differences. Further, they suggest that intra individual variation in the use of context is likely to be stronger with high cloze probability sentences, such as those used in the current investigation, relative to lower cloze probability sentences.

The greater intra individual timing variability on an item by item basis observed in our older listeners, relative to the young listeners is also consistent with a growing body of literature that supports the increased role of cognitive functioning in the speech processing performance of older adults Arlinger, Lunner, Lyxell [59] Lunner [62]. For example in a recent investigation examining young and older listeners of varying hearing abilities, Benichov et al. [65] observed that the impact of hearing acuity on speech recognition performance decreased as sentence level contextual support increased. However, variability in the cognitive functioning capabilities of older adults contributed to speech processing performance at all levels of context. As a group, older adults often exhibit a greater range of cognitive abilities than their younger counterparts Cabeza, et al. [66].

Compounding this greater variability in cognitive abilities at the group level is the greater trial-by-trial variability in response times with increasing age that is commonly found in cognitive research Li, Aggen, Nesselroade [67]; Li et al. [68] Papenberg et al. [69]. These previous observations, together with the current results on the impact of trial-by-trial latency jitter, suggest that a variety of mechanisms likely contribute to the observed age-related changes in the N400 amplitude. Estimating the contribution of latency jitter will facilitate greater understanding of the relative role of other mechanisms Baddeley A [70] Cattell RB [71].

The NSTF methodology used in this investigation offers a unique approach to extracting variability information from ERP investigations. This procedure is applicable to other ERP investigations outside of the N400 and language domain in which differences in intra-individual variability may be a factor. Furthermore, the method can be applied to data sets previously collected to test for variability effects that may have been missed in the initial analyses Horn JL [72] Jasper HHV [73].

An improved understanding of how language comprehension changes with age is important for achieving a complete picture of cognitive aging. Individual differences in processing variability represent underutilized information present in ERP data ripe for exploration. An improved understanding of age-related changes in language processing will allow a proactive approach to be taken in developing technologies that reduce, eliminate, or avoid problems arising from age-related speech processing difficulties. Further, the details of comprehension changes may be used to inform the design of technologies to mitigate impairment McArdle JJ [74-79].

References


