Processing lexicality in healthy aging and Alzheimer's disease: P3 ERP amplitude as an index of lexical categorization

Nancy Azevedo[1,2], Ph.D, Ruth Ann Atchley, Ph.D.[3], N.P. Vasavan Nair, MD, FRCPC[4,5] and Eva Kehayia, Ph.D.[1,2]

[1] School of Physical & Occupational Therapy, McGill University

[2] Center for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR) - Jewish Rehabilitation Hospital of the CISSS de Laval

[3] Department of Psychology, University of South Florida

[4] Department of Psychiatry, McGill University

[5] Memory Clinic, Douglas Mental Health University Institute

Abstract

To explore how processing lexicality may change with aging and in the presence of Alzheimer's disease (AD), we conducted two experiments investigating lexicality judgements using an on-line behavioural psycholinguistic methodology and electrophysiological/event-related potential (ERP) methods; oddball lexical decision tasks. Results from these lexical decision tasks showed that while those with AD show similar rates of accuracy for their lexical decision as compared older adults (OA), they are particularly slowed when making judgements for pseudowords. Our results from the ERP tasks also showed that the two groups behaved differently with regard to elicitation of the P3 ERP response, which indicates differences in how these two groups form lexical categories. The pattern of ERP responses suggests that older adults are sensitive to the orthography/phonology of the stimuli during the course of lexical processing as compared to participants with AD who show less sensitivity to orthographic/phonological cues. Additionally, the ERP P3 amplitude results suggest further linguistically related differences between healthy older adults and those with AD, and highlight the importance and usefulness of combining behavioural psycholinguistic and ERP methodologies.

Keywords: lexical processing, lexicality, lexical decision, event-related potentials (ERP), oddball task, P3 component, aging, Alzheimer's disease

Visual word recognition, the ability to accurately and quickly determine whether a string of letters refers to a real word or not, is a core component of reading and comprehension for written language (Gold, Andersen, Jicha, & Smith, 2009). The recognition of one of the many thousands of written words that adults know is a complex, multi-step process (Eisenhauer, Fiebach, & Gagl, 2018; Yap & Balota, 2019) involving visual-perceptual, pre-lexical (orthographic and phonological), and lexical-semantic processing levels (e.g., Carreiras, Armstrong, Perea, & Frost, 2014; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). However, despite this complexity, this process is accomplished rapidly in young adults, with word-specific information beginning to be accessed in the first 200 milliseconds (ms) after presentation (Sereno & Rayner, 2003). One excellent way to study these component processes within word recognition is to ask readers to perform lexicality judgements using a stimulus list that includes words, pseudowords, and nonword stimuli. Pseudowords are legal and pronounceable strings of letters, such as *poble*, that are not part of the English lexicon. This kind of research allows us to look at differences based on the orthographic/phonological legality of the stimuli, which sheds light on the pre-lexical stage of word recognition. Through these experiments, we see that young adults are faster and more accurate in rejecting non-pronounceable letter strings than those that are pronounceable (Azevedo, Kehayia, Atchley, & Nair, 2015; Ratcliff, Thapar, Gomez, & McKoon, 2004). This difference suggests that pseudowords may trigger information (i.e. familiarity of letter combinations, phonological/orthographic legality, or the capacity to access word representations), that is either not activated or less activated by nonwords.

Visual Word Recognition in Healthy Aging

As people age, many types of cognitive processes begin to slow down. While tasks in some non-lexical cognitive domains tend to show more age-related slowing than lexical tasks (Hale & Myerson, 1996; Lima, Hale, & Myerson, 1991), even with many years of exposure to words over a lifetime, the time taken to make lexicality judgements also increases with age. As early as 1981, Bowles and Poon (1981) found that older adults had longer response times (RTs) when performing a lexical decision task requiring the discrimination between high and low frequency nouns and orthographically and phonologically legal pseudowords. More recently, additional studies have also reported findings showing that older adults are consistently slower to perform lexicality judgements when compared to young adults across several languages and across a variety of word types controlled for different lexical-semantic properties such as frequency, neighbourhood density and/or frequency, and syntactic class (Azevedo et al., 2015; Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Davies, Arnell, Birchenough, Grimmond, & Houlson, 2017; Kavé & Levy, 2005; Moberg, Ferraro, & Petros 2000; Ratcliff,et al., 2004; Robert & Mathey, 2007; Stadtlander, 1995; Taler & Jarema, 2007).

When seeking to estimate the rate of slowing of performance on lexical decision tasks caused by aging, Madden (1992) found that RTs increased with age at a rate of 4 ms per year for perceptually intact targets and 10 ms per year for degraded targets. Similarly, a meta-analysis of 22 lexical decision experiments found that older adults' RTs were 250 to 300 ms longer than those of young adults (Myerson, Ferraro, Hale, & Lima, 1992). However, while older adults are consistently slower when making lexicality judgements, this increase in response latency is not necessarily accompanied by a decline in accuracy rates. Some lexical decision studies report no difference in error rates between older and younger adult groups (Balota et al., 2004; Duñabeitia, Marín, Aviles, Perea, & Carreiras, 2009; Lima et al., 1991; Myerson et al., 1997; Robert & Mathey, 2007). However, Taler and Jarema (2007) reported a higher overall error rate for older adults while Stadtlander (1995), Ratcliff and colleagues (2004), and Azevedo and colleagues (2015) found that the older adults were more accurate when compared to young adults. Therefore, in summary, it is clear that healthy aging slows one or more of the component processes that contribute to visual word recognition, but in many of the studies, there is less evidence that aging causes word recognition to be less accurate.

Visual Word Recognition in Alzheimer's Disease (AD)

Language impairments are reliably observed in those with AD, even early in the course of the disease (Taler & Phillips, 2008). It has been suggested that the language impairment observed in those with AD is associated with degraded semantic memory, which includes impaired word-list generation and other anomia like deficits (see Braaten, Parsons, McCue, Sellers, & Burns, 2006 for review). Consistent with this theoretical argument, it has been found that the majority of individuals with AD exhibit word-finding difficulties from the onset of the disease (Appell, Kertesz & Fisman, 1982; Bayles & Kaszniak, 1987; Duong, Whitehead, Hanratty, & Chertkow, 2006; Joubert, Joncas, Barbeau, Joanette, & Ska, 2006; Kavé & Goral, 2018). When looking at tasks where individuals with AD and healthy older adults are required to make lexicality judgements, numerous studies have shown that those with AD are consistently slower when compared to older adults across different languages and across a range of words with different lexical-semantic properties such as frequency, neighbourhood density, and number of semantic associates (Caza & Moscovitch, 2005; Duñabeitia, Marín, & Carreiras, 2009; Duong et al., 2006; Madden, Welsh-Bohmer, & Tupler, 1999; Nikolaev, Higby, Hyun & Ashaie, 2019). However, one might ask if it is *only* the lexical-semantic stage of word recognition affected by AD, or is it the case that earlier stages in the word recognition process are also impacted by the cognitive changes associated with AD?

Past research has also suggested the presence of a selective deficit in the ability of individuals with AD to process pseudowords, as compared to words or nonwords (Cuetos et al., 2010; Glosser, Kohn, Friedman, Sands, & Grugan, 1997; Madden et al., 1999), which would suggest that AD might also impact pre-lexical processing of phonology and orthography. These studies reported an over-acceptance of pseudoword stimuli compared to what was observed in healthy aging. In a lexical decision study that we previously conducted (Azevedo, Kehayia, Atchley, & Nair, 2015), we did not observe a difference in error rate for pseudowords between healthy older adults and those with AD, but we did see a selective slowing for pseudowords in the AD participant group. Nonetheless, questions remain regarding individuals' with AD ability to process lexicality. Given that the behavioural methods used above can only provide the total time taken to visually process a lexical stimulus (including the time taken to make a response), the use of other, more sensitive, methods to study the time course of lexical decision may offer better insight regarding the process and results obtained. Event-related potentials (ERPs) offer the possibility of describing the process as it unfolds over time; thus, while from a behavioural point of view both healthy older adults and those with AD may appear to be processing lexicality in a similarly accurate manner, it remains possible that the steps taken to achieve an accurate response may differ.

Event-Related Potentials (ERPs) Investigating Visual Word Recognition

Electrophysiological measurements of event-related brain potentials (ERPs) have contributed complementary information regarding the cognitive processes that underlie visual lexical processing and word recognition that would not be available using traditional psycholinguistic methodologies alone. While the spatial resolution of electroencephalography (EEG) is poor, its temporal resolution is exceptional and permits the study of brain activation associated with ongoing cognitive processes in real time. One way to study visual word recognition is to ask participants to perform lexicality judgements on strings of letters presented within an oddball task. In a traditional oddball task, participants are asked to attend to a rarely occurring stimulus when it is presented among a different, and frequently occurring, type of stimulus (circles presented among squares, for example). In an oddball lexical decision task, the rare stimuli are words presented among frequently occurring nonwords or pseudowords (or vice-versa). While this methodology has been used with young adults (Azevedo, Atchley, & Kehayia, 2015; Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999), we are not aware of any studies that have employed it with healthy older adults or individuals with AD.

The traditional oddball paradigm described above is known to elicit the P3 component (also referred to as the P300) which represents the brain's electrophysiological response to a stimulus that is unexpected and can be elicited by "low-probability task-relevant stimuli during stimulus classification tasks in auditory, visual, and somatosensory modalities"

(Olichney, Yang, Taylor, & Kutas, 2011). The P3 has been documented as an index of attentional resources associated with a number of different cognitive processes, most importantly in the context of the current study, of stimulus-categorization (Ashford, Coburn, Rose, & Bayley, 2011; Azizian et al., 2006; Juckel, Karch, Kawohl, Kirsch, Jäger, et al., 2012; Mecklinger & Ullsperger, 1993; Polich, 2007). Since P3 amplitude is significantly larger when an individual encounters a stimulus that differs in a salient fashion from its antecedent context (called the P3 response), we maintain that a P3 response can be used as a metric of an individuals' ability to use lexicality as a salient feature, when performing the lexical decision oddball tasks. A reduction in P3 amplitude (i.e. no significant P3 response) can be interpreted as indicative of a significant change in the individual's ability to selectively attend to the task-critical stimulus type or effectively create categories that can be used to distinguish between the different stimulus types in an oddball list. Thus, if the P3 amplitude is diminished in a particular lexical decision oddball stimulus task (for example, if the task is to find rare pseudowords among a list of frequent words) then we would interpret this finding to indicate that creating the "word" and "not a word" categories in this situation is significantly more difficult. Further, based on the extensive psycholinguistic literature employing the lexical decision task, we would theorize that this disruption in lexical categorization is most likely due to the orthographic/phonological overlap between the rare and frequent categories, which render it difficult to quickly make the lexicality judgment. Because the P3 response is sensitive to such a wide range of stimulus differences, when the individual encounters a rare stimulus that differs in a salient fashion from its antecedent context, we can use the P3 response to examine the degree to which healthy aging adults and individuals with AD can make use of lexicality as a salient categorization feature in a lexical decision oddball task.

In a review of P3 research with individuals with AD, Polich and Corey-Bloom (2005) showed that despite the generalized neurocognitive decline, which is the hallmark of AD, a reliable P3 can be observed in individuals with AD. This is a critical observation, given that some other patient populations do not readily show a reliable P3 (for example, research by Bruder, Tenke, Stewart, Towey, Leite, et al.,1995 and Röschke, Wagner, Mann, Fell, Grözinger, et al., 1996 has demonstrated a significant attenuation of the P3 in patients with depression). Furthermore, the Korean version of the Consortium to Establish a Registry for Alzheimer's disease (CERAD-K) assessment, which includes various neuropsychological tasks, showed that changes in the P3 component were reflecting the deterioration in areas of cognition including of language, memory, and executive function in individuals with AD (Lee, Lee, Moon, Moon, Kim, et al., 2013). Although there have been a number of P3 studies conducted with individuals with AD (Ally, Jones, Cole, & Budson, 2006; Chapman, Nowlis, McCrary, Chapman, Sandoval et al., 2007; Katada, Sato, Ojika, & Ueda, 2004; Lee et al., 2013; Parra, Ascencio, Urquina, Manes, & Ibáñez, 2012; Pokryszko-Dragan, Słotwiński, & Podemski, 2003; Polich & Corey-Bloom, 2005), we are not aware of any research that has employed the P3 oddball task to specifically examine changes in the ability to process lexicality. In this way, though the P3 task has been established as a viable task to use with individuals with AD, the current research is innovative in that it uses this established task to study deficits in the processing of lexicality in this patient population.

The Current Study

In order to further explore how processing lexicality may change with the presence of AD and seeking to identify linguistically-related differences that could be used to differentiate healthy aging adults and those with AD, we conducted two experiments that investigate lexicality judgements in these two populations using on-line behavioural psycholinguistic methodology (Experiment 1) and event-related potential (ERP) methods (Experiment 2). Furthermore, in order to rule out any overt deficits in selective attention that may be affecting performance on the lexical decision oddball tasks, we also evaluated these cognitive abilities using two attention tasks (please see Appendix). We anticipate that this combination of methodologies can enhance our understanding of lexicality processing in individuals with AD and provide insight on how these individuals may differ from healthy older adults.

Experiment 1 (Lexical Decision Task)

Methods

Participants were 18 individuals with probable AD (11F, 7M) and 24 healthy older adults (17 F, 7 M). All were dominant English speakers and had normal or corrected-to-normal vision and were primarily recruited via the Memory Clinic of the Douglas Mental Health University Institute in Montréal, Québec. The older adult (OA) group had a mean age of 68 years (range: 50-88) and, on average, had 14.5 years of education (range: 10-20). All scored 26/30 or above on the Montreal Cognitive Assessment (MoCA, Nasreddine, Phillips, Bédirian, Charbonneau & Whitehead, 2005) thus ruling out the presence of cognitive impairments in this group (Luis, Keegan & Mullan, 2009; Nasreddine et al., 2005).

The probable AD group had a mean age of 75 years (range: 57-83) and averaged 15 years of education (range: 7-21 years). The criteria specified by the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA; McKhann, Drachman, Folstein, Katzman, Price et al., 1984; McKhann, Knopman, Chertkow, Hyman, Jack Jr. et al., 2011) were used for the diagnosis of probable AD that was performed by a psychiatrist with extensive experience with this population. The Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975; Folstein, Robins, & Helzer, 1983) was used to evaluate AD severity. Only individuals with AD whose score on the MMSE ranged from 20 to 29, indicating mild severity (Feldman & Woodward, 2005; Feldman, Van Baelen, Kavanagh, & Torfs, 2005), were included in this study. All participants were carefully assessed by study staff to ensure that they could comply with the requirements of cognitive testing. The study was approved by the Douglas Research Centre Research Ethics Board and written informed consent was obtained from all participants.

The on-line visual lexical decision task was performed in a single session using E-Prime experiment software (Psychological Software Tools, Pittsburgh, PA). The task consisted of 320 trials in total: 80 experimental words (concrete monomorphemic nouns); 80 filler words (verbs); 80 pseudowords; and 80 nonwords. All stimuli were matched for length (4-6 letters and 1-2 syllables) and the words had a frequency of use of 11 occurrences per million (range: 2 and 50) according to the USENET database (Shaoul & Westbury, 2006). Pseudowords were strings of letters that are permissible in English spelling and are pronounceable (e.g. *fodum*) and nonwords were non-permissible and unpronounceable letter strings (e.g. *mlopf*). Both pseudowords and nonwords were included in this study in order to investigate the difference in processing between the phonologically/orthographically legal and illegal stimulus types and to allow for a better comparison with the stimuli in Experiment 2. Each trial began with the presentation of a fixation cross for 500 ms and then a stimulus was presented in lowercase letters. The stimulus remained on the screen until the participant made a response by pressing one of two buttons on a keyboard. Once a response was made, there were 500 ms before the next stimulus was presented. Stimuli were randomized for each participant and participants were given a break midway through the experiment. Participants were instructed to decide as quickly and as accurately as possible whether the stimulus was a word or not. We conducted 10 practice trials before beginning the experiment to ensure that participants were comfortable with the task. Response time (RT, in ms, measured from stimulus onset until a response was made) and error rate were collected for all trials.

Statistical Analysis

Participants who made 50% or more errors for any of the stimulus types were excluded from all analyses. This cut-off led to the exclusion of one (1) individual with AD and resulted in a final total of 41 participants. For the older adults, trials with an incorrect response or with a lexicality judgement response time over 2 standard deviations from the group mean (for each stimulus type) were removed from the RT analysis and led to an average of 3% of trials being excluded. For the AD participants, trials with an incorrect response or a response time over 2 standard deviations from their individual mean (for each stimulus type) were not included in the RT analysis and led to the exclusion of an average of 4% of trials. Data for the filler trials (the verbs) were not included in the analyses for either of the two participant groups. Mean correct lexical decision response times and percent error for each stimulus type are presented in Table 1.

Table 1. Error Rates and Response Times (RTs) in ms for lexicality judgements for Older Adults and Individuals with AD (Exp. 1)

Stimulus Types	Older Adults				Individuals with AD			
	Error Rate		RT		Error Rate		RT	
	median	MAD	mean	SD	median	MAD	mean	SD
Words	1%	1%	710	69	1%	1%	963	336
Pseudowords	4%	3%	917	150	8%	5%	1381	483
Nonwords	0%	0%	669	57	1%	1%	998	414

Given that the error rates did not follow a normal pattern of distribution we chose to proceed with analysis of this variable using the Wilcoxon-Mann-Whitney nonparametric test. Differences in response latencies between the participant groups on the lexical status of the stimuli were investigated using a 2 (Group: older adults (OA) and individuals with AD (AD)) x 3 (Stimulus type: words, pseudowords and, nonwords) mixed-model analysis of variance (ANOVA).

Results

Error Rates.

When looking at the overall error rates, older adults and individuals with AD both had a low error rate (3%, range: 0-11% and 4%, range: 0-12% respectively). Results from the Wilcoxon-Mann-Whitney test show that the two groups displayed a similar pattern of performance across the 3 stimulus types. There was no difference in error rates between groups for words: Z= -0.01, $p = 0.98$; pseudowords: Z= 1.05, $p = 0.29$; or nonwords: Z= 1.41, $p = 0.16$.

However, when comparing within groups, each group made significantly more errors to pseudowords than to words (both groups: *p's* < 0.0001) and to pseudowords than to nonwords (both groups: *p's* < 0.0001). While those with AD did not show a significant difference in errors between nonwords and words ($p = 0.98$), the older adults made fewer errors to nonwords than to words ($p = 0.03$).

Response Time (RT).

Results from the 2 (Group: OA, AD) x 3 (Stimulus type: words, pseudowords, and nonwords) ANOVA showed that individuals with AD were significantly slower ($F(1, 39) =17.73$; $p = 0.0001$) than the healthy older adults. There was also a significant main effect of Stimulus Type ($F(2,78) = 103.50$; $p < 0.0001$)). We found that pseudowords were significantly slower than the other two conditions (1109 ms), while the two groups of participants responded to words and nonwords at a similar rate (words = 815 ms, nonwords = 805ms). These two main effects were further modulated by a significant Group x Stimulus type interaction ($F(2,78) = 8.95$; $p < 0.001$). Further analysis of this interaction showed that those with AD were significantly slower compared to older adults for each of the stimulus types (p's < .01). However the degree of difference between the two participant groups is far more pronounced for the pseudowords (AD were 464 ms slower than OA) than for words (difference = 253 ms) or for nonwords (difference = 329 ms). One could argue that this graded separation between the groups is due simply to task difficulty. However, neither group showed a significant difference in RT when comparing nonwords and words (AD 35 ms, $t(78) = 0.90$, $p = 0.37$; OA 41 ms, $t(78) = 1.24$, $p = 0.22$).), which argues against the variability in RT between groups being only a function of task difficulty.

Therefore, our results, as was also observed in our earlier research (Azevedo, et al., 2015), provide some evidence that the AD participants are struggling more with pseudowords, relative to the healthy older adult participants. However, this finding is not as conclusive as we may have anticipated because it remains possible that our finding of differential slowing could be explained at least in part by simply pointing to selective slowing based on task difficulty. One is thus left wondering whether there may be an underlying linguistic deficit also contributing to our findings. To complement these data and obtain a better understanding of the processes that underlie lexicality judgements, we turn to ERP tasks in order to explore what is occurring during an earlier stage of lexical processing that occurs prior to the participant making a behavioural response.

Experiment 2 (Event-Related Potential (ERPs) - Oddball Lexical Decision Tasks)

The ERP oddball lexical decision tasks in Experiment 2 require the coordination of different cognitive abilities. These include linguistic abilities (to make the lexicality judgements) as well as abilities in other areas of cognition such as selective attention. In order to address this, we had participants perform a brief attention screening that comprised two short tasks measuring selective attention (the number Stroop and the Useful Field of View) prior to performing the ERP tasks. These provided a general evaluation of these non-linguistically related abilities and helped rule out an overt impairment at this level that may be interacting with individuals' performance on the oddball lexical decision tasks. The methods and results for these additional tasks, which show that the older adults and those with AD are not showing fundamental differences in basic attention tasks, are provided in the Appendix.

Methods

Participants were 10 individuals with probable AD (6F, 4M) and 17 healthy older adults (11 F, 6 M). The older adult group had a mean age of 69 years (range: 55- 85) and, on average, had 14.1 years of education (range: 10- 18). The probable AD group had a mean age of 76 years (range: 70- 83) and averaged 17.6 years of education (range: 12- 21 years). MMSE scores ranged from 20 to 29, indicating mild severity (Feldman & Woodward, 2005; Feldman et al., 2005). Recruitment location, inclusion criteria, and diagnosis criteria, were identical to those described for Experiment 1. All individuals who participated in the second experiment also first participated in Experiment 1. Testing for each experiment occurred in separate sessions, with approximately two weeks between sessions. Experiment 1 was always performed before Experiment 2.

Four blocks of the 20/80 oddball lexical decision task were administered in one single session using E-Prime experiment software (Psychological Software Tools, Pittsburgh, PA). Presentation was counterbalanced for each participant in order to control for possible order effects. Each block had 200 trials: 40 (20%) rare trials and 160 (80%) frequent trials and participants were asked to respond "YES" if the stimulus they saw in the trial was a word and "NO" if it was not. Eight hundred unique stimuli were used, 200 nonwords, 200 pseudowords, and 400 real words (see Azevedo, Atchley, & Kehayia, 2015 for a full list of stimuli). All stimuli were matched for number of letters (3 to 7 letters and 1 or 2 syllables), pseudowords and words were matched for neighbourhood density (0 to 21 lexical neighbours), and the words were concrete nouns with a frequency of occurrence in English (mean of 8 per million) according to the USENET database (Shaoul & Westbury, 2006). The stimulus types were manipulated based on their orthographic/phonological structure and their lexical status: words were both orthographically/phonologically legal and had a real lexical status (e.g. *cigar*); pseudowords were orthographically/phonologically legal but did not have a lexical status (e.g. *acle*); and nonwords were orthographically/phonologically illegal and did not have a lexical status (e.g. *nduy*). Block 1 consisted of rare (20%) nonwords among frequent (80%) words (Nw-W), Block 2 consisted of rare (20%) pseudowords with frequent (80%) words (Ps-W), Block 3 consisted of rare (20%) words among frequent (80%) nonwords (W-Nw), and Block 4 consisted of rare (20%) words with frequent (80%) pseudowords (W-Ps). There were no statistically significant differences in mean frequency of occurrence for word stimuli among the 4 blocks as determined by a one-way ANOVA ($F = 0.20$, $p = 0.9$); there were also no statistically significant differences in mean number of letters between any of the stimulus types for the 4 blocks as determined by a one-way ANOVA ($F = 1.44$, $p = 0.19$); lastly there were no statistically significant differences in mean number of lexical neighbours (N) across the phonologically legal stimulus types (word and pseudowords) in the 4 blocks as determined by one-way ANOVA ($F = 0.36$, $p = 0.88$).

A Latin-square design was used to balance block presentation. Each trial began with the presentation of a stimulus for 750 ms, followed by a blank screen for 250 ms. Next, a response prompt, indicating that participants could make their behavioural response, was presented for 1800 ms. We conducted 20 practice trials before beginning each experimental block to ensure that participants were comfortable with the task. Behavioural data (accuracy) were collected in each block of the oddball task using E-Prime experiment software. Response time was not collected since the task required that individuals delay making a behavioural response until they saw the response prompt. For the ERP analyses, the 40 rare trials and 40 frequent trials were used for analysis. The 40 frequent trials chosen for analysis were randomly selected and represented trials throughout the block from beginning to end. They were tagged for analysis during the programming of the experiments. Thus, the 40 rare and frequent trials for each block are the same for every person who participated in the experiment.

Electroencephalogram (EEG) data were collected for each trial using the Neuroscan data collection software and a 40-channel NuAmps amplifier, using silver-silver chloride electrodes. A QuickCap electrode cap with 34 monopolar electrodes was placed according to the 10/20 system. Each scalp site was referred to linked mastoids. Electrodes were placed above and below the left eye and at the outer canthi to monitor blinks and eye movements (electro-oculogram; EOG). Electrode impedances were measured using a criterion of 5 kΩ, per manufacturer guidelines. The EEG and EOG data were digitized on-line at a sampling rate of 250 Hz, and were filtered with bandpass cuttoffs of 0.1 - 30 Hz. EEG waveforms were time-locked to each stimulus onset and were segmented from 200 ms prior to stimulus onset to 1000 ms after stimulus onset. Eye-movement artifacts due to blinks were corrected off-line (Gratton, Coles & Donchin, 1983). A trial was identified as bad if it had movement artifacts of greater than ±70 μV and was rejected prior to averaging. In order to isolate the P3 component associated with the successful categorization of a stimulus as being either a "word" or "not a word", a traditional windowed analysis was conducted on individual average files. The a priori time window of 500-650 ms was chosen and the Cz, Fz, and Pz electrode channels were selected for analysis. The selection of this window was based on previous P3 work utilizing linguistic stimuli (Azevedo, Atchley, & Kehayia, 2015).

Statistical Analysis

To examine whether oddball lexical decision tasks evoked a P3 response to the rare trials (i.e. a significantly larger P3 component to the rare stimulus type when compared to the frequent stimulus type) for each participant group in any of the four blocks, we compared the mean amplitude of the rare trials to frequent trials in the critical time window of 500-650 ms. Furthermore, to investigate whether there was a difference in P3 amplitude between the two participant groups we used a 2 (Group: OA, AD) X 2 (Trial type: rare, frequent) X 4 (Block: Nw-W, Ps-W, W-Nw, W-Ps) X 3 (Electrode channel: Fz, Cz, Pz) mixed-model ANOVA. Trials with a correct response or a non-response were included in the analyses while trials that were responded to incorrectly were not included. For the AD participants, this led to the exclusion of an average of 2% of trials for the Nw-W block, of an average of 6% for the Ps-W block, of an average of 3% for the W-Nw block, and of an average of 5% for the W-Ps block.

Results

ERP, Behavioural Response Accuracy.

Error rates for each block of oddball lexical decision tasks for the 80 experimental trials are presented in Tables 2 and 3. Overall, older adults were highly accurate in each of the four blocks while individuals while AD had a higher error rate for each block. It is important to note that the delayed response required for the ERP tasks seems to have worked against those with AD; the bulk of their behavioural errors were actually non-responses, i.e. they made their response prior to the response prompt and were thus not recorded by E-Prime. We observed a high rate of non-response errors for all four experimental blocks. For the 80 experimental trials, we observe that 90% of errors in the Nw-W block, 64% of errors in the Ps-W block, 87% of errors in the W-Nw block, and 69% of errors in the W-Ps block made by the AD group were due to non-responses.

Table 2. *Mean Error Rates for ERP Lexical Decision Oddball tasks for Older Adults (Exp. 2)*

Block	Rare Trials (n = 40)		Frequent Trials (n = 40)		Total (n = 80)	
	mean	SD	mean	SD	mean	SD
Nw-W	2%	6%	3%	6%	2%	6%
Ps-W	3%	5%	5%	4%	4%	4%
W-Nw	2%	3%	3%	9%	3%	7%

W-Ps	4%	3%	7%	9%	6%	7%

Table 3. *Mean Error Rates for ERP Lexical Decision Oddball tasks for Individuals with AD (Exp. 2)*

Block	Rare Trials (n = 40)		Frequent Trials (n = 40)		Total (n = 80)	
	mean	SD	mean	SD	Mean	SD
Nw-W	11%	10%	25%	17%	18%	16%
Ps-W	13%	11%	22%	16%	17%	14%
W-Nw	20%	21%	26%	22%	23%	21%
W-Ps	13%	13%	15%	14%	14%	13%

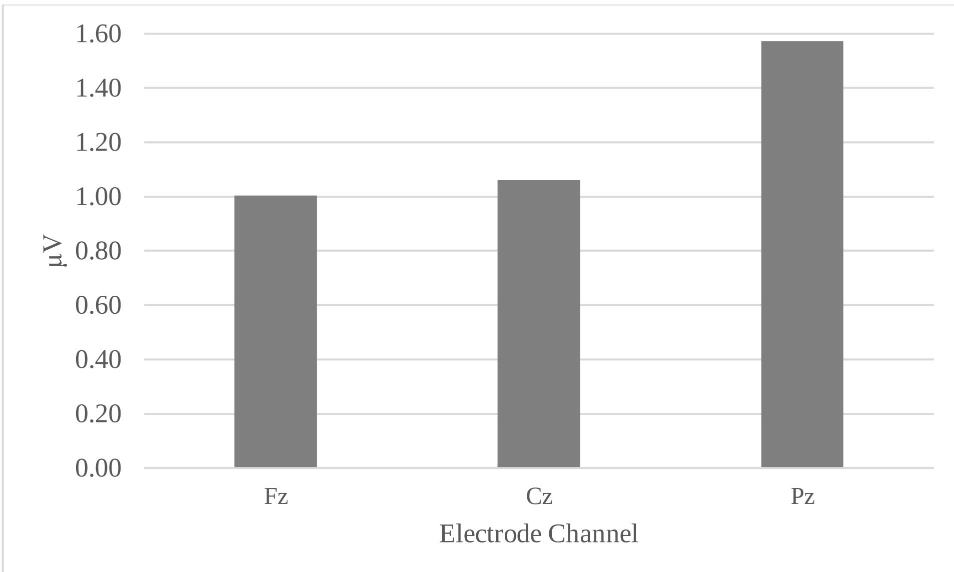
ERP, P3 Component.

Results from the mixed model 2 (Group: OA, AD) X 2 (Trial type: rare, frequent) X 4 (Block: Nw-W, Ps-W, W-Nw, W-Ps) X 3 (Electrode Channel: Fz, Cz, Pz) omnibus ANOVA showed a significant main effect of Block ($F(3,75) = 2.88$, $p = 0.04$), of Trial Type ($F(1,25) = 11.47$, $p = 0.002$), and of Electrode channel ($F(2,50) = 4.67$, $p = 0.01$). There was no main effect for the variable of participant group ($F < 1$). There was one statistically significant interaction between Block and Group ($F(3,75) = 3.01$, $p = 0.04$), a marginally significant interaction between Block and Trial Type ($F(3,75) = 2.45$, $p = 0.07$), and a marginally significant Block by Trial Type by Channel by Group interaction ($F(6,150) = 2.10$, $p = 0.06$). No other effects were significant.

The three main effects observed were consistent with our predictions and confirm that our ERP experiment produced P3 components that are comparable to earlier research in this area. First, the reliable Trial Type main effect indicates that we are seeing the expected P3 component given that the rare trials are more positive and larger in amplitude (Mean μV = 2.21) than the frequent trials (Mean μV = 0.99). As shown in Figure 1, our P3 component is numerically largest at the Pz electrode (Mean μV = 1.57) and grows smaller as we look at the component at the midpoint on the head (Cz, Mean μV = 1.06) and over a more frontal, midline electrode location (Fz, Mean μV = 1.00). Because we observed that our P3 component is consistent in its scalp distribution with the previous literature, stating that P3 amplitude is largest at the Pz electrode channel (Bentin et al., 1999; Parra et al., 2012), all additional analyses and planned comparisons were executed using data from Pz. Finally, we observed that the Nw-W condition was reliably more positive-going than the other three block conditions. This block main effect is best understood in the context of the additional analyses conducted for the PZ channel only and the planned comparisons discussed further below.

In order to more carefully consider the ERP findings observed at the Pz electrode, we next ran a mixed model 2 (Group: OA, AD) X 2 (Trial type: rare, frequent) X 4 (Block: Nw-W, Ps-W, W-Nw, W-Ps) ANOVA utilizing only the data from Pz. In this analysis, we again observed a significant main effect of Block ($F(3,75) = 3.07$, $p = 0.03$) and Trial Type ($F(1,25) = 17.50$, $p < 0.001$). We also observed a significant interaction between Block and Group ($F(3,75) = 2.99$, $p = 0.04$), which indicates that our OA and AD participants were processing the combinations of word, nonword, and pseudoword trial types in different ways. Like with the omnibus ANOVA, the highest order interaction (Group X Block X Type) resulted in an F value greater than 2 ($F(3,75) = 2.17$, $p = 0.099$), it did not reach the level for concluding confidently that we have a reliable higher order interaction.

Figure 1. *P3 component amplitude at the three electrode channels of interest (Exp 2)*



The ERP waveforms for both participant groups for each block are presented in Figures 2-9. Although the overall interactions were only marginally significant ($p = 0.06$ and p = 0.099), given that our a priori goal was to examine whether any block of the oddball lexical decision task evoked a P3 component for either participant group, we compared the rare trials to the frequent trials in the critical 500-650 ms time window at the Pz electrode for each block using simple planned comparisons implemented separately for each of the two participant groups.

Figure 2. Grand-average ERPs time-locked to stimulus onset at the PZ electrode for the OA group for block 1, Nw-W (Exp. 2).

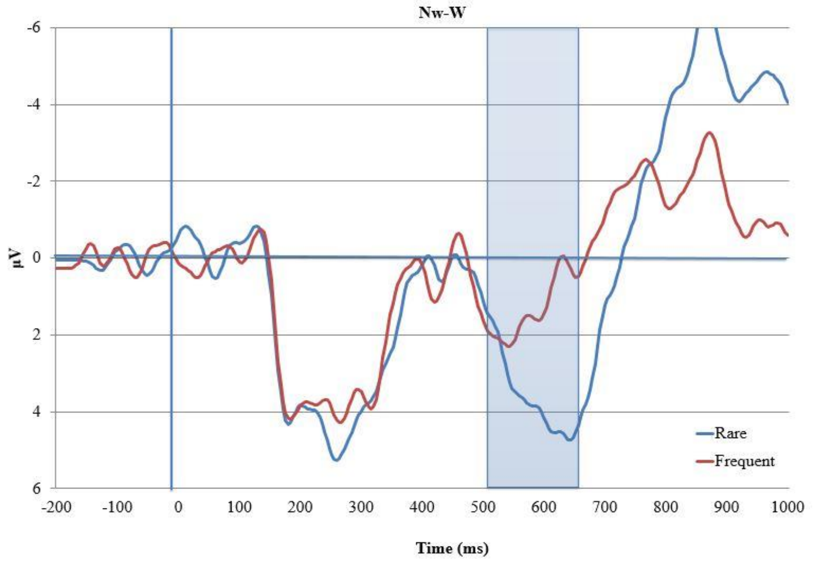


Figure 3. Grand-average ERPs time-locked to stimulus onset at the PZ electrode for the AD group for block 1, Nw-W (Exp. 2).

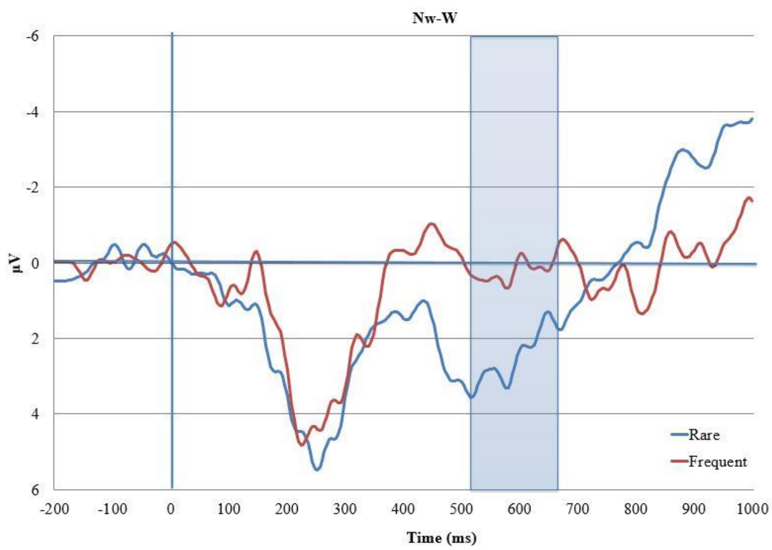


Figure 4. Grand-average ERPs time-locked to stimulus onset at the PZ electrode for the OA group for block 2, Ps-W (Exp. 2).

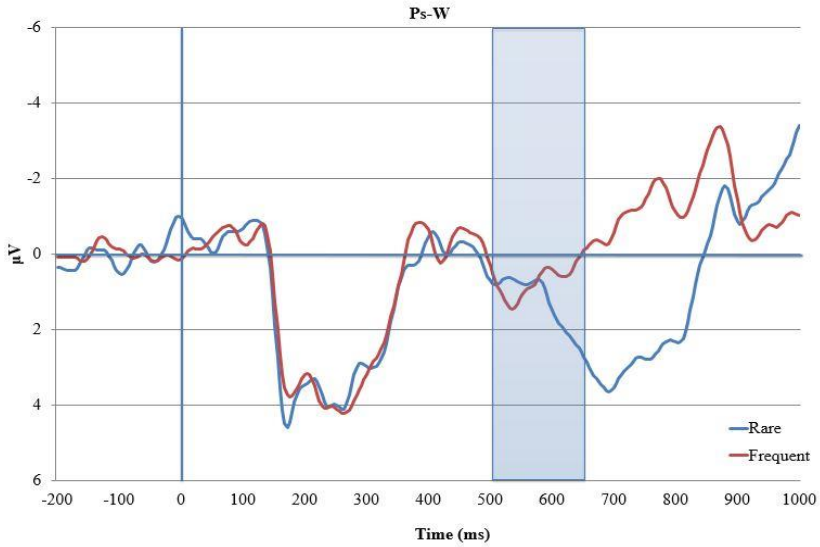


Figure 5. Grand-average ERPs time-locked to stimulus onset at the PZ electrode for the AD group for block 2, Ps-W (Exp. 2).

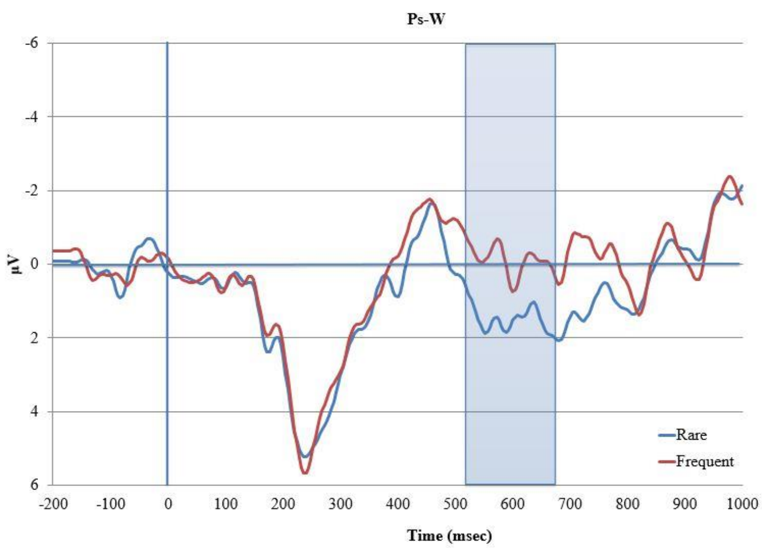


Figure 6. Grand-average ERPs time-locked to stimulus onset at the PZ electrode for the OA group for block 3, W-Nw (Exp. 2).

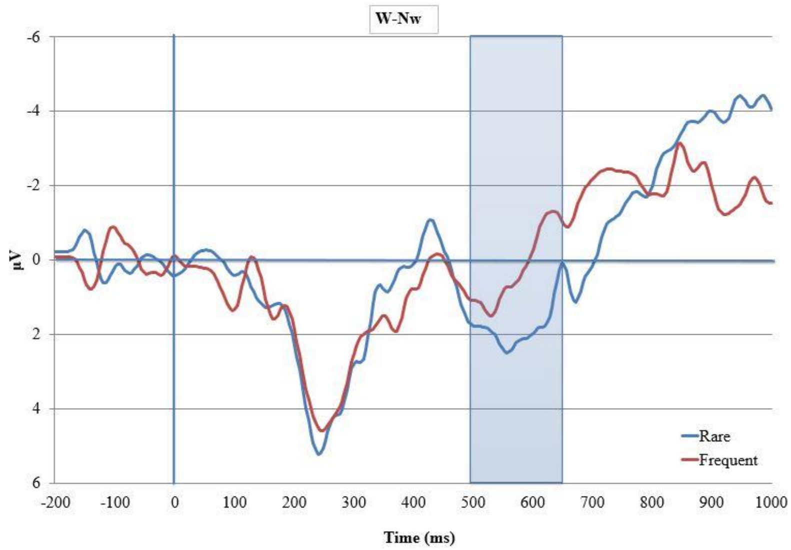


Figure 7. Grand-average ERPs time-locked to stimulus onset at the PZ electrode for the AD group for block 3, W-Nw (Exp. 2).

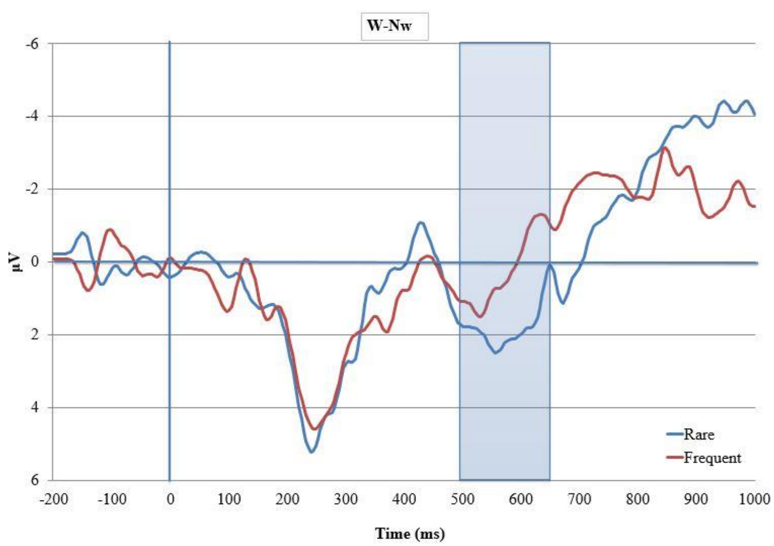


Figure 8. Grand-average ERPs time-locked to stimulus onset at the PZ electrode for the OA group for block 3, W-Ps (Exp. 2).

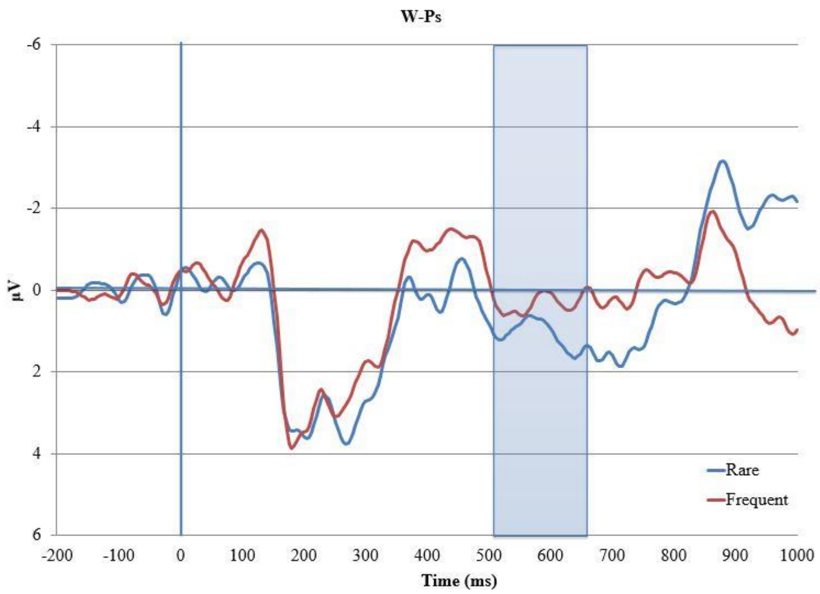
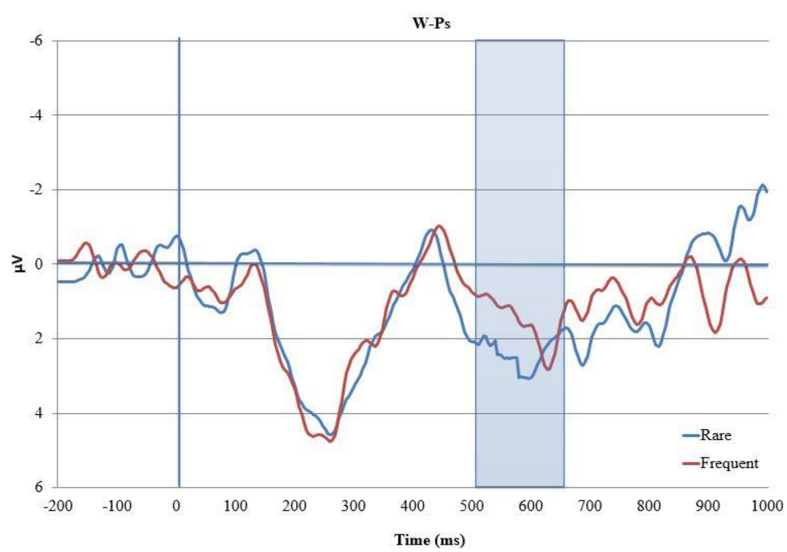


Figure 9. Grand-average ERPs time-locked to stimulus onset at the PZ electrode for the AD group for block 3, W-Ps (Exp. 2).



Given that our aim was to investigate which blocks, if any, elicited a P3 response, and then to compare the pattern of response between the two groups, we first present the results for the older adults and then those for the individuals with AD.

Beginning with the older adults, results of the planned comparisons showed that, for the two blocks that contrast words with nonwords (Nw-W and W-Nw), older adults displayed a significant P3 response. The P3 amplitude was significantly larger for rare nonwords (Mean μV = 3.52) compared to the frequent words (Mean μV = 1.37) in the Nw-W block ($F(1, 16) = 9.28$; $p = 0.008$) and for rare words (Mean μV = 2.64) compared to the frequent nonwords (Mean μV = -0.01) in the W-Nw block ($F(1, 16) = 8.622$, $p = 0.01$) (see Figures 2 and 6).

However, for blocks contrasting words with pseudowords (Ps-W and W-Ps) older adults did not display a significant P3 response. In the Ps-W block there was a non-significant difference in P3 component for the rare pseudowords (Mean μV = 1.31) compared to the frequent words (Mean μV = 0.72), $F(1, 16) = 0.54$, $p = 0.47$) and in the W-Ps block there was also a non-significant difference in P3 component for the rare words (Mean μV = 1.06) compared to the frequent pseudowords (Mean μV = 0.30), $F(1, 16) = 1.49$, $p = 0.24$) (see Figures 4 and 8).

In contrast, individuals with AD showed a different pattern in their ERP responses. Results of the planned comparisons showed that, for the two blocks in which words make up the frequent trials (Nw-W and Ps-W), those with AD displayed a significant P3 response. The P3 amplitude was significantly larger for rare nonwords (Mean μV = 2.61) compared to the frequent words (Mean μV = 0.25) in the Nw-W block ($F(1, 9) = 5.86$, $p = 0.04$) and for rare pseudowords (Mean μV = 1.29) compared to the frequent words (Mean μV = -0.26) in the Ps-W block ($F(1, 9) = 8.52$, $p = 0.02$) (see Figures 3 and 5).

Conversely, when nonwords or pseudowords comprised the frequent trials (W-Nw and W-Ps) with words being the rare trials, individuals with AD did not display a significant P3 response. In the W-Nw block there was also a non-significant difference in the P3 component for the rare words (Mean μV = 1.61) compared to the frequent nonwords (Mean μV = 0.27), $F(1, 9) = 3.53$, $p = 0.10$) and for the W-Ps block there was a non-significant difference in P3 component amplitude for the rare words (Mean μV = 2.48) compared to the frequent pseudowords (Mean μV = 1.30), $F(1, 9) = 2.19$, $p = 0.17$) (see Figures 7 and 9).

Discussion

With the onset of AD, changes occur in the way people process lexicality that are mainly manifested in a slowdown when performing lexicality judgements. However, while there is strong evidence that the semantic level of lexical processing is affected in AD, behavioural studies have suggested that the orthographic/phonological stage of this process may also be affected. In order to further explore this, we conducted two experiments investigating lexicality judgements in healthy older adults and in individuals with mild AD. For the behavioural lexical decision task in Experiment 1, consistent with results in the current literature, we found that individuals with AD were slower in making lexicality judgements than older adults. Furthermore, we observe that the three types of stimuli were not impacted equally but rather a gradation was seen in the amount of slowing with the pseudowords being the most affected. While the behavioural results from Experiment 1 do point to a potential change in the orthographic/phonological stage of lexical processing with AD, the ERP P3 amplitude results from Experiment 2 highlighted a possible change in this stage of lexical processing in the AD group. This finding suggests the presence of changes happening at a linguistic level that point to a difference between the healthy older adults and those with AD and highlights the importance and usefulness of combining behavioural psycholinguistic and ERP methodologies.

The presence of a robust P3 response in the current study is believed to reflect an initial stage of lexical categorization, i.e. the participants successfully made use of lexicality as a salient stimulus feature to form categories that allowed them to quickly differentiate between the stimulus types ("word" or "not a word") during the ERP oddball lexical decision task. While the ERP results, and any conclusions that we draw from them, should be interpreted with caution given the small sample size of the AD group, we found that while both groups successfully made use of lexicality as a salient stimulus feature (as reflected by a robust P3 response in certain experimental blocks) the conditions that elicited this response were different for the two groups thus suggesting that a change in how lexicality is processed is occurring in those with AD.

The pattern of observed ERP responses suggests that older adults are sensitive to the orthography/phonology of the stimuli early in the course of processing while performing ERP oddball lexical decision tasks. In the context where the rare and frequent stimuli are maximally distinct (W-Nw and Nw-W blocks), with stimulus types that are dissimilar from one another in both orthographic/phonological legality and lexical status, older adults were able to quickly create effective “word” and “not a word” categories (as reflected by a significant P3 response to the rare stimuli trials in these blocks). However, older adults no longer showed a P3 response in the critical 500-650 ms time window in the less saliently different blocks that had an overlap in orthographic/phonological legality between the rare and frequent stimulus types (W-Ps and Ps-W blocks). This suggests that sensitivity to orthographic/phonological legality interfered with the older adults’ ability to create effective “word” and/or “not a word” categories that could distinguish between the two orthographically/phonologically legal stimulus types quickly enough to be used by selective attention processes in these blocks. One could argue that the older adults are showing the expected early reliance on orthographic/phonological information before they go on to search for additional information in the semantic lexicon (as needed) and suggest that processing orthographic and phonological legality is more automatic than performing a lexical semantic search. Therefore, with overall slower RTs given their age and given that P3 is a relatively early waveform, our results suggest that the healthy older adults relied mainly on bottom up processing that utilises orthographic and phonological information to make these lexical decisions.

While the modulation in P3 response seen for older adults in the current study is very similar to what has previously been observed in younger adults (Azevedo, Atchley, & Kehayia, 2015), it contrasts with the ERP results of individuals with AD. Our AD participants showed a P3 response only in the context where words were the frequently presented stimuli (Ps-W and Nw-W blocks). It is important to note that a P3 response was elicited to both the pseudoword and nonword rare trials in these blocks despite the two stimulus types differing in their orthographic/phonological legality. The presence of a P3 response to both of these stimulus types suggests that those with AD are less sensitive to the orthographic/phonological legality of the rare stimuli when they are presented among an abundance of words; instead, they seem to have relied on whether or not the stimulus had a lexical status (and semantic content). Thus, both the orthographically/phonologically legal pseudowords and the orthographically/phonologically illegal nonwords were effectively characterized as being the same (i.e. “not a word”) and hence both stimulus types elicited a P3 response.

The modulation of the P3 response in contrasting contexts is consistent with interactive activation and competition (IAC) models of word recognition, such as the multiple read-out model (MROM; Grainger & Jacobs, 1996) and the dual-route cascaded model (DRC; Coltheart et al., 2001), that posit the existence of two criteria that can be used to make a lexicality judgement. A “Yes” response can be made when a specific word level unit reaches an activation threshold (i.e. a lexical activation) or based on a “fast guess” criterion that is sensitive to the level of global activation at the word level (i.e. if there is a high overall level of activation at the word level). Alternatively, a “No” response is given if, after a pre-specified time, no word representation is sufficiently activated to meet a threshold.

We propose that the degree of dependence on each of these two criteria differs between the two groups based on their sensitivity to and dependence on phonology/orthography in the initial steps of lexical processing. Older adults appear to favour the criterion of global activation but task expectancy (whether words were the frequent or rare trials in a given block) does not appear to influence their choice of criterion to use for each block. In the blocks where there is overlap in orthographic/phonological legality between the two stimuli types being contrasted (Ps-W and W-Ps blocks), the legality of both the words and pseudowords leads to an increase in global lexical activation for both stimulus types which appears to make the creation of effective “word” and “not a word” categories difficult in the time required to be available for use by selective attention; hence we do not observe a P3 response in these conditions. However, since there is little or no global activation being generated by the nonwords that can impede the creation of effective “word” and “not a word” categories, a robust P3 response is elicited for the blocks where there is no orthographic/phonological overlap between the stimulus types (Nw-W and W-Nw blocks).

One the other hand, those with AD seem to be insensitive or less sensitive to the global activation caused by orthographic/phonological legality of stimuli in the initial stages of lexical processing which leads them to instead rely on the lexical activation criterion. We observe a P3 response only in the blocks where words are the frequent stimulus type. It is worth noting that a P3 response was elicited to both the pseudoword and nonword rare trials in these blocks despite the fact that the two stimulus types differ in their orthographic/phonological legality, thus pointing to a change in sensitivity to orthography/phonology. Furthermore, we do not observe the P3 response in the W-Nw block as we did with the OAs, also pointing to a reduction in sensitivity to orthography/phonology in the AD group. If individuals with AD were fully sensitive to global activation generated by the orthography/phonology of the pseudowords then we would expect to observe a P3 response in both the Nw-W and W-Nw blocks but not in the

Ps-W or W-Ps blocks, as was observed in OAs. We interpret this finding to suggest that individuals with AD are more likely than the OAs to rely on the criteria of lexical activation, i.e. they are performing a lexical search for all stimuli types, even nonwords that are illegal and do not warrant a search. However, success with this strategy appears to be mediated by task expectancy, i.e., whether the frequent trials are words or not. When the frequent trials are words, adopting such an approach seems to be successful as evidenced by a P3 response to the rare trials, but when the frequent trials are not words this strategy appears to no longer be successful as they do not seem to be able to retrieve a lexical entry for words in time to elicit a P3 response. We thus hypothesize that it is the lexical activation criterion that drives the AD performance. In addition, this apparent decrease in sensitivity to orthography/phonology, and the consequent reliance on the criterion of lexical activation, appears to be contributing, above and beyond any effects of general slowing in processing speed that is known to be present with AD, to the slowing of behavioural response times when making lexicality judgements. This would suggest that a linguistically related deficit is also contributing to the slowing of response times when individuals with AD make a lexicality judgement.

Additional Observations Regarding our P3 Results

A reduction in P3 component amplitude can occur due to an alteration in attentional processes, yet the absence of a P3 component observed in the AD group for any ERP block of Experiment 2 is not likely to be attributable to a deficit in this cognitive domain. While the attention screening performed by the participants was not comprehensive, the results of the two attention tasks we administered (reported in the Appendix) suggest that neither participant group displayed a marked alteration in attentional processes that would interfere with their ability to selectively attend to task critical stimuli in the ERP tasks. Nonetheless, the pattern of P3 response observed for those with AD contrasts with what we observed in this study for the older adults and what has been observed for young adults in our previous research (Azevedo, Atchley, & Kehayia, 2015). Furthermore, this pattern appears to be mediated by a decrease in sensitivity to the orthography/phonology of the stimuli in those with AD. Thus, we believe that these findings suggest that the alteration in performance is more likely to be linguistically related rather than related to attention.

Furthermore, while we did observe that those with AD appeared to have made a significant number of behavioural errors in Experiment 2, upon closer inspection these were revealed to mostly consist of responses made before the response prompt was presented. As a result, these responses were not recorded and were counted as errors. It is not possible to determine how many of these non-responses were true errors. However, we would argue that, given the high rate of accuracy for Experiment 1, it is likely that the high error rate observed across all trial types for those with AD in Experiment 2 is an artifact caused by task demands. We propose that impulsivity, a condition that is associated with AD but is unrelated to deficits in attention, led to the failure in their ability to withhold response and is responsible for this pattern of responses. Although this difficulty in withholding their behavioural response until the presentation of the response prompt in the ERP tasks was unforeseen, it was not entirely surprising given that individuals with AD frequently show such difficulty even early in the course of the disease (Amieva, Lafont, Auriacombe, Carret, Dartigues, et al. 2002; Collette, Amieva, Adam, Hogge, Van der Linden, Fabrigoule, & Salmon, 2007).

While a traditional P3 component typically has a latency of approximately 300-600 ms when elicited to non-linguistic stimuli, in the current study we observed that task conditions requiring the processing of linguistic stimuli have led to a delay in latency of the component. This delay was also observed in our previous lexical decision oddball task conducted with young adults (Azevedo, Atchley, & Kehayia, 2015). It is also in line with previous ERP studies that investigate the timeline of lexical processing steps that must be performed before a “word” or “not a word” categorization can be made (and thus must occur before a P3 component can be elicited in any of our ERP tasks). Before a P3 component could be elicited in the two blocks contrasting words and nonwords, each stimulus first had to be processed at least until orthographic/phonological legality could be determined. ERP research has shown that a divergence in waveforms associated with phonologically legal versus illegal stimulus types occurs relatively late, i.e. in the 300-350 ms timeframe (Bentin et al., 1999; Massol, Midgley, Holcomb, & Grainger, 2011; Massol, Grainger, Midgley, & Holcomb, 2012; Spironelli, Penolazzi & Angrilli, 2010). Furthermore, before a P3 component could be elicited in the two blocks contrasting words and pseudowords, an attempt to determine the lexical status of the stimulus needed to be performed since both stimulus types were legal. ERP research points to an even later divergence in waveforms associated with an attempt to access a lexical representation for legal stimuli, i.e. in the 400 ms timeframe (Coch & Mitra, 2010; Deacon, Dynowska, Ritter, & Grose-Fifer, 2004). Therefore, since elicitation of the P3 component was dependent on initial steps involved in processing lexical stimuli we were not surprised that this component was delayed in our tasks.

Study Limitations

The primary limitation of the current research is the small sample of AD participants that we were able to recruit for our study. In many ways, the current experiments should be viewed as a large case study design, rather than a typical research experiment as might be run with large numbers of healthy individuals that are more easily recruited. In addition to having fewer participants than would be optimum, we also have a reliable relationship between AD status and age that potentially creates a situation where age, rather than AD status might be driving some portion of the observed differences between our two participant samples. While we cannot rule out that age may be playing a role in the differences we observe between the groups, we do believe that age is not the main driver of these differences. We make this claim because we have rerun our ANOVA with age as a covariate and we observed the same general pattern of results. We also reran the ANOVA including only the older adults aged 65 years or over and again we observed the same pattern of results[1]. Future research is clearly needed to address these concerns regarding experimental power and inter-relationships between age and AD status which

[1] Results from the mixed model 2 X 2 X 4 X 3 omnibus ANOVA for only those participants over age 65 showed a significant main effect of Trial Type ($F(1,21) = 11.51$, $p = 0.003$), and of Electrode channel ($F(2,42) = 3.28$, $p = 0.047$). There was no main effect for the variable of Participant group or Block. There was a significant interaction between Channel and Trial Type ($F(2,42) = 4.60$, $p = 0.02$), and a marginally significant Block by Trial Type by Channel by Group interaction ($F(6,126) = 2.09$, $p = 0.06$). No other effects were significant.

make it necessary for researchers to consider our findings with some degree of caution. Nonetheless, we believe that the current experiments provide guidance to this future research and indicate that even with small sample sizes some important and informative differences might be observed when we ask healthy older adults and individuals with AD to make lexical decisions.

Future Direction

The ERP results point to a change in how lexicality is processed in individuals with AD that may only become visible in behavioural responses later in the course of the disease. More specifically, these results suggest a change that occurs at the orthographic/phonological stage of processing. The current study was designed specifically to elicit the P3 component and consequently we employed a 20/80 ratio of rare to frequent stimuli. While the oddball task used here is the ideal experimental task to elicit the P3 component it does not easily permit us to investigate subsequent ERP components that may reflect orthographic/phonological processing. One possible way to further investigate the extent to which individuals with AD are sensitive to orthography/phonology could be to investigate the N350 component. This component has been linked to phonological processing (Ruz and Nobre, 2008; Spironelli & Angrilli, 2007; 2009; Spironelli et al., 2010) and is thought to distinguish between pronounceable and non-pronounceable stimuli since it is elicited only by orthographically/ phonologically legal stimuli (Bentin et al., 1999) in healthy young adults.

While additional research is needed to further probe the nature of this apparent deficit in processing lexicality in those with AD, further research is also warranted to further investigate if (and how) this linguistic deficit may also be contributing to the behavioural slowing of RTs that was observed in the lexical decision task, over and above general slowing in processing speed.

Acknowledgement

This study was supported by a joint doctoral training award to the first author from the Alzheimer Society of Canada and the FRQS (Fonds de recherche du Québec- Santé). Preparation of this manuscript was supported by the Social Sciences and Humanities Research Council of Canada Words in the World training initiative.

References

Ally, B. A., Jones, G. E., Cole, J. A., & Budson, A. E. (2006). The P300 component in patients with Alzheimer's disease and their biological children. *Biological Psychology*, *72*(2), 180-187.

Amieva, H., Lafont, S., Auriacombe, S., Carret, N. L., Dartigues, J. F., et al. (2002). Inhibitory breakdown and dementia of the Alzheimer type: a general phenomenon? *Journal of Clinical and Experimental Neuropsychology, 24*(4), 503-516.

Appell, J., Kertesz, A., & Fisman, M. (1982). A study of language functioning in Alzheimer patients. *Brain and Language*, *17*(1), 73-91.

Ashford, J. W., Coburn, K. L., Rose, T. L., & Bayley, P. J. (2011). P300 energy loss in aging and Alzheimer's disease. *Journal of Alzheimer's* Disease, 26, 229–238.

Azevedo, N., Atchley, R. A., & Kehayia, E. (2015). Electrifying the lexical decision: Examining a P3 ERP component reflecting early lexical categorization. *The Mental Lexicon*, *10*(3), 339-363.

Azevedo, N., Kehayia, E., Atchley, R. A., & Nair, V. P. N. (2015). Lexicality judgements in healthy aging and in individuals with Alzheimer's disease: Effect of neighbourhood density. *Mental Lexicon, 10*(2), 286-311.

Azizian, A., Freitas, A. L., Watson, T. D., & Squires, N. K. (2006). Electrophysiological correlates of categorization: P300 amplitude as index of target similarity. *Biological Psychology, 71*(3), 278-288.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283.

Bayles, K. A. & Kaszniak, A. W. (1987). The brain and age-related dementing diseases. In K. A. Bayles & A. W. Kaszakniak (Eds.), C*ommunication and Cognition in Normal Aging and Dementia*. Boston: College-Hill Press.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M., Echallier, J., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: time course and scalp distribution. *Journal of Cognitive Neuroscience, 11*(3), 235-260.

Bowles, N. L., & Poon, L. W. (1981). The effect of age on speed of lexical access. *Experimental Aging Research, 7*(4), 417-425.

Braaten, A. J., Parsons, T. D., McCue, R., Sellers, A., & Burns, W. J. (2006). Neurocognitive differential diagnosis of dementing diseases: Alzheimer's dementia, vascular dementia, frontotemporal dementia, and major depressive disorder. *International Journal of Neuroscience*, *116*(11), 1271-1293.

Bruder, G. E., Tenke, C. E., Stewart, J. W., Towey, J. P., Leite, P., Voglmaier, M., & Quitkin, F. M. (1995). Brain event-related potentials to complex tones in depressed patients: Relations to perceptual asymmetry and clinical features. *Psychophysiology*, *32*(4), 373-381.

Carreiras, M., Armstrong, B. C., Perea, M., Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends Cognitive Science 18*(2), 90-98.

Caza, N., & Moscovitch, M. (2005). Effects of cumulative frequency, but not of frequency trajectory, in lexical decision times of older adults and patients with Alzheimer’s disease. *Journal of Memory and Language, 53*(3), 456-471.

Chapman, R. M., Nowlis, G. H., McCrary, J. W., Chapman, J. A., Sandoval, T. C., Guillily, M. D., Gardner, M .N., & Reilly, L. A. (2007). Brain event-related potentials: Diagnosing early-stage Alzheimer's disease. *Neurobiology of Aging*, *28*(2), 194-201.

Coch, D., & Mitra, P. (2010). Word and pseudoword superiority effects reflected in the ERP waveform. *Brain Research, 1329*, 159-174.

Collette, F., Amieva, H., Adam, S., Hogge, M., Van der Linden, M., Fabrigoule, C., & Salmon, E. (2007). Comparison of inhibitory functioning in mild Alzheimer's disease and frontotemporal dementia. *Cortex*, *43*(7), 866-874.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204-256.

Cuetos, F., Herrera, E., & Ellis, A. W. (2010). Impaired word recognition in Alzheimer's disease: The role of age of acquisition. *Neuropsychologia*, *48*(11), 3329-3334.

Davies, R. A., Arnell, R., Birchenough, J. M., Grimmond, D., & Houlson, S. (2017). Reading through the life span: Individual differences in psycholinguistic effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(8), 1298.

Deacon, D., Dynowska, A., Ritter, W., & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: Implications for theories of N400 and word recognition. *Psychophysiology*, *41*(1), 60-74.

Duñabeitia, J. A., Marín, A., & Carreiras, M. (2009). Associative and orthographic neighborhood density effects in normal aging and Alzheimer's disease. *Neuropsychology, 23*(6), 759-764.

Duñabeitia, J. A., Marín, A., Aviles, A., Perea, M., & Carreiras, M. (2009). Constituent priming effects: Evidence for preserved morphological processing in healthy old readers. *European Journal of Cognitive Psychology*, *21*(2-3), 283-302.

Duong, A., Whitehead, V., Hanratty, K., & Chertkow, H. (2006). The nature of lexico-semantic processing deficits in mild cognitive impairment. *Neuropsychologia, 44*(10), 1928-1935.

Edwards, J. D., Vance, D. E., Wadley, V. G., Cissell, G. M., Roenker, D. L., & Ball, K. K. (2005). Reliability and validity of useful field of view test scores as administered by personal computer. *Journal of Clinical and Experimental Neuropsychology*, *27*(5), 529-543.

Edwards, J. D., Ross, L. A., Wadley, V. G., Clay, O. J., Crowe, M., Roenker, D. L., & Ball, K. K. (2006). The useful field of view test: normative data for older adults. *Archives of Clinical Neuropsychology, 21(*4), 275-286.

Eisenhauer, S., Fiebach, C. J., & Gagl, B. (2018). Dissociable prelexical and lexical contributions to visual word recognition and priming: Evidence from MEG and behavior. *bioRxiv*, 410795.

Feldman, H. H., & Woodward, M. (2005). The staging and assessment of moderate to severe Alzheimer disease. *Neurology, 65*(6 suppl 3), S10-S17.

Feldman, H. H., Van Baelen, B., Kavanagh, S. M., & Torfs, K. E. (2005). Cognition, function, and caregiving time patterns in patients with mild-to-moderate Alzheimer disease: a 12-month analysis. *Alzheimer Disease & Associated Disorders*, *19*(1), 29-36.

Flowers, J. H., Warner, J. L., & Polansky, M. L. (1979). Response and encoding factors in "ignoring" irrelevant information. *Memory & Cognition*, *7*(2), 86-94.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-Mental State." A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189-198.

Folstein, M. F., Robins, L. N., & Helzer, J. E. (1983). The mini-mental state examination. *Archives of General Psychiatry*, *40*(7), 812.

Fozard, J. L., & Gordon-Salant, S. (2001). Changes in vision and hearing with aging. Handbook of the psychology of aging, 5, 241-66.

Glosser, G., Kohn, S. E., Friedman, R. B., Sands, L., & Grugan, P. (1997). Repetition of single words and nonwords in Alzheimer's disease. *Cortex 33*, 653-666.

Gold, B. T., Andersen, A. H., Jicha, G. A., & Smith, C. D. (2009). Aging influences the neural correlates of lexical decision but not automatic semantic priming. *Cerebral Cortex*, *19*(11), 2671-2679..

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468-484.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, *103*(3), 518-565.

Hale, S., & Myerson, J. (1996). Experimental evidence for differential slowing in the lexical and nonlexical domains. *Aging, Neuropsychology, and Cognition, 3*(2), 154-165.

Joubert, S., Joncas, S. Barbeau, E., Joanette, Y. & Ska, B. (2006). Cognition. In S. Gauthier (Ed), *Clinical diagnosis and management of Alzheimer's disease, third edition* (pp. 165-176). Abington: Informa Healthcare.

Juckel, G., Karch, S., Kawohl, W., Kirsch, V., Jager, et al. (2012). Age effects on the P300 potential and the corresponding fMRI BOLD-signal. *Neuroimage, 60*(4), 2027-2034.

Katada, E., Sato, K., Ojika, K., & Ueda, R. (2004). Cognitive event-related potentials: useful clinical information in Alzheimer's disease. *Current Alzheimer Research, 1*(1), 63-69.

Kavé, G., & Levy, Y. (2005). The processing of morphology in old age: Evidence from Hebrew. *Journal of Speech, Language, and Hearing Research, 48*(6), 1442-1451.

Kavé, G., & Goral, M. (2018). Word retrieval in connected speech in Alzheimer's disease: a review with meta-analyses. *Aphasiology, 32*(1), 4-26.

Lee, M. S., Lee, S. H., Moon, E. O., Moon, Y. J., Kim, S., Kim, S. H., & Jung, I. K. (2013). Neuropsychological correlates of the P300 in patients with Alzheimer's disease. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 40*, 62-69.

Lima S. D., Hale S., & Myerson J. (1991). How general is general slowing? Evidence from the lexical domain. *Psychology and Aging, 6,* 416-425.

Luis, C. A., Keegan, A. P., & Mullan, M. (2009). Cross validation of the Montreal Cognitive Assessment in community dwelling older adults residing in the Southeastern US. *International Journal of Geriatric Psychiatry, 24*(2), 197-201.

Madden, D. J. (1992). Four to ten milliseconds per year: age-related slowing of visual word identification. *Journal of Gerontology*, *47*, 59-68.

Madden, D. J., Welsh-Bohmer, K. A., & Tupler, L. A. (1999). Task complexity and signal detection analyses of lexical decision performance in Alzheimer's disease. *Developmental Neuropsychology, 16*(1), 1-18.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin, 109*(2), 163-203.

Massol, S., Midgley, K. J., Holcomb, P. J., & Grainger, J. (2011). When less is more: feedback, priming, and the pseudoword superiority effect. *Brain Research, 1386,* 153-164.

Massol, S., Grainger, J., Midgley, K. J., & Holcomb, P. J. (2012). Masked repetition priming of letter-in-string identification: An ERP investigation. *Brain Research*, *1472,* 74-88.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease Report of the NINCDS-ADRDA Work Group* under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology, 34*(7), 939-939.

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack Jr, C. R., et al. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, *7*(3), 263-269.

Mecklinger, A., & Ullsperger, P. (1993). P3 varies with stimulus categorization rather than probability. *Electroencephalography and Clinical Neurophysiology, 86*(6), 395-407.

Moberg, M., Ferraro, F. R., & Petros, T. V. (2000). Lexical properties of the Boston Naming Test stimuli: Age differences in word naming and lexical decision latency. *Applied Neuropsychology*, *7*(3), 147-153.

Myerson, J., Ferraro, F. R., Hale, S., & Lima, S. D. (1992). General slowing in semantic priming and word recognition. *Psychology and Aging, 7*(2), 257.

Myerson, J., Hale, S., Chen, J., & Lawrence, B. (1997). General lexical slowing and the semantic priming effect: The roles of age and ability. *Acta Psychologica, 96*(1-2), 83-101.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695-699.

Nikolaev, A., Higby, E., Hyun, J., & Ashaie, S. (2019). Lexical decision task for studying written word recognition in adults with and without dementia or mild cognitive impairment. *JoVE (Journal of Visualized Experiments), (148),* e59753.

Olichney, J. M., Yang, J. C., Taylor, J., & Kutas, M. (2011). Cognitive event-related potentials: biomarkers of synaptic dysfunction across the stages of Alzheimer's disease. *Journal of Alzheimer's Disease, 26*, 215-228.

Parra, M. A., Ascencio, L. L., Urquina, H. F., Manes, F., & Ibáñez, A. M. (2012). P300 and neuropsychological assessment in mild cognitive impairment and Alzheimer dementia. *Frontiers in Neurology, 3*, 172.

Pokryszko-Dragan, A., Słotwiński, K., & Podemski, R. (2003). Modality-specific changes in P300 parameters in patients with dementia of the Alzheimer type. *Medical Science Monitor Basic Research*, *9*(4), CR130-CR134.

Polich, J., & Corey-Bloom, J. (2005). Alzheimer's disease and P300: review and evaluation of task and modality. *Current Alzheimer Research*, *2*(5), 515-525.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*, 2128–2148.

Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging, 19*(2), 278.

Robert, C., & Mathey, S. (2007). Aging and lexical inhibition: The effect of orthographic neighborhood frequency in young and older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *62*(6), P340-P342.

Röschke, J., Wagner, P., Mann, K., Fell, J., Grözinger, M. & Frank, C. (1996). Single trial analysis of event related potentials: A comparison between schizophrenics and depressives. *Biological Psychiatry, 40*(9), 844-852.

Ruz, M., & Nobre, A. C. (2008). Attention modulates initial stages of visual word processing. *Journal of Cognitive Neuroscience, 20*(9), 1727-1736.

Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends in Cognitive Sciences, 7*(11), 489-493.

Shaoul, C., & Westbury C. (2006) USENET Orthographic Frequencies for 111,627 English Words. (2005-2006). Edmonton, AB: University of Alberta (downloaded from http://www.psych.ualberta.ca/~westburylab/downloads/wlfreq.download.html).

Spironelli, C., & Angrilli, A. (2007). Influence of phonological, semantic and orthographic tasks on the early linguistic components N150 and N350. *International Journal of Psychophysiology, 64*(2), 190-198.

Spironelli, C., & Angrilli, A. (2009). Developmental aspects of automatic word processing: language lateralization of early ERP components in children, young adults and middle-aged subjects. *Biological Psychology, 80*(1), 35-45.

Spironelli, C., Penolazzi, B., & Angrilli, A. (2010). Gender differences in reading in school-aged children: an early ERP study. *Developmental Neuropsychology, 35*(4), 357-375.

Stadtlander, L. M. (1995). Age differences in orthographic and frequency neighborhoods. *Advances in Psychology, 110*, 72-86.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643-662.

Taler, V., & Jarema, G. (2007). Lexical access in younger and older adults: the case of the mass/count distinction. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *61*(1), 21-34.

Taler, V., & Phillips, N. A. (2008). Language performance in Alzheimer's disease and mild cognitive impairment: a comparative review. *Journal of Clinical and Experimental Neuropsychology*, *30*(5), 501-556.

Yap, M. J., & Balota, D. A. (2019). *Visual word recognition*. Oxford University Press.

APPENDIX

Attention Screening

Two standardized neuropsychological tasks measuring attention processes were administered: the number Stroop task (Flowers et al., 1979; Stroop, 1935; MacCleod, 1991) and the Useful Field of View Task (UFOV 6- Visual Awareness, Inc., Chicago). These provided a general evaluation of these non-linguistic related abilities and helped rule out an overt impairment at this level that may be interacting with individuals' performance on the oddball lexical decision tasks.

Methods

The number Stoop, sometimes called the counting Stroop, used in our study is a variant of the classic Stroop task that involves numbers rather than colors. Participants were presented with number words (either "one", "two" or "three") on a computer screen. Four sets of 20 trials were presented in random order; in two of the trial sets the participants were asked to name the number written on the screen while in the other two trials they were asked to count the number of words on the screen. Each set of trials could be either congruent (the name and the count match (e.g. "two two") or incongruent (e.g. "two two two"). Instructions were provided before the beginning of each set of trials and participants responded by pressing the appropriate number key on a keyboard. We conducted 4 practice trials before beginning each trial to ensure that participants were comfortable with the task. Response time and error rate were collected for each trial. While the classic Stroop task is more commonly used, it can be difficult to administer with older and/or cognitively impaired individuals as the response (a key press that corresponds to a colour) requires them to learn the mapping between the colour on the screen-key on the keyboard. Furthermore, there is a greater variability in colour perception with older adults that might impact the results of the task (Fozard & Gordon-Salant, 2001). By using the number Stroop we hoped to overcome these potential limitations. The Useful Field of View Task (UFOV) comprises three subtests that each measure age-related changes in information processing speed, proficiency in dividing attention, and the ability to ignore irrelevant information (Edwards, Vance, Wadley, Cissell, Roenker, et al., 2005). In the first subtest, participants were asked to identify a visual target (a car or a truck) presented at fixation. In the second, participants had to identify a central target and localize a peripheral target while in the third subtask participants were asked to make the same two responses, but with the added difficulty of the peripheral target being embedded among distractors.

Statistical Analyses

In the number Stroop, some individuals with AD had difficulty in response mapping, i.e. they were not consistently able to make a button-press response in the time allotted by the task. They were, however, able to verbally make their response. Given this unforeseen difficulty, individuals with AD who showed this difficulty were given the option to respond verbally and their response was imputed by the experimenter. Consequently, only response accuracy was included for analysis. Trials from the two congruent sets and those from the two incongruent sets were collapsed to form two overall sets (congruent and incongruent). Due to the presence of many perfect scores we elected to use A’ instead of d’ to evaluate stimulus sensitivity.

For the UFOV, a report providing an overall threshold value for each of the three subtests of the test is automatically generated following completion of the task. The three-test total for each participant was compared to normative scores (Edwards, Ross, Wadley, Clay, Crowe, et al., 2006) to determine if participants were within the performance norms expected for their age and level of education.

Results

Results for both tasks used in the attention screening are presented below. Results from the A' analysis for the accuracy data show that both groups are performing well above threshold in this task. The mean A' score for the OA group was 0.98 (range: 0.94 - 1) and 0.95 (range: 0.88 - 1) for individuals with AD. Two sample t-tests showed no significant difference in mean A' scores between the two groups ($t = -1.69$, $p = 0.12$) indicating that neither group appears to show an attention deficit when performing the number Stroop task. Furthermore, while individuals with AD had difficulty in making an overt button press response in this task and this differentiated them from the older adults, it is likely that a higher-order deficit (namely response mapping), and not an attention deficit, is responsible for this difficulty given that both groups showed a high sensitivity to the stimuli (A') in this task. For the UFOV, when compared to norms for their age and level of education (Edwards et al., 2006), only one healthy older adult and three individuals with AD had a three-test total score that was below their expected norms. All other participants' performance indicated normal attention as measured by this task. Overall, the combined results from these two tasks suggest that both participant groups' attention processes are within normal levels and that any potential disruptions in the P3 component, if found, are not be likely to be primarily attributable to marked deficits in these cognitive domains.

Number Stroop task

	Errors Congruent Trial Set (n = 40)	Errors Incongruent Trial Set (n = 40)	A' (Stimulus Sensitivity)
OA 1	3%	5%	0.98
OA 2	0%	0%	1.00
OA 3	0%	3%	0.99
OA 4	0%	0%	1.00
OA 5	0%	0%	1.00
OA 6	0%	18%	0.96
OA 7	0%	3%	0.99
OA 8	0%	0%	1.00
OA 9	0%	0%	1.00
OA 10	0%	5%	0.99
OA 11	10%	13%	0.94
OA 12	0%	3%	0.99
OA 13	0%	23%	0.94
OA 14	0%	0%	1.00
OA 15	3%	0%	0.94

OA 16	3%	0%	0.94
OA 17	0%	0%	1.00
Mean	1%	4%	0.98
St-Dev	3%	7%	

Performance on the Number Stroop task for older adults

Number Stroop task

	Errors Congruent Trial Set (n = 40)	Errors Incongruent Trial Set (n = 40)	A' (Stimulus Sensitivity)
AD 1	8%	33%	0.88
AD 2	0%	0%	1.00
AD 3	30%	8%	0.89
AD 4	0%	0%	1.00
AD 5	15%	25%	0.88
AD 6	0%	0%	1.00
AD 7	8%	20%	0.92
AD 8	0%	0%	1.00
AD 9	5%	15%	0.95
AD 10	0%	10%	0.98
Mean	7%	11%	0.95
St-Dev	10%	12%	

Performance on the Number Stroop task for individuals with AD

Useful Field of View (UFOV)

	Subtest 1 (Processing Speed-Stimulus Identification)	Subtest 2 (Divided Attention)	Subtest 3 (Selective Attention)	3-Subtest Total	Within Norms for age/ed*
OA 1	17	23	86	126	yes
OA 2	17	60	127	204	yes
OA 3	17	17	87	121	yes
OA 4	17	178	178	373	yes
OA 5	23	37	287	347	yes
OA 6	17	20	203	240	yes
OA 7	17	17	150	184	yes
OA 8	53	500	not administered	n/a	no
OA 9	17	40	97	154	yes
OA 10	37	87	243	367	yes
OA 11	37	157	433	627	yes
OA 12	23	23	130	176	yes

OA 13	17	63	83	163	yes
OA 14	17	20	83	120	yes
OA 15	17	53	77	147	yes
OA 16	17	40	180	237	yes
OA 17	17	70	253	340	yes

Performance on the Useful Field of View (UFOV) task for older adults

** For the UFOV, the three-test total for each participant was compared to normative scores (Edwards et al., 2006) to determine if participants were within the performance norms expected for their age and level of education.*

Useful Field of View (UFOV)

	Subtest 1 (Processing Speed-Stimulus Identification)	Subtest 2 (Divided Attention)	Subtest 3 (Selective Attention)	3-Subtest Total	Within Norms for age/ed*
AD 1	140	500	not administered	n/a	no
AD 2	17	60	263	340	yes
AD 3	40	500	not administered	n/a	no
AD 4	113	113	330	556	yes
AD 5	30	157	316	503	yes
AD 6	17	50	257	324	yes
AD 7	17	103	270	390	yes
AD 8	153	220	340	713	yes
AD 9	133	500	not administered	n/a	no
AD 10	120	223	300	643	yes

Performance on the Useful Field of View (UFOV) task for individuals with AD

Address of correspondence:

Nancy Azevedo, School of Physical & Occupational Therapy, McGill University, 3654 Promenade Sir-William-Osler, Montréal (Québec) H3G 1Y5. Phone: 450-688-0550 ext. 536.

Email: nancy.azevedo@mail.mcgill.ca

Submitted on 27 June 2020; Accepted on 07 May 2021