Published in final edited form as:
J Exp Psychol Learn Mem Cogn. 2013 November ; 39(6): . doi:10.1037/a0032860.

Cross-age comparisons reveal multiple strategies for lexical ambiguity resolution during natural reading

Mallory C. Stites[1,2], **Kara D. Federmeier**[1,2,3], and **Elizabeth A. L. Stine-Morrow**[1,2,4]
[1]Department of Psychology, University of Illinois, Urbana-Champaign

[2]Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign

[3]Program in Neuroscience, University of Illinois, Urbana-Champaign

[4]Department of Educational Psychology, University of Illinois, Urbana-Champaign

Abstract

Eye-tracking was used to investigate how younger and older (60+) adults use syntactic and semantic information to disambiguate noun/verb (NV) homographs (e.g. *park*). In event-related potential work using the same materials, Lee and Federmeier (2009, 2011) found that young adults elicited a sustained frontal negativity to NV-homographs when only syntactic cues were available (i.e., in syntactic prose); this effect was eliminated by semantic constraints. The negativity was only present in older adults with high verbal fluency. The current study shows parallel findings: young adults exhibit inflated first fixation durations to NV-homographs in syntactic prose, but not semantically congruent sentences. This effect is absent in older adults as a group. Verbal fluency modulates the effect in both age groups: high fluency is associated with larger first fixation effects in syntactic prose. Older, but not younger, adults also show significantly increased rereading of the NV-homographs in syntactic prose. Verbal fluency modulates this effect as well: high fluency is associated with a reduced tendency to reread, regardless of age. This relationship suggests a tradeoff between initial and downstream processing costs for ambiguity during natural reading. Together, the eye-tracking and ERP data suggest that effortful meaning selection recruits mechanisms important for suppressing contextually inappropriate meanings, which also slow eye movements. Efficacy of fronto-temporal circuitry, as captured by verbal fluency, predicts the success of engaging these mechanisms in both young and older adults. Failure to recruit these processes requires compensatory rereading or leads to comprehension failures (Lee & Federmeier, 2012).

Keywords

eye-tracking; ERPs; noun/verb homographs; aging; individual differences

Introduction

The English language is full of one-to-many mappings, in which an identical written wordform can represent a number of distinct meanings. Yet skilled readers have the ability to move easily from these written forms to meaning, often failing to notice the ambiguity at all. It remains one of the main goals of psycholinguistic research, then, to understand the

Address correspondence to: Mallory Stites, Department of Psychology, University of Illinois, 603 E. Daniel St., Champaign, IL 61820, Phone: 636.208.5809, stites2@illinois.ed.

processes underlying lexical ambiguity resolution and the factors that affect how these processes unfold.

The most widely studied type of lexical ambiguity involves words whose meanings fall within a single syntactic class, generally nouns (e.g., *calf*; henceforth referred to as NN homographs). Work across methodologies has found that two key factors best determine the processing costs associated with NN homographs: meaning frequency and the information provided by the context in which the ambiguous word is embedded. For words with two equally-likely meanings (i.e., balanced homographs) in a semantically neutral context, eye-tracking measures have consistently revealed increased reading times compared to unambiguous words, an effect thought to reflect competition for selection between the two equally-activated meanings (Rayner & Frazier, 1989). In contrast, when the context biases interpretation of a balanced homograph towards one meaning, reading times are indistinguishable from those for unambiguous words. Reading times on biased NN homographs (whose meanings are of unequal frequencies) are longer relative to unambiguous words, but only when the preceding context instantiates the subordinate, or less frequent, meaning. This well-characterized eye-tracking result has been called the *subordinate bias effect* (Rayner, Pacht & Duffy, 1994), a phenomenon explained by the reordered access model as reflecting competition between the two meanings, in which the dominant sense is not completely suppressed (Duffy, Morris, Rayner, 1988). Using event-related potentials (ERPs), Swaab, Brown and Hagoort (2003) demonstrated that in spoken language comprehension, both meanings of ambiguous words are partially activated initially, regardless of context. However, similar to the reading results, they found that the dominant meaning is harder to suppress and active longer in an inappropriate context than the subordinate meaning. There is some evidence that suppression is an effortful process, in that a reader's working memory span has been found to predict the ability to suppress an ambiguous word's dominant meaning in an inappropriate context (Gunter, Wagner, & Friederici, 2003).

Across methods, therefore, the literature on comprehension of NN homographs has converged to show that the dominant meaning of a biased ambiguous word is easier to activate and harder to suppress than its subordinate meaning. Importantly, preceding context exerts a large influence over how much activation each meaning receives, thus determining whether there will be competition between the two meanings and associated processing costs. Given the importance of context in meaning resolution, a critical question is how and how much different aspects of context contribute to on-line ambiguity resolution. Biasing sentential contexts typically provide both semantic and syntactic cues as to the identity of the upcoming word. In the case of NN homographs, whose meanings fall into the same syntactic category, the role of syntax in ambiguity resolution cannot be determined. However, meaning ambiguity can also co-occur with ambiguity about a word's part of speech, as in the case of words whose meaning varies across noun and verb senses (e.g., *duck*; henceforth NV homographs).

In the case of NV homographs, it becomes possible to examine the contributions of semantic and syntactic aspects of context separately. In contrast to work with NN homographs, studies using NV homographs have yielded contradictory results, but have also used different paradigms and materials. One important goal of the current study, therefore, is to take a critical first step toward reconciling those differences, particularly regarding the question of whether syntax alone can produce selective access for NV homographs, by measuring eye movements during natural reading using the materials and design of a study that previously measured ERPs during the processing of NV homographs. Part of the discrepancy across methods may be due to the fact that different strategies are afforded when people can control the pacing and order with which they encounter information and

can revisit problematic parts of a text, as during natural reading, as opposed to during listening and word-by-word reading tasks often used with ERP measures, where such control is not an option. These differences in control may also impact how ambiguity resolution is accomplished by people of different ages and cognitive abilities; as detailed below, a second goal of the study is therefore to investigate the role that aging and individual differences play in determining ambiguity effects for NV homographs in natural reading. In this way, our study can begin to unite the disparate findings from the study of NV homographs, further clarify the role of syntax in ambiguity resolution, and move towards producing a more general account of the processes underlying ambiguity resolution.

Eye-tracking results have been taken to suggest that disambiguating syntactic information can allow selective access for NV homographs. Folk and Morris (2003) embedded biased NN (e.g., *Before the game, the baseball team went to the **diamond** to practice*) and NV (e.g. *Biking through Utah, the cyclist lost a **spoke** in the mountains*) homographs in sentences that constrained the word's meaning through both semantic and syntactic context. Although the typical subordinate bias effect was observed for the NN homographs, this effect was absent for the NV homographs. That is, gaze durations were not inflated on the NV homographs when the subordinate meaning was intended, suggesting that the context provided enough information to selectively activate only the subordinate meaning of the word. In a second experiment, balanced NV homographs were embedded in sentences that were semantically neutral but that syntactically instantiated the word's noun sense (e.g. *Laurie took the **prune** out of the fruit bowl and ate it*). Results showed that there were no reading time costs for the NV homographs compared to unambiguous controls, either on the critical words or later in the sentence. Taken together with the lack of a subordinate bias effect, these results led the authors to conclude that syntax alone provided enough information to allow readers to preselect the correct meaning of the NV homograph so as to eliminate the costs typical of ambiguity resolution based on semantic constraints alone.

In contrast, behavioral studies using cross-modal priming have suggested that both meanings of NV homographs are initially active despite disambiguating syntax. For example, Tanenhaus, Leiman and Seidenberg (1979) found facilitated naming times for targets related to either sense of NV homographs embedded in semantically neutral but syntactically constraining sentences (i.e. *I bought the **watch**/I will **watch***) when the targets were presented immediately, but not after a 200 ms delay. Other, similar studies have replicated this basic finding (Seidenberg, Tanenhaus, Leiman & Bienkowski, 1982; Tyler & Marslen-Wilson, 1977). These results have been taken as evidence for a multiple-access account of lexical ambiguity resolution in which both meanings of a NV homograph are initially activated regardless of syntactic cues, before suppression of the contextually inappropriate meaning.

Finally, work using ERPs has also suggested that syntactic cues alone do not allow selective access, and have further suggested that the disambiguation of NV homographs unfolds in a qualitatively different manner in the presence of syntactic cues alone as compared to when semantic cues are also available. Lee and Federmeier (2009) embedded NV homographs (with distinct noun and verb meanings) and matched unambiguous words in two types of sentences: congruent sentences that were both semantically and syntactically coherent with only one interpretation (noun or verb sense) of the homograph (e.g., *A key strategy to winning in a poker game is to **bluff***) and syntactic prose sentences that maintained the same syntactic structure but lacked coherent semantics (e.g., *A surprised shirt to winning in a time girl is to **bluff***). When ERPs time-locked to the onset of the target words were examined, the ambiguous NV homographs elicited responses that were more negative than the unambiguous controls over frontal channels between approximately 200 and 700 milliseconds post-stimulus onset, but only in the syntactic prose sentences. This finding replicated several previous studies, which also found frontal negativity to NV homographs

(both in sentence-final and sentence-medial positions) in the presence of a variety of types of syntactically, but not semantically, constraining contexts (Federmeier, Segal, Lombrozo & Kutas, 2000; Lee & Federmeier, 2006).

By contrast, in the congruent sentences, the waveforms elicited by the two word types were indistinguishable over frontal channels, suggesting that the addition of semantics to the well-defined syntax eliminated the processing costs that are indexed by the frontal negativity for the class-ambiguous words. Semantic processing was clearly facilitated for both ambiguous and unambiguous words in the congruent contexts compared to the syntactic prose ones, as seen in an amplitude reduction in the N400, a negative-going waveform that peaks over central-parietal electrodes around 400 milliseconds post-stimulus onset and is part of the normal brain response to words and other meaningful stimuli (see Kutas & Federmeier, 2011, for a review). N400 responses to ambiguous words used in their dominant sense were facilitated to the same degree as unambiguous words. When, instead, the semantic context picked out the nondominant sense of an ambiguous word, there was less facilitation of the N400, suggesting residual activation of semantic features associated with the word's dominant sense (i.e., a subordinate bias effect). Critically, however, even in this case there was no frontal negativity effect, showing that presence of semantically constraining information eliminates the need for certain meaning selection processes, even when access is not fully selective.

The ERP findings thus suggest that in syntactically well-defined but semantically neutral or impoverished contexts, the disambiguation of NV homographs requires the engagement of additional neurocognitive resources. This effect has been posited to reflect the recruitment of frontally-mediated selection mechanisms that serve to suppress the context-inappropriate meaning of the target word when readers have only syntactic cues. To test this hypothesis, Lee and Federmeier (2011, 2012) examined older adults' ability to disambiguate NV homographs using syntactic cues alone. Consistent with the idea that normative aging is associated with deterioration of the frontal lobe (Raz et al., 2005) that produces reductions in the ability to engage controlled, top-down processes (DiGirolamo et al., 2001), older adults did not show the frontal negativity effect exhibited by the young adults in Lee and Federmeier (2009). This thus suggests that older adults were less able to recruit frontally-mediated selection mechanisms. Indeed, in a paradigm probing downstream activation of the two meanings of the NV-homographs (Lee & Federmeier, 2012), younger adults showed selective activation of the contextually-appropriate meaning, whereas older adults showed continued activation of the contextually-inappropriate meaning. However, prior work has shown that high verbal fluency is associated with young-like language processing abilities in older adults (Federmeier et al., 2002; Federmeier, Kutas, & Schul, 2010). Consistent with these findings, older adults with higher verbal fluency -- and thus those who were likely to have more preserved frontal lobe functioning (e.g., Henry & Crawford, 2004; Stuss & Levine, 2002) -- did show the frontal negativity. Moreover, participants with larger frontal negativities show greater downstream suppression of the contextually-inappropriate meaning (Lee & Federmeier, 2012).

Thus, the behavioral and ERP evidence converge to show that selection mechanisms need to be recruited to effect lexical ambiguity resolution when NV homographs are constrained by syntactic context in the absence of semantically biasing information. The data from eye-tracking stand apart in suggesting that syntactic cues may be more powerful than semantic cues for resolving lexical ambiguity, even to the extent of allowing the selection of subordinate meanings. Discrepancies in the experimental conditions employed across methodologies may have contributed to differences in outcomes. For example, the semantically neutral sentence contexts used by Folk and Morris (2003) in their eye-tracking study always picked out the noun interpretation of the NV homographs. If subjects

implicitly learned this pattern, it may have changed their reading strategies on these words, possibly reducing ambiguity effects that would have manifested otherwise. On the other hand, natural reading as measured with eye-tracking affords strategic control over the intake of information that is not available in the word-by-word reading paradigms typically used with ERP research. It is therefore possible that comprehenders approach the task of ambiguity resolution differently when, for example, they do or do not have the opportunity to revisit words. This difference in the opportunity to self-regulate input so as to engage reading strategies might be particularly important in the context of aging (Stine-Morrow, Gagne, Miller, & Hertzog, 2008), as the availability of cognitive resources and the timing of cognitive processes change (Salthouse, 2010; Tucker-Drob, 2011). Thus, employing an eye-tracking task with older adults will allow us to explore the question of whether some older adults are truly unable to recruit the resources necessary to resolve the ambiguity of NV homographs in the syntactic prose sentences, as suggested by the findings of Lee & Federmeier (2011, 2012) or whether the additional control that natural reading, as measured by eye-tracking, allows older adults over their intake of information might allow these processes to take place over a different (possibly later) timecourse than seen in younger adults.

Given the possibility that prior eye-tracking results might have been discrepant, at least in part, because of the stimuli used, in the current research we used the materials and design that were shown to recruit selection mechanisms in Lee and Federmeier's (2009, 2011) studies. In order to provide a more unified account of NV homograph resolution, we aim to discover whether there are similar ambiguity effects when those materials are read naturally and readers are free to control their own eye movements over the sentences. In Experiment 1, with young adults, we expect to see eye movement correlates of the processing difficulties measured by ERPs, perhaps in a slow-down in reading times on ambiguous words in the absence of biasing semantics. These results will also help answer the question whether syntactic cues alone can allow for selective access during natural reading. To the extent that any eye-tracking effects observed in the young adults reflect the same selection processes as those indexed by the ERPs, we should then expect, in Experiment 2, that these effects will be absent in older adults as a group. However, building on the patterns seen in the ERP work, we will also test for relationships between ambiguity effects and individual difference measures, particularly verbal fluency. Additionally, a significant advantage of using eye-tracking to study NV homograph ambiguity resolution is that it affords the unique opportunity to see if older adults, who may be unable to recruit the necessary selection mechanisms on their first encounter with the ambiguous word, can employ other resolution strategies, possibly in the form of returning to that word later. To the extent that eye movement effects associated with the processing of ambiguity track the specific patterns seen on the ERP frontal negativity -- i.e., are seen in syntactic prose but not congruent sentences for young adults and are reduced/absent in older adults, but modulated by verbal fluency -- we can feel confident that we are tapping into overlapping processes involved in the resolution of ambiguity across methods, an important step forward in providing a more general understanding of ambiguity resolution. This study represents the first eye-tracking study to use the same materials and design as a previous ERP study, providing us the chance to observe whether electrophysiological markers of difficult ambiguity resolution have behavioral correlates, and allowing the rare opportunity for the two measures to inform interpretations of one another.

Experiment 1: Young Adults

Methods

Participants—Eighteen UIUC undergraduate and graduate students (9 males; mean age 20.4 years, range 18–26) participated for cash ($8 per hour) or course credit. All were right-

handed, as assessed by the Edinburgh inventory (Oldfield, 1971) and were monolingual speakers of English, with no consistent exposure to another language before the age of five. Participants also had no history of neurological disease, psychiatric disorders or brain damage. Each was randomly assigned to a different experimental list.

Participants were assessed on working memory (reading span test; Daneman & Carpenter, 1980), response suppression/inhibition (Hayling test; Burgess & Shallice, 1996), and executive function (verbal fluency tests, including letter fluency [FAS] and category fluency [animals, fruits and vegetables and first names]; Benton & Hamsher, 1978). For the reading span test, participants read aloud sentences in groups progressing from two up to six sentences (with five sets within each span size), and at the end of each group were asked to recall the last word in each sentence. Performance was measured by assessing each participant's set size (the highest set size for which the participant recalled all items in at least three of the five groups), and the total number of items recalled (see Table 2 for a summary of all neuropsychological measures). In the Hayling test, participants were given sentences frames missing the last word, and had to provide a congruent ending (for phase I of the test) or an anomalous ending (for phase II). Performance was measured by counting the total number of trials for which participants provided the correct response type. For the fluency tests, participants were instructed to orally generate words that began with the letters F, A, and S in three separate 1-minute periods (for the letter fluency test) and to orally generate words that fell into the categories of animals, fruits and vegetables, and first names (for the category fluency test), again in three separate 1-minute periods.

Materials—The stimuli were the same as those used in Experiment 1 of Lee and Federmeier (2009). There were two categories of target words: NV homographs, whose meanings were both semantically and syntactically ambiguous (e.g. *the park/to park*), as well as unambiguous controls that were neither semantically nor syntactically ambiguous (e.g. *the logic/to choose*). Context sentences that preceded these targets were of two distinct types: *congruent*, which provided both coherent semantic and syntactic cues for disambiguation (e.g. *The children loved the swings and were excited when their mom took them to the park*), or *syntactic prose*, which provided the same syntactic cues but lacked meaningful semantics (e.g. *The horses broke the try-outs and were weighed when their woman loved them to the park*). To create syntactic prose sentences, content words from the congruent sentences were randomly exchanged with those from other congruent sentences, while the function words and sentence-final target words remained intact. In this way, the syntactic prose sentences were syntactically well-formed and thus clearly instantiated the part of speech for the ambiguous words, but lacked any clear message-level semantics. Word type and sentence type were fully crossed, creating four experimental conditions. A filler condition was also created in which the words from the syntactic prose sentences were scrambled, up to the target word. The scrambled sentences were included so that readers would not be able to predict whether syntactic or semantic information would be available from trial to trial.

Each critical sentence was continued with a short follow-up sentence, so that the target word was never the last word in the entire passage. This allowed for regressive movements back to the target words. Follow-up sentences for the congruent context were coherent continuations of the topic of the critical sentence, while the syntactic prose and random follow-up sentences were created in the same way as the syntactic prose sentences, as follow-up sentences for these two conditions were identical. Table 1 contains example critical and follow-up sentences for the congruent and syntactic prose contexts.

Subjects read 172 two-sentence passages, each consisting of a critical and a follow-up sentence. Six lists were generated so that target words could be rotated through all three

sentence conditions in both their noun and verb sense. Each list was randomized once, and then presented to subjects in the same randomized order. Subjects saw only one of the lists, so that they saw every target word once, and saw at least 28 items in each of the six conditions. The noun-verb homographs were equally likely to appear as a noun or a verb, and were always placed in a sentential context that clearly supported one meaning. The unambiguous words always appeared in appropriate syntactic frames. There were equal numbers of appearances of ambiguous and unambiguous words within each sentential context. The number of nouns and verbs was also matched across the ambiguous and unambiguous word sets.

The target words were matched across conditions, both experiment-wide and within each list, for several important linguistic characteristics: log-frequency (Francis & Kucera, 1967), word length, and usage-specific concreteness (Lee & Federmeier, 2006). Specifically, a control word was selected to match each sense of the target homograph in usage-specific characteristics, in order to match the targets and controls as closely as possible. In addition, the semantic distinctiveness of the homographs (i.e., the similarity of their noun and verb meanings), which was normed by Lee and Federmeier (2006), was controlled across context types. Only homographs that were previously rated as having semantically distinct meanings were included in the study.

The critical sentences were also matched on a range of features. Sentence length was equated across conditions, as were cloze probability and plausibility (Lee & Federmeier, 2009). To assess cloze probability, a different group of subjects was given the sentence frames up to (but excluding) the target word, and asked to complete the sentence. For the plausibility ratings, another group of subjects was given the sentences to read and asked to rate them on a scale of 1–7, where 1 represents “makes no sense at all” and 7 represents “makes perfect sense.” The sentences chosen for use in the current study had an average cloze probability rating of 50% and an average plausibility rating of 6.5 (for further details, refer to Lee & Federmeier, 2009).

Procedure—Subjects were tested individually in a quiet room, where they were seated 97 cm away from a 19-inch ViewSonic P225f monitor (resolution of 1024 × 768), with a refresh rate of 120 Hz. Before the experiment began, the head-mounted SR Research Eyelink II eye-tracker was fitted and calibrated for each subject using a nine-point calibration system. A chin rest was used to reduce head movements. Drift correction was done at the beginning of each trial. Recordings were monocular, taken from the eye determined by the Eyelink software to have the more accurate calibration and validation readings.

Subjects were presented with written instructions, after which they received nine practice trials before the experiment began to allow them to become acclimated to the trial sequence. Each trial began with the appearance of a drift check target, which appeared in the upper-left corner of the screen. Subjects controlled the start of each trial by fixating the drift-check target while simultaneously pressing an advance button on the hand-held controller. Each passage consisting of a critical sentence and its follow-up sentence was presented in left-justified, white, uppercase letters on a black background in size 30 Courier New Font, in which three characters subtended 1 degree of visual angle. Subjects were instructed to read the sentences normally and press the advance button on the controller when they were finished. The experimental sentences were followed by a probe word that was presented in the center of the screen; subjects judged whether the probe word was or was not in the previous trial by pressing designated “Yes” and “No” buttons on the controller, both of which were pressed with the right hand. Half of the probes were new and half were old. Of the old words, half were content words chosen randomly from the previous trial and half

were the sentence final target words. This task was included to insure that subjects were attending to the entire passage, especially in the syntactic prose and filler conditions. With the button press, the trial was completed, and the next trial began with the appearance of the drift-check in the upper-left corner of the screen. The experiment was divided into four blocks consisting of 42–44 trials, each of which took an average of 7–8 minutes.

After each block, subjects were given a sentence recognition task to complete on paper. In this task, they were presented with sentences, with equal numbers in each of the three context types, half of which were presented in the previous block and half of which were not. Of the new sentences, all were taken from other experimental lists, meaning that there would be lexical item overlap between the old and new sentences, ensuring that subjects had to attend to more than word-level information to answer correctly. Subjects were asked to place a check next to items that they believed were seen in the previous block. This task, like the probe word task, was intended as a check to ensure that subjects were attending to all sentences. The time spent completing this task also served as a break between blocks, as subjects were allowed to move their heads out of the chin rest, and the headset could also be removed, if desired.

The entire experimental session lasted approximately 120 minutes, including set-up and practice trials, and neuropsychological assessment. Participants were given a short break following the eye-tracking portion of the experiment before the neuropsychological tests were administered.

Data Analysis—Within the Eyelink II data analysis package, consecutive fixations less that 80 ms in duration and less than .5° apart in visual angle were combined into one fixation. Single fixations that were shorter than 80 ms or longer than 800 ms were then automatically excluded. Fixations shorter than 80 ms are unlikely to represent meaningful cognitive processing (Rayner, 1998), while fixations longer than 800 ms are likely the result of cases in which the tracker temporarily lost the eye, causing inaccurately inflated reading times. Approximately 1.5% of trials were excluded from the analysis due to track loss or program error.

Results

Behavioral Results

Word recognition task: Overall accuracy for the word recognition task was 94%, showing that participants were attending to the sentences closely enough to remember individual words. Memory performance was assessed using the discriminability index d', with scores and standard deviations shown in Table 3. For this analysis, and for all subsequent behavioral results, a 2×2 within-subjects Analysis of Variance (ANOVA) with factors of context (congruent vs. syntactic prose) and target word type (ambiguous vs. unambiguous) was conducted. Results revealed a significant main effect of context type, $F(1,17) = 11.41$, $p < .01$, with better memory for words appearing in congruent than in syntactic prose sentences. There was no main effect of target word type or significant interaction between the two factors (Fs < 1). This overall level of performance is comparable to that in the ERP study using these materials (in which participants were 98% accurate on average; Lee & Federmeier, 2009), and the pattern of effects is the same.

Sentence recognition task: Overall accuracy in the sentence recognition task was 83%, which, while lower than that for the word recognition task, still shows that participants were attending to the stimuli and encoding the sentences as a unit (see Table 3 for condition means and standard deviations). An omnibus ANOVA revealed a main effect of context, $F(1,17) = 17.6$, $p < .01$, with better memory for congruent sentences. There was no main

effect of target word type, but there was a significant interaction between the two factors, $F(1,17) = 4.82$, $p < .05$, indicating that within the syntactic prose context, memory performance was better following sentences containing ambiguous words, while the two were roughly equated in the congruent context. This interaction could be due to the reading time differences, whereby relatively longer reading times on ambiguous target words in the syntactic prose context (which will be discussed in detail below) may have led to better encoding of those sentences. This interaction was not present in the previous ERP experiment using these materials (Lee & Federmeier, 2009), in which words were presented at a predetermined rate. Additionally, the overall performance on the sentence recognition task was lower for natural reading than for the corresponding ERP experiment (in which average accuracy on the sentence recognition task was 93%), which is likely explained by the fact that every trial in the current stimulus set contained both a critical and filler sentence, thus doubling the number of sentences presented to participants.

Summary: Behavioral measures confirmed that participants paid attention to the individual words that they were reading while also deriving a coherent message-level meaning for each sentence. Adding coherent meaning to syntactic structure aided memory performance both on short-term single word recall and longer-term entire-sentence recall, a pattern that replicates that seen in Lee and Federmeier's (2009) corresponding ERP experiment. However, different from the pattern seen when these stimuli were encountered with the fixed presentation rate used for ERP recordings, the natural reading results indicate that young participants choose to spend more time reading syntactic prose sentences that contain ambiguous words than those without ambiguous words, contributing to better memory for these sentences overall.

Eye-Tracking Results—Reading times were examined to the target region, which consisted of the target word as well as the immediately preceding word. Target words in this stimulus set were always preceded by a pre-target word class cue, which provided clear class-disambiguating information. This pre-target cue was most often a function word (e.g., *to* or *the*), but was sometimes an adjective or other determiner before a noun instantiation (e.g., *his, another, first*), or an auxiliary or adverb before a verb instantiation (e.g., *could, not*). Importantly, the pre-target cue always clearly disambiguated the word-class of the subsequent NV homograph. These pre-target cues were skipped in 50.5% of cases before readers fixated the target word (less than the estimated 65% skipping rate of function words reported by Rayner (1998)). When readers did fixate the pre-target words in the current study, their short average word length means that at least part of the subsequent ambiguous word is likely to also be in the reader's field of view, as the estimated word identification span is 7–8 characters to the right of fixation for English readers (Rayner, 1998). Because readers' first apprehension of the target word was thus roughly evenly distributed between fixations to the target word and to its preceding word class cue, we created a region of interest encompassing both words. Furthermore, if readers encountered processing difficulty downstream, they may be likely to return to either the ambiguous word itself, or to the cue word directly preceding it, which contains the disambiguating syntactic information.

First fixation durations and gaze durations were collected for the target region as an index of initial processing time. First fixation duration is the length of the reader's first fixation on a word, whereas gaze duration is the sum of all fixations on the target word before the eyes leave it in either direction. Additionally, we examined two measures of later processing. The first is the probability that a regression will be made back to the word from a later point in the sentence at any time after the first pass (*regressions in*). The second measure is rereading time, which was calculated for each word by subtracting gaze duration from the word's total reading time, in order to create a metric of how long a word was reread after it was fixated during the first-pass. A word was assigned a rereading value of 0 if it was not refixated, in

order to capture the probability of refixating target words, which importantly varied across conditions. For all measures, a 2×2 by-subjects ANOVA (F_1) was conducted with the within-subjects factors of context (congruent vs. syntactic prose) and word type (ambiguous vs. unambiguous), and a 2×2 by-items ANOVA (F_2) was conducted with the within-items factor of context and the between-subjects factor of word type (because target words were only either ambiguous or unambiguous).

First fixation duration: In light of ERP results showing frontal negativity to the ambiguous words in the syntactic prose sentences beginning as early as 150–200 milliseconds post-stimulus onset, first fixation durations were examined as a measure of initial processing. Mean first fixation durations (averaged over participants) can be found in Table 4. Analyses revealed a significant main effect of context, $F_1(1,17) = 5.51$, $p < .05$, $F_2(1,170) = 4.62$, $p < .05$, indicating that reading times were longer in the syntactic prose than in the congruent context. There was no main effect of word type, $F_1 < 1$, $F_2(1,170) = 1.32$, $p = .25$. The interaction between the two factors was significant by subjects, $F_1(1,17) = 7.22$, $p < .05$, and marginal by items, $F_2(1,170) = 3.20$, $p = .07$. The marginal interaction by items is likely due to the fact that the ambiguity manipulation was a between items variable, and thus the by-items analysis had less power than the by-subjects analysis. Importantly, follow-up t-tests revealed that the critical 11 ms cost for ambiguous relative to unambiguous words in the syntactic prose context was significant by *both* subjects and items[1], $t_1(17) = 2.48$, $p < .05$, $t_2(171) = 2.07$, $p < .05$, whereas first fixations did not differ between word types in the congruent context, $t_1(17) = -.89$, $p = .39$, $t_2(171) = .3$, $p = .76$. When first fixations to the pre-target word class cue and to the target word (with no prior fixations to the cue) are examined separately, the same numeric pattern is seen (pre-target word: $t_1(17) = 2.08$, $p = .05$; target word: $t_1(17) = 1.87$, $p = .08$; note that each of these comparisons contains roughly half of the number of trials as in the combined region).

Gaze duration: Gaze durations include all fixations to the target region before the eyes first leave it. Mean gaze durations can be found in Table 5. Analyses showed a significant main effect of context, $F_1(1,17) = 65.26$, $p < .001$, $F_2(1,170) = 64.39$, $p < .001$, with longer gaze durations in syntactic prose sentences. No main effect of word type was found (F_1 and $F_2 < 1$), although the interaction between the two factors was significant, $F_1(1,17) = 9.85$, $p < .001$, $F_2(1,170) = 7.89$, $p < .01$. Follow-up t-tests showed that unambiguous words were read significantly longer than ambiguous words by subjects (and marginally so by items) in the congruent context, $t_1(17) = -3.13$, $p < .01$, $t_2(171) = -1.88$, $p = .06$, but that the 20 ms cost for ambiguous words in the syntactic prose context, while numerically in the same direction as that for first fixation, was not significant, $t_1(17) = 1.30$, $p = .21$, $t_2(171) = 1.15$, $p = .25$.

Regressions in: The next measure of interest is the probability of regressing in to the target region from a later word in the sentence. Since both the target and the pre-target cue contain disambiguating information, this measure will provide a way to capture how much downstream processing was necessary after readers had moved on (see Table 5 for means). Analyses showed a main effect of context, $F_1(1,17) = 16.68$, $p < .01$, $F_2(1,170) = 8.43$, $p < .01$, as readers were more likely to regress back to the target region in the syntactic prose context. There was no significant effect of word type, F_1 $(1,17) = 2.76$, $p = .12$, $F_2(1,170) =$

[1]In order to confirm that the first fixation ambiguity effect was not driven by just a few outlier items, we examined how many ambiguous words showed increased first fixation durations in the syntactic prose relative to the congruent context (note that no individual item could show an *ambiguity* effect, because a target word could only be ambiguous or unambiguous). We found that 60% of the NV homographs showed this increase. Furthermore, when we removed the 25% of items at the upper and lower ends of the first fixation effect size range, leaving only those items in the middle 50%, we still observed a 13.8 ms average difference between the syntactic prose and congruent contexts for the ambiguous words (as opposed to the 14 ms difference when all items were included). These results provide further evidence that the first fixation effect is not driven by a small subset of items.

2.05, $p = .15$ (numerically, there were more regressions in to ambiguous than unambiguous words), and no interaction between context and word type, F_1 and $F_2 < 1$.

Rereading time: As an additional later measure of processing, rereading times on the target region were calculated by subtracting the gaze duration from the total time spent reading the region. As previously mentioned, these values include 0 ms rereading times for cases in which neither word in the region was refixated. Since the probability of refixating target words varied across conditions, and calculating rereading time assumes that a word will be refixated, only including cases in which a word was reread will overestimate the actual average rereading times for the words (see Rayner, Slattery, Dreighe, and Liversedge, 2011). Average rereading times are listed in Table 5. Analyses revealed a significant main effect of context, $F_1(1,17) = 20.35$, $p < .01$, $F_2(1,170) = 45.27$, $p < .001$, indicating that readers spent more time rereading the words in the syntactic prose context. There was also a significant effect of word type, $F_1(1,17) = 7.03$, $p < .05$, which was marginal by items, $F_2(1,170) = 3.28$, $p = .07$, indicating that readers spent more time rereading the ambiguous compared to unambiguous words overall. The interaction between the two factors was not significant, $F_1(1,17) = 2.07$, $p = .17$, $F_2(1,170) = 1.11$, $p = .29$.

Discussion

In Experiment 1, we found increased reading times, specifically first fixation durations, for ambiguous target words and the preceding word-class cues compared to unambiguous target regions, providing converging evidence with previous ERP studies (Lee and Federmeier, 2009, 2011) showing ambiguity costs to NV homographs in this condition. As was also true for the ERP ambiguity cost, the ambiguity cost on first fixation durations was not seen when the same words appeared in semantically congruent sentences. Not only the pattern, but also the timecourse of the two effects align. Across studies, the ERP frontal negativity tends to begin prior to 200 ms post-stimulus-onset (Federmeier et al., 2000; Lee & Federmeier, 2006, 2009). First fixation durations in this study were in the 200–250 ms range. It is therefore plausible that the neural processes reflected in the ERP frontal negativity and the processes responsible for delaying saccades off of the target region could share a common source.

Rereading times were longer for ambiguous than unambiguous words across context types, suggesting that when readers have control over their intake of information (different from the corresponding ERP experiment), they tend to spend more time when they go back to ambiguous words, perhaps reflecting cases in which these items were not fully disambiguated during the first pass. In addition, later gaze measures consistently reveal costs associated with reading in the syntactic prose sentences as compared with the congruent sentences, with longer gaze durations, more regressions in, and longer rereading times for syntactic prose sentences, regardless of whether or not they contain an ambiguous or unambiguous target word. As the syntactic prose sentences are semantically incoherent, these costs are not surprising, and, indeed, are analogous with patterns seen in the ERP data. Lee and Federmeier (2009) observed a general reduction in N400 amplitude (250–500 ms) in the congruent context compared to the syntactic prose context for both ambiguous and unambiguous words (cf., Van Petten & Kutas, 1990). The reduction in N400 amplitude likely reflects the build-up of message-level semantics in the congruent – but not syntactic prose – contexts, which eased word processing.

Although the later eye gaze measures mirrored the general pattern of context effects on the N400, additional effects of ambiguity on the N400 in Lee & Federmeier (2009) were not directly found in eye movement patterns in the present study. Lee and Federmeier (2009) found that in the congruent context, the N400 elicited by the unambiguous words showed a greater reduction than that seen to the ambiguous words, even though they were matched for

cloze probability (i.e., predictability). In a second experiment, this pattern was shown to be specific to cases in which the context picked out the subordinate meaning of the ambiguous words, and it was hypothesized to reflect residual activation of dominant meaning features. In the present experiment, gaze durations were facilitated (not lengthened) for the ambiguous words in the congruent context. This finding could add support to Lee and Federmeier's (2009) hypothesis that the N400 pattern indexes residual activation of the alternative meaning within semantic memory. Shorter gaze durations might be a consequence of this increased activation, which might allow readers to reach some threshold of activation faster for ambiguous words, causing their eyes to move forward in the text sooner than for unambiguous words (Reichle, Pollatsek, Rayner, 2006). Furthermore, within the two contexts, readers were equally likely to make a regression to the target word, regardless of whether it was ambiguous or unambiguous. The fact that neither gaze durations nor probability of regression in showed ambiguity effects like those seen on the N400 in congruent contexts is consonant with Lee and Federmeier's (2009) finding that in the congruent context, ambiguity is resolved without the need to recruit additional neural resources – and, thus, perhaps, without concomitant behavioral consequences.

Importantly, then, we did find eye gaze effects, specifically on first fixation durations, that matched those seen for the frontal negativity ERP effect. This pattern suggests that when syntactic, but not semantic, cues to disambiguation are available, there is a cost for processing ambiguous compared to unambiguous words. This cost is eliminated when semantic constraints augment syntactic ones. Thus, in contrast to previous eye tracking studies (Folk & Morris, 2003), the present results suggest that syntactic cues alone are insufficient to allow selective access of one meaning of an NV homograph.

Experiment 2: Older Adults

The results from the younger adults show that effects seen in the ERP findings linked to selection processes have parallels in first fixation times, both of which come online around 200 ms after first encountering an ambiguous word in the absence of coherent semantics. These parallel effects suggest that inhibitory processes used to suppress contextually inappropriate meanings may also slow the eyes during natural reading. To test this hypothesis, we examined the reading patterns of older adults using these same materials.

In collecting eye movement measures from older adults, we have the opportunity to examine how ambiguity resolution processes may change with age. To our knowledge, no prior study has investigated how older adults resolve the ambiguity associated with NV homographs in natural reading. Studies using other methods have found mixed results, but the overarching pattern seems to suggest that older adults are similar to younger adults in their ability to use semantic information to resolve ambiguities online (Balota & Duchek, 1991; Hopkins, Kellas, & Paul, 1995; Meyer & Federmeier, 2010; Swaab, Brown and Hagoort, 1998) but that they fail to use syntactic cues online when the context is semantically neutral (Dagerman, MacDonald, & Harm, 2006; Lee & Federmeier, 2009; 2012). Furthermore, it has been suggested that older adults have inhibition deficits (Hasher & Zacks, 1988), particularly in their ability to keep task-irrelevant information from entering into working memory and in suppressing incorrect behavioral responses to a task (Craik and Bialystok, 2008). Although inhibition deficit theory mainly refers to more controlled tasks, it is possible that general inhibition deficits in older adults could limit their ability to suppress the context-irrelevant meaning of the NV homographs. In order to test this possibility, we collected individual difference measures of inhibition through the Hayling sentence completion task (Burgess & Shallice, 1996) to determine whether there is an association between the ability to inhibit predictable responses to sentences and the ability to do on-line meaning suppression in the service of ambiguity resolution.

Although older adults may have more difficulty using syntactic cues on-line and/or suppressing context-irrelevant meanings, it is the case that, as already discussed, natural reading affords additional strategies, such as rereading. Indeed, there is evidence that older adults strategically allocate their processing resources over text differently than young adults do (Stine-Morrow et al., 2006). Older adults sometimes make longer fixations on individual words (Kliegl, Grabner, Rolfs, & Engbert, 2004) and have been shown to have smaller perceptual spans than young adult readers (Rayner, Castelhano, & Yang, 2009). They also have been found to skip words more frequently than younger adults. Although word skipping is typically thought to indicate that the skipped word was processed on the reader's previous fixation (Rayner, 1998), this pattern in older adults has been described as "risky" reading (Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006), as older adults also make more regressions back in text (Kemper, Crow, & Kemtes, 04), suggesting they may not have fully processed the skipped words. There is also evidence that older adults take more advantage of rereading to resolve ambiguity (Shake & Stine-Morrow, 2011). Thus, if older adults have difficulty engaging top-down resources, such as selection mechanisms, in a timely fashion, they may adopt a strategy of delaying lexical ambiguity resolution until more information is available or after they go back to critical parts of the text. Thus, for older adults, effects of lexical ambiguity may manifest on different eye gaze measures, such as rereading times.

Lee and Federmeier (2011) found that the frontal negativity ERP effect was absent in older adults as a group, although it could be seen in older adults with high verbal fluency. If the two effects do indeed share a source, then we would expect the first fixation effects to similarly be reduced or absent in older adults as a group -- but subject to individual differences related to verbal fluency.

Methods

Participants—Eighteen older adults (5 men; mean age 69.5 years; range 62–83 years) participated in the eye-tracking experiment for cash payment of $8 per hour. All participants were right-handed as assessed by the Edinburgh inventory (Oldfield, 1971). All were also monolingual speakers of English, with no consistent exposure to other languages before age 5. Participants had no history of neurological/psychiatric disorders or brain damage. We conducted the Montreal Cognitive Assessment (MoCA) to screen participants for cognitive impairments. Participants scored an average of 27.1 out of 30 possible points (SD = 1.88; range 23–30), which falls within the normal range for this test (suggested cutoff for impairment < 23; Luis, Keegan, & Mullan, 2009).

We also assessed older adults using the same battery of neuropsychological tests as administered to the young adults (see Table 2 for performance measures). Individual difference measures are available for all but two of the older adults subjects, who were unable to complete the battery due to time. Furthermore, one older adult was excluded from the analysis involving the Hayling: anomalous completions task, for refusal to comply with task demands during test administration. To assess age-related effects on the individual difference measures, we performed a one-way ANOVA, with the between-subjects factor of age, for each of the measures collected. The older adults had a significantly higher average d-prime score on the magazine recognition test, $F(1,32) = 6.45$, $p < .05$, as well as more correct responses on Hayling: congruent completions test, $F(1,32) = 6.98$, $p < .05$. For all other individual difference measures, performance was well-matched between the two groups (all F's < 1).

Materials, Procedure, and Data Analysis—Experimental design, stimulus materials, and analytic approach were the same as in the previous study. Approximately 7.4% of trials were eliminated from the analyses due to track loss.

Results

Behavioral Results

Word recognition task: Overall accuracy on the word recognition task for older adults was 92%, and memory performance in each condition was again analyzed using the d' index (see Table 3 for means). Results showed a main effect of context, $F(1,17) = 39.98$, $p < .01$, with better word recognition in congruent sentences than in syntactic prose, but no main effect of word type nor interaction between the two factors (F's < 1). This level of performance is slightly lower than the 98% accuracy seen in Lee and Federmeier's (2011) ERP experiment, although the pattern of effects is the same across experiments.

The older adults' performance was also numerically lower than that of the young adults. To test for significant differences between the groups, an omnibus ANOVA was run with the between-subjects factor of age (young and old), and within-subject factors of context (congruent vs. syntactic prose) and target word type (ambiguous vs. unambiguous). The test showed a significant main effect of context, $F(1,34) = 48.91$, $p < .01$, but no reliable effect of target word type ($F < 1$) or age $F(1,34) = 2.42$, $p = .13$. However, age significantly interacted with context, $F(1,34) = 6.65$, $p < .05$, indicating that, although both groups demonstrated poorer memory performance for words appearing in the syntactic prose context, this deficit was larger for the older adults.

Sentence recognition task: Overall accuracy in the sentence recognition task for older adults was 82%. Analyses of d' scores (listed in Table 3) showed a main effect of context, $F(1,17) = 7.24$, $p < .05$, with reduced memory for syntactic prose than for congruent sentences. There was no main effect of word type nor interaction between the two factors (Fs < 1). Overall accuracy was again slightly lower than the 89% observed by Lee and Federmeier (2011), although, as previously mentioned, the ERP version contained half the number of sentences overall.

As with the word recognition task, older adults' performance on the sentence recognition task was slightly lower than that of the young adults in Experiment 1. To directly compare performance, an omnibus ANOVA was run with the between-subjects factor of age (young and old), and within-subject factors of context (congruent vs. syntactic prose) and target word type (ambiguous vs. unambiguous). The test showed a significant main effect of context, $F(1,34) = 63.51$, $p < .01$, but no reliable effect of target word type or age (Fs < 1). However, age showed a marginal interaction with context, $F(1,34) = 2.92$, $p = .097$, indicating again that older adults were more affected by the difficulty of the syntactic prose sentences than were younger adults.

Summary: Like younger adults, older adults attended to and were able to remember both word level and message level information. As was true for younger adults, older adults' memory performance was significantly better for both single words and entire sentences appearing in the congruent contexts. Additionally, although there were no significant main effects of age, interactions between age and context type suggest that the lack of coherent semantics is more detrimental to older adults' ability to remember words and sentences, after both short and longer delays. This interaction was not present in the ERP study, perhaps because the current study had double the number of sentences, with both critical and filler sentences in every trial; this may have taxed older adults' memory, thus magnifying the memory deficits caused by the lack of coherent semantics. It could also be that when

given the choice about how to allocate time and attention across the words in the sentence types, older adults use strategies less conducive to their future memory in the syntactic prose condition.

Eye-Tracking Results

First fixation: Like the younger adults, the older adults skipped the pre-target word-class cue 50.4% of the time, creating a situation in which their first apprehension of the target word is distributed almost evenly between fixations to the pre-target cue and fixations to the target word. As such, first fixations for the older adults were also examined for the target region (see Table 4 for means). Results revealed a main effect of context, $F_1(1,17) = 6.75$, $p < .05$, $F_2(1,170) = 5.47$, $p < .05$, but no effect of word type or interaction (all $Fs < 1$). These results indicate that first fixation durations were longer in the syntactic prose context, but importantly, that they did not differ between ambiguous and unambiguous words. Similar analyses conducted on the target and pre-target cues separately produced very similar results, with longer first fixations in the syntax-only contexts but no main effects or interactions with word type.

To directly compare the age groups, two omnibus ANOVAs were conducted. For the by-subjects analysis, there was a between-subjects factor of age (young and old), and within-subject factors of context (congruent vs. syntactic prose) and target word type (ambiguous vs. unambiguous). For the by-items analysis, there was a between-items factor of word type, and within-items factors of age and word type. Results revealed a main effect of context, $F_1(1,34) = 11.80$, $p < .01$, $F_2(1,170) = 9.73$, $p < .01$, indicating that first fixation durations in both groups were longer in the syntactic prose than congruent context. There was no main effect of word type (F_1 and $F_2 < 1$). Context and word type showed a marginal interaction by subjects, $F_1(1,34) = 3.75$, $p = .06$, that was not significant by items, $F_2(1,170) = 1.34$, $p = .25$. This trend could reflect the larger ambiguity effect in the syntactic prose condition present in the young adults but not in the older adults. There was a significant main effect of age, $F_1(1,34) = 8.78$, $p < .01$, $F_2(1,170) = 103.48$, $p < .001$, with longer fixation times overall in older adults. Age did not interact with context, $F_1(1,34) = 1.03$, $p = .32$, $F_2 < 1$, or word type, (F_1 and $F_2 < 1$), nor was there a significant three-way interaction, $F_1(1,34) = 2.22$, $p = .15$, $F_2 < 1$ (the numerical trend reflects the pattern difference seen across the within group analyses). Effects of age on first fixation patterns will be further examined below, in the context of individual differences.

Gaze duration: As with the young adults, gaze durations were examined for the target region (see Table 5 for means). Results revealed a main effect of context, $F_1(1,17) = 33.91$, $p < .01$, $F_2(1,170) = 45.29$, $p < .001$, with longer gaze durations in syntactic prose sentences. There was no main effect of word type, $F_1(1,17) = 2.41$, $p = .14$, $F_2 < 1$, nor an interaction between the two factors (F_1 and $F_2 < 1$). When gaze durations were compared to those with younger adults, the analyses revealed a main effect of context, $F_1(1,34) = 84.27$, $p < .001$, $F_2(1,170) = 80.46$, $p < .001$, indicating that both age groups had longer gaze durations on target words in the syntactic prose compared to congruent context. There was no main effect of word type, $F_1(1,34) = 2.73$, $p = .11$, $F_2 < 1$, although there was a significant interaction between context and word type, $F_1(1,34) = 8.82$, $p < .01$, that was marginal by items, $F_2(1,170) = 2.89$, $p = .09$. When collapsing across age, gaze durations in the congruent context were longer for unambiguous relative to ambiguous words, whereas gaze durations in the syntactic prose context showed the opposite pattern. Furthermore, there was a main effect of age, $F_1(1,34) = 9.67$, $p < .01$, $F_2(1,34) = 178.37$, $p < .001$ indicating that older adults' gaze durations were generally longer than those of young adults. Age did not interact with context or word type (all Fs < 1), but there was a marginal three-way interaction by subjects, $F_1(1,34) = 3.29$, $p = .08$, that was not reliable by items, $F_2(1,170) = 2.41$, $p = .12$.

This marginal interaction reflects the fact that older adults' gaze durations were numerically longer for unambiguous relative to ambiguous words in both contexts, whereas for young adults, this pattern was present in the congruent context but reversed in the syntactic prose context.

Regressions in: The probability of regressions in to the target region is shown in Table 5. Analyses showed a main effect of context, $F_1(1,17) = 9.95$, $p < .01$, $F_2(1,170) = 15.31$, $p < .001$ but no main effect of word type (F_1 and $F_2 < 1$) or interaction between the two factors, $F_1(1,17) = 2.15$, $p = .16$, $F_2 < 1$. Older adults were more likely to regress in to target words in the syntactic prose relative to congruent context, but, as for younger adults, this effect was not modulated by word type. When directly compared with the younger adults, there was an overall main effect of context, $F_1(1,34) = 21.97$, $p < .001$, $F_2(1,170) = 17.99$, $p < .001$, indicating that both age groups made more regressions back to target words in the syntactic prose context. There was no main effect of word type, $F_1(1,34) = 1.63$, $p = .21$, $F_2(1,170) = 1.54$, $p = .22$, nor did context and word type significantly interact, $F_1(1,34) = 1.87$, $p = .18$, $F_2 < 1$. There were no significant interactions with age (all Fs < 1), although there was a marginal main effect of age by subjects that was significant by items, $F_1(1,34) = 3.88$, $p = .06$, $F_2(1,170) = 22.33$, $p < .001$, showing that older adults were generally more likely to make a regression than young adults.

Rereading time: Finally, analyses of rereading times (means in Table 5) for the target region revealed a significant main effect of context, $F_1(1,17) = 12.82$, $p < .01$, $F_2(1,170) = 84.87$, $p < .001$, a main effect of word type, $F_1(1,17) = 11.86$, $p < .01$, $F_2(1,170) = 4.07$, $p < .05$, as well as an interaction, $F_1(1,17) = 9.87$, $p < .01$, $F_2(1,170) = 3.98$, $p < .05$. Follow-up comparisons show that the interaction is driven by the significant 70 ms difference in rereading times between the ambiguous and unambiguous targets in the syntactic prose condition[2], $t_1(17) = 3.62$, $p < .01$, $t_2(171) = 2.26$, $p < .05$, whereas the 2 ms difference between word types in the congruent context is not reliable, $t_1(17) = .20$, $p = .84$, $t_2(171) = .08$, $p = .93$. In the between-groups analysis with the younger adults, there was a main effect of context, $F_1(1,34) = 22.46$, $p < .001$, $F_2(1,170) = 101.83$, $p < .001$, indicating that all participants spent more time rereading target words in the syntactic prose context. There was also a main effect of word type, $F_1(1,34) = 18.61$, $p < .001$, $F_2(1,170) = 5.52$, $p < .05$, showing that across contexts, participants spent more time rereading the ambiguous words than the unambiguous words. Additionally, there was a significant interaction between context and word type in the by-subjects analysis, $F_1(1,34) = 3.99$, $p = .05$, that was not reliable by-items, $F_2(1,170) = 1.22$, $p = .27$, indicating that the differential rereading of ambiguous words was larger in the syntactic prose than congruent context. Furthermore, age significantly interacted with context, $F_1(1,34) = 4.46$, $p < .05$, $F_2(1,170) = 31.89$, $p < .001$, which demonstrated that although both groups increased their rereading times in the syntactic prose context, this increase was larger for older than younger adults. Age did not interact with word type, $F_1(1,34) = 2.63$, $p = .11$, $F_2 < 1$, but there was a significant three-way interaction between context, word type, and age, $F_1(1,34) = 11.82$, $p < .01$, $F_2(1,170) = 6.86$, $p < .05$. This effect indicates that the interaction between context and word type

[2]Although the current dataset was not designed to test the effect of dominance, a post-hoc analysis was performed to determine whether the rereading time effect for older adults was only driven by instances in which the NV homograph was used its non-dominant sense. To calculate dominance, each homograph was given a bias rating, by comparing the frequency difference between the word's noun and verb usages to the sum of these frequencies. If the difference was more the 30% of their summed frequency, the word was considered biased; if not, it was considered balanced. By this scoring criteria, 40% of the homographs used were balanced and 60% were biased. Results revealed that rereading times on homographs used in their dominant sense (285 ms) were roughly equivalent to those on homographs used in their non-dominant sense (i.e., all subordinate and balanced words; 291 ms), both of which were longer than rereading times on unambiguous words (219 ms). Since reading times on both groups of ambiguous words were substantially longer than those on the unambiguous words, we can conclude that usage-specific dominance did not have an effect on rereading time effects in older adults.

manifested differently in the two age groups. Older adults showed a large effect of context, as well as a significant ambiguity effect within the syntactic prose condition. Young adults also exhibited a context effect, albeit more modest than in the older adults, and they only showed a numeric ambiguity effect in the syntactic prose context.

Therefore, like the younger adults, older adults do more rereading in syntactic prose contexts and also of ambiguous words. However, different from the younger adults, the older adults show a pattern in which they selectively spend more time rereading ambiguous words than unambiguous words in the syntactic prose contexts -- a pattern that is eliminated in the congruent sentences. Thus, the pattern observed on the frontal negativity and in first fixation times for the younger adults is seen instead in rereading times for older adults as a group.

Combined Analysis: The Role of Individual Differences—In the ERP study using the same materials, Lee and Federmeier (2011) found that among older adults there was a significant relationship between the size of the frontal negativity effect and verbal fluency, such that higher verbal fluency scores were associated with a larger frontal negativity effect. Given the similarities of the frontal negativity and the first fixation effects in terms of timing and sensitivity to context and age, it seems likely that they are indexing the same underlying process. If this is the case, then older adults with high verbal fluency should exhibit the first fixation effect, whereas those with low verbal fluency would not. This individual difference might also hold for young adults; although Lee and Federmeier (2009) did not find this correlation in their young adult ERP data, it is possible that this was in part because of the more restricted range of verbal fluency obtained in their young adult sample. In the present sample, verbal fluency scores were well matched between younger and older adults: young adults had a mean fluency of 70.2 with a range of 51–114 and older adults had a mean fluency of 64.2 with a range from 45–101. Fluency did not differ as a function of age, $t(32) = -1.26$, $p = .22$.

Therefore, to better understand the relationship between eye-movement patterns and fluency, eye movement measures were regressed against verbal fluency measures for both age groups together. To examine first fixation effects, a measure of ambiguity cost in the syntactic prose condition was calculated for all participants by subtracting first fixation durations to the target region for unambiguous from ambiguous words (producing a positive value when a participant had longer first fixations on ambiguous words). Overall fluency was marginally correlated with the ambiguity cost, $r = .30$, $p = .08$, and category fluency in particular was significantly correlated, $r = .34$, $p < .05$, with participants who were able to produce more words showing larger ambiguity costs (Figure 1). The other subcomponent of verbal fluency, FAS fluency, was not correlated with the ambiguity cost, $r = .17$, $p = .33$.

In order to explore whether the other individual difference measures were important predictors of the first fixation effects, we performed a multiple regression analysis including eight individual difference measures (category fluency, letter fluency, reading span: setsize, reading span: total, Hayling: correct congruent completions, Hayling: correct anomalous completions, author recognition d-prime score, and magazine recognition d-prime score). We used a backwards-stepping procedure with the exclusion criteria of $p > .1$, such that factors were entered into the model and then removed one at a time if their significance level was above .1 (starting with the highest *p*-value). This continued until only factors with a *p*-value of less than .1 remained in the model. This resulted in a model containing the factors letter fluency and Hayling: correct congruent completions. The overall model explained roughly 23% of the variance in the first fixation effect, $R^2 = .23$, $F(2,30) = 4.36$, $p = .02$. The individual factor of letter fluency reached significance, $\beta = .38$, $t = 2.32$, $p < .05$, as did the Hayling: correct congruent score, $\beta = -.35$, $t = -2.14$, $p < .05$. Interestingly, the Hayling correct congruent score reflects a subject's ability to produce words that correctly fit with

the semantic and grammatical category required by a sentence frame, which would seemingly measure many of the same underlying skills as the verbal fluency task.

Furthermore, since the ambiguity cost was expressed in the older adults as longer rereading times on the ambiguous words, we wanted to explore whether this cost was also correlated with verbal fluency, or whether it was simply an expression of group differences regardless of individual differences. A second measure of ambiguity cost was calculated for all participants by subtracting the rereading times on the unambiguous words from ambiguous words in the syntactic prose condition (again producing a positive value when a participant has longer rereading times for the ambiguous words). Overall fluency showed a marginal negative correlation with rereading cost, $r = -.29$, $p = .097$, and again, category fluency had a significant negative correlation, $r = -.34$, $p < .05$, with participants who produced fewer words showing larger ambiguity costs in rereading times (Figure 2). As before, the other subcomponent of verbal fluency, FAS fluency, was not correlated with the rereading ambiguity costs, $r = -.16$, $p = .37$.

In order to test for relationships between the rereading ambiguity cost and the other individual difference measures, a second multiple regression analysis was performed on the rereading cost, in the same manner as described above. This analysis produced a final model that contained only the category fluency score and explained roughly 12% of the variance in the rereading effect, $R^2 = .12$, $F(2,30) = 4.05$, $p = .05$. The individual factor of category fluency was significant, $\beta = -.34$, $t = -2.01$, $p = .05$. All other factors had p-values of larger than .1 and were not retained in the model.

To further explore the nature of the relationship between age, verbal fluency, and ambiguity effects, we divided the two age groups into high- and low-verbal fluency subgroups. The young adult high verbal fluency group produced 80.3 words on average ($SD = 13.6$; range 71–114), whereas the low verbal fluency group produced 60.1 words on average ($SD = 5.7$; range 52–68). The high verbal fluency older adults produced an average of 73.1 words ($SD = 12.4$; range 63–101), and the low verbal fluency older adults produce an average of 55.4 words ($SD = 5.4$; range 45–61). Reading time measures to the target region in the syntactic prose sentences could then be examined as a function of both age and verbal fluency group (see Table 6 for values). In order to assess how age and verbal fluency jointly affected the size and timing of the ambiguity effects, a repeated-measures ANOVA was conducted on the ambiguity effects for first fixation duration and rereading time (i.e. the difference between ambiguous and unambiguous reading measures on the target region in the syntactic prose context), with the within-subjects factor of timing (initial -- first fixation vs. late -- rereading time), and between-subjects factors of age (young vs. old) and verbal fluency group (high vs. low). Results showed a main effect of timing, $F(1,30) = 12.18$, $p < .01$, indicating that the ambiguity effects were larger for the rereading than first fixation measure. There was a significant interaction between timing and age, $F(1,30) = 14.53$, $p < .01$, showing that as a group, younger adults elicited larger initial effects than older adults, whereas the older adults elicited larger late effects than the young. Furthermore, there was a significant interaction between timing and verbal fluency group, $F(1,30) = 13.39$, $p < .01$, indicating that regardless of age, high fluency readers elicited larger initial effects, whereas low fluency readers elicited larger later effects. The three-way interaction between timing, age and verbal fluency group was not significant, $F(1,30) = 1.35$, $p = .25$. The pattern observed across the four groups further demonstrates the trade-off between first fixation and downstream costs (see Figure 3), with both age and fluency contributing to how ambiguity affects fixations over time.

Discussion

As predicted, the eye-tracking results revealed that older adults do not exhibit inflated first fixation durations for ambiguous words in the syntax-only context. This finding strengthens the tie between the first fixation ambiguity effect seen in the young adults and the ERP frontal negativity that has been seen under the same conditions with these same stimuli (Lee & Federmeier, 2009), as both effects are absent in older adults as a group (Lee & Federmeier, 2011). Similar to the pattern seen for the young, older adults' gaze durations are shorter for words in congruent contexts as well as for ambiguous words overall, and they show a higher probability of regressing in to both word types in the syntactic prose context. Interestingly, and different from younger participants, older adults spend more time rereading ambiguous words in the syntactic prose (but not the congruent) contexts. The pattern seen on first fixations for young adults thus appears on a much later reading measure in older adults, suggesting that older adults are impaired at recruiting resources important for meaning selection during their initial reading, necessitating that they return to the ambiguous word for further processing.

Further evidence for this trade-off between first fixation and rereading effects can be seen in the pattern of individual differences. Lee and Federmeier (2011) found that although older adults as a group failed to elicit the frontal negativity effect, a subset of older adults with higher verbal fluency showed the young-like pattern. We regressed the size of the ambiguity effect on first fixations with verbal fluency, and found that -- for both age groups -- higher category fluency scores indeed predicted larger first fixation effects. At the same time, fluency was negatively correlated with ambiguity effects on rereading. Thus, at the individual level as well as at the group level (as a function of age and verbal fluency), readers exhibit a trade-off between slowing down during their first apprehension of ambiguous words in contexts demanding effortful meaning selection, or, instead, spending more time rereading those words after the first pass through the text.

General Discussion

The current study investigated how young and older adults use semantic and syntactic context information to resolve the ambiguity associated with noun/verb homographs (e.g. *park*) during natural reading. Overall results showed that when only syntactic information was available, young adults exhibited inflated first fixation durations on their first apprehension of the ambiguous words, whereas older adults as a group did not. The presence of coherent semantic information eliminated this cost for young participants. These findings parallel previous ERP work by Lee and Federmeier (2009, 2011), which found sustained frontal negativity elicited by the ambiguous words in the syntactic prose context for younger, but not older, adults -- an effect that was also eliminated by coherent semantics. The fact that the first fixation and frontal negativity effects arise in the same conditions and age groups and come online in the same time window, between 150–250 ms after first apprehending the target word, suggests that there is overlap in the process(es) that underlie these effects.

Lee and Federmeier (2009, 2011) posited that the frontal negativity indexes fronto-temporally mediated meaning selection processes that are necessary in the absence of coherent semantic information and that likely involve suppression of the word's context-inappropriate meaning. Given that the frontal negativity has been observed for NV homographs in multiple types of syntactically constraining but semantically neutral contexts, and multiple positions within the sentence, it seems to reflect processes that are routinely brought on-line to help effect semantic selection (and thus are not specific to unusual sentence types, such as the syntactic prose used here, or to wrap-up processes for words in sentence-final position). To investigate more specifically how the frontal negativity effect

impacts downstream processing in order to better understand its functional role, Lee and Federmeier (2012) embedded NV homographs in semantically neutral but syntactically constraining sentences. The homographs were immediately followed by a prepositional phrase whose head noun was more plausible for one interpretation of the homograph than the other. For example, the sentences, "Ben tried the duck in the **dish** prepared by the famous chef" and, "Ben tried to duck in the **alley** to avoid the paparazzi" are plausible, whereas the sentences, "Ben tried to duck in the **dish** prepared by the famous chef" and "Ben tried the duck in the **alley** to avoid the paparazzi" are much less plausible. ERPs measured at the head noun can thus serve as a probe of the outcome of ambiguity resolution. Young adults elicited a frontal negativity that began with the onset of the NV homograph and was sustained up to the onset of the head noun of the prepositional phrase, at which point they showed a clear N400 plausibility effect (i.e., reduced N400 for plausible relative to implausible nouns). The young adults showed this N400 plausibility effect regardless of whether the context picked out the dominant or subordinate sense of the word. The fact that there was only downstream activation for nouns that were plausible for the context-appropriate meaning of the target word suggests that the frontal negativity indexed the young adults' ability to successfully suppress the context-inappropriate meaning. We postulate that this suppression also causes the first fixation effects in the present study by generating an inhibition signal that slows the eyes when readers first apprehend those words.

One prior eye-tracking study came to different conclusions about the impact of syntactic context information on lexical ambiguity resolution during natural reading. Folk and Morris (2003) found that in semantically neutral but syntactically constraining sentences, gaze durations and spillover times did not differ between NV homographs and their unambiguous controls, which they cited as evidence that syntactic information is sufficient to resolve the ambiguity associated with NV homographs. An important difference between that study and the present one is that Folk and Morris' materials always picked out the noun meaning of the NV homographs. It is thus possible that this regularity, rather than the syntactic information in the sentences themselves, allowed readers to avoid the need to engage selection mechanisms that would be necessary when such experimental constraints were not available. More generally, however, the studies cannot actually be compared directly, as Folk and Morris (2003) did not report first fixation durations, which is where the effects were present in our young adult population, nor did they separate instances in which readers first fixated the function word *the*, which preceded all of their ambiguous words, as in the present study. That neither study found costs on gaze durations to the NV homographs might indicate that the these effects are quickly resolved or that, by the time readers refixate the word, other processes exert more important influences over the eye movement program. What is clear is that the present data, in conjunction with behavioral and ERP evidence, are inconsistent with the claim that syntactic constraints alone eliminate ambiguity effects for NV-homographs.

In ERP studies, older adults as a group fail to elicit the frontal negativity when they encounter NV homographs with clearly disambiguating syntactic, but not semantic, constraints (Lee & Federmeier, 2011; 2012). Consistent with the idea that this effect indexes the engagement of selection processes, older adults correspondingly show downstream plausibility effects (on the head noun of the subsequent prepositional phrase) only for nouns consistent with the dominant meaning of an ambiguous word (Lee & Federmeier, 2012). Thus, it seems they selected a meaning based on dominance rather than utilizing the syntactic cues. Furthermore, the size of the frontal negativity correlates with the size of the plausibility effect on an individual basis, such that more frontal negativity is associated with more selective activation of the contextually appropriate meaning of the NV-homograph.

We therefore predicted that if the first fixation effects seen in the young participants in the present study index processes that are overlapping with the frontal negativity, then older

adults as a group should fail to elicit this effect. Further strengthening the connection between the first fixation effect and the frontal negativity, these predictions were confirmed: older adults as a group did not show lengthened first fixation durations to ambiguous compared to unambiguous words in the syntactic prose sentences. In ERPs, when presentation rates are fixed and there is no opportunity to look back, a failure to disambiguate the NV-homograph when it is encountered is likely to lead to downstream comprehension problems (as in Lee & Federmeier, 2012) that may be difficult to overcome. However, in natural reading, comprehenders are free to move their eyes back in the text to areas that may have caused difficulties, thus affording the use of different strategies for disambiguation. In fact, the older adults as a group spent more time rereading the ambiguous words in the syntax-only context, an effect not present in the young adults. Thus, in natural reading, when initial meaning selection is unsuccessful, ambiguity resolution may instead be effected when readers return to those words to engage in further processing.

A recent eye-tracking study by Shake and Stine-Morrow (2011) also found a similar tradeoff between initial effects in young adults and later effects in older adults. In sentences containing a role noun with a strong gender bias (e.g. firefighter, electrician), young adults showed initial costs on gaze duration when they encountered a reflexive pronoun that violated their gender expectations (i.e. *herself*). Older adults, on the other hand, were more likely to launch a regression out of the region and/or spend more time rereading the pronoun. As in the current dataset, older adults as a group relied on later reprocessing strategies to resolve ambiguities that younger adults appear able to resolve when they initially encounter the word.

Beyond differences across the age groups, we found that individual differences played an important role in determining reading behaviors. Motivated by Lee and Federmeier's (2011) finding that older adults with high verbal fluency showed a young-like pattern of frontal negativity, we examined correlations between verbal fluency and ambiguity effects in gaze measures. We found that high verbal fluency readers showed larger first fixation ambiguity effects than low verbal fluency readers. Verbal fluency has been linked to the integrity of frontal cortical areas and, more specifically, to the efficacy of fronto-temporal circuits (Stuss & Levine, 2002). Verbal fluency may therefore reflect the speed and/or efficacy of the transmission of incoming sensory information to frontal areas important for selection and eye movement control, making it a useful predictor of the first fixation ambiguity effects. Interestingly, the correlation between ambiguity effects and scores on the Hayling: correct anomalous completions test, a measure of response inhibition, was not significant. Zacks and Hasher (1997) have noted that the type of inhibition involved in lexical meaning selection, which happens outside of conscious awareness, is quite different from that described by the inhibition deficit theory, which refers to controlled selective attention processes usually engaged in during dual-task situations. The lack of correlation observed between the more effortful inhibition indexed by the Hayling test and the selection-related ambiguity effects in the present data provide some support for the idea that these two processes may have some non-overlapping properties. It is also the case that the verbal fluency task partially measures response inhibition (i.e., participants must inhibit word repetition, as well as the production of words that fall outside of the appropriate category) in addition to frontal lobe function, and so likely indexes more of the underlying processes necessary for recruitment of the frontally-mediated selection mechanisms.

We also found an inverse relationship between verbal fluency and rereading times, such that readers with low verbal fluency spend more time rereading the ambiguous words in the syntactic prose context. This correlation suggests that the low fluency readers spend more time rereading the ambiguous words, presumably in order to resolve meaning ambiguity that was not resolved on their first pass through the sentence. Taken in tandem, these patterns

demonstrate an interesting trade-off between these two effects: high fluency readers tend to show larger initial, and smaller later, effects, whereas low fluency readers show smaller initial, but larger later, effects.

The trade-off is most clearly seen in the two most extreme groups: the high verbal fluency young adults and the low verbal fluency older adults. The high fluency young exhibit clear first fixation effects (as seen in Figure 4), suggesting that they can engage their selection mechanisms both quickly and effectively. They are also the only group who show no further processing costs on the ambiguous words. This pattern of early, but not late, effects suggests that the initial meaning selection process was successful for these readers. In contrast, the low fluency old show no tendency for any first fixation costs, suggesting they do not initially process the ambiguous words differently -- and thus may not select a meaning or may select based on lexical frequency/dominance rather than context. Instead, they spend significantly more time rereading the ambiguous target region. Thus, we see that readers either slow down initially to select a meaning or, instead, reread the word later to resolve the ambiguity.

On the other hand, the low fluency young and high fluency older adults exhibit remarkably similar patterns to each other for every measure analyzed, and these patterns are intermediate in comparison to the two extreme groups already discussed. Both groups show a small tendency toward first fixation effects on the target region, as seen in Figure 4 (probably explaining why there was no interaction with age group even though the overall effect was significant in the young but not significant in the old). Interestingly, when fixations to the target and pre-target word class cue are examined separately (Table 6; Figure 5), these groups show a pattern of first fixation ambiguity effects for fixations to the pre-target cue but not to the target words (whereas the high fluency young show the pattern in both cases and the low fluency older adults in neither). Considering the length of the cue words (typically 2–3 letters), these readers were very likely able to apprehend the target word simultaneously with the word class cue in the cases where they exhibited the effect, since the range of word identification extends 7–8 characters to the right of fixation (Rayner, 1998). This pattern suggests that these two groups were only able to recognize the ambiguity and bring their selection mechanisms online when they received both the syntactic cue and the NV-homograph together, making clear the context-appropriate sense. In contrast, when readers skip the pre-target cue, they either have to derive information about the cue word from the left of the center of gaze, or integrate information about the word class cue that was gleaned from a previous fixation. Only high fluency young participants showed effects on first fixations in this case.

In terms of their downstream processing costs, both the low fluency young and the high fluency older adult groups show elevated rereading times to ambiguous words, although their rereading effects are notably smaller than those for the low fluency older adults. Interestingly, as can be seen in Figure 6, these are the only groups to show a pattern of elevated regressions in to the ambiguous, compared to unambiguous, target region. Thus, these groups show both early and late effects, suggesting that they are less effective at resolving ambiguity during first pass reading (perhaps in part because of restrictions on the conditions in which they show first fixation effects) and, correspondingly, are more likely to need to return to and spend time rereading the ambiguous words. Although these group patterns could only be explored on the surface in the current data set, they point to interesting differences in reading styles/strategies that are affected by age and the availability of cognitive resources.

In summary, the current study found parallel effects when materials previously examined during word-by-word reading with ERP measures were used in natural reading with eye

gaze measures. When young adults encounter NV homographs in contexts containing syntactic, but not semantic, cues for disambiguation, they show early-onset ambiguity effects, in the form of longer first fixations and frontally-distributed ERP activity, linked to the recruitment of frontally-mediated, inhibition-based meaning selection mechanisms. Older adults as a group do not exhibit either effect. In ERP studies, this is known to lead to downstream comprehension problems (Lee & Federmeier, 2012). However, during natural reading, older adults spend more time rereading the ambiguous words in the same context, likely to compensate for their failure to initially recruit these mechanisms and select a meaning. These differences are not driven only by age, however, as effects were modulated by verbal fluency in both age groups and with both measures. In gaze measures, high fluency was correlated with an increased tendency to show first fixation effects and, correspondingly, a reduced tendency to show rereading effects. Taken together, the results across methods suggest age-related decline in the efficacy of top-down meaning selection mechanisms, which higher verbal fluency can partially protect against. They also highlight the importance of comparing effects across modalities, by showing that although not all older adults exhibit young-like patterns of ambiguity resolution, those readers who are less successful at ambiguity resolution on the first pass will strategically revisit ambiguous words when given the opportunity to do so -- as during natural reading.

Acknowledgments

The authors would like to acknowledge C. Lee for her contributions to stimulus development and X. Gao for assistance using the eyetracker. This research was supported by NIH Predoctoral Training Grant T32-HD055272 and NSF Graduate Research Fellowship NSF DGE 11-44245 FLLW to MCS, and a James S. McDonnell Foundation Scholar Award and NIH grants AG026308 to KDF and AG013935 to EALS.

References

Balota DA, Duchek JM. Semantic priming effects, lexical repetition effects, and contextual disambiguation effects in healthy aged individuals and individuals with senile dementia of the Alzheimer type. Brain and Language. 1991; 40:181–201. [PubMed: 2036582]

Benton, AL.; Hamsher, K. Multilingual aphasia examination manual. Iowa City, IA: University of Iowa; 1978.

Binder KS. Sentential and discourse topic effects on lexical ambiguity processing: An eye movement examination. Memory & Cognition. 2003; 31(5):690–702. [PubMed: 12956234]

Burgess PW, Shallice T. Response suppression, initiation and strategy use following frontal lobe lesions. Neuropsychologia. 1996; 34:263–272. [PubMed: 8657357]

Burke, DM.; Harrold, RM. Automatic and effortful semantic processes in old age: Experimental and naturalistic approaches. In: Light, LL.; Burke, DM., editors. Language, memory, and aging. New York: Cambridge University Press; 1993. p. 100-116.

Burke DM, Peters L. Word associations in old age: Evidence for consistency in semantic encoding during adulthood. Psychology and Aging. 1986; 1:283–292. [PubMed: 3267408]

Craik, FIM.; Bialystok, E. Lifespan cognitive development: The roles of representation and control. In: Craik, FIM.; Salthouse, TA., editors. The Handbook of Aging and Cognition. 3rd ed.. New York: Psychology Press; 2008. p. 557-601.

Daneman M, Carpenter PA. Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior. 1980; 19:450–466.

Dagerman KS, MacDonald MC, Harm MW. Aging and the use of context in ambiguity resolution: Complex changes from simple slowing. Cognitive Science. 2006; 30:311–345. [PubMed: 21702817]

DiGirolamo GJ, Kramer AF, Barad V, Cepeda NJ, Weissman DH, Milham MP, et al. General and task-specific frontal lobe recruitment in older adults during executive processes: A fMRI investigation of task-switching. NeuroReport. 2001; 12:2065–2071. [PubMed: 11435947]

Duffy SA, Morris RK, Rayner K. Lexical ambiguity and fixation times in reading. Journal of Memory and Language. 1988; 27(4):429–446.

Federmeier KD, Kutas M. Aging in context: Age-related changes in context use during language comprehension. Psychophysiology. 2005; 42:133–141. [PubMed: 15787850]

Federmeier KD, Kutas M, Schul R. Age-related and individual differences in the use of prediction during language comprehension. Brain and Language. 2010; 115:149–161. [PubMed: 20728207]

Federmeier KD, McLennan DB, De Ochoa E, Kutas M. The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. Psychophysiology. 2002; 39:133–146. [PubMed: 12212662]

Federmeier KD, Segal JB, Lombrozo T, Kutas M. Brain responses to nouns, verbs and class-ambiguous words in context. Brain. 2000; 123:2552–2566. [PubMed: 11099456]

Folk JR, Morris RK. Effects of syntactic category assignment on lexical ambiguity resolution in reading: An eye movement analysis. Memory & Cognition. 2003; 31(1):87–99. [PubMed: 12699146]

Frazier L, Rayner K. Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. Journal of Memory and Language. 1987; 26:505–526.

Gunter TC, Wagner S, Friederici AD. Workng memory and lexical ambiguity resolution as revealed by ERPs: A difficult case for activation theories. Journal of Cognitive Neuroscience. 2003; 15(5): 643–657. [PubMed: 12965038]

Hasher, L.; Zacks, RT. Working memory, comprehension, and aging: A review and a new view. In: Bower, GH., editor. The psychology of learning and motivation. Vol. Vol. 22. New York: Academic Press; 1988. p. 193-225.

Henry JD, Crawford JR. A meta-analytic review of verbal fluency performance following focal cortical lesions. Neuropsychology. 2004; 18:284–295. [PubMed: 15099151]

Hopkins KA, Kellas G, Paul ST. Scope of word meaning activation during sentence processing by young and older adults. Experimental Aging Research. 1995; 21:123–142. [PubMed: 7628507]

Kemper S, Crow A, Kemtes K. Eye-fixation patterns of high- and low-span young and older adults: Down the garden path and back again. Psychology and Aging. 2004; 19:157–170. [PubMed: 15065939]

Kliegl R, Grabner E, Rolfs M, Engbert R. Length, frequency and predictability effects of words on eye movements in reading. Cognitive Psychology. 2004; 16(1/2):262–284.

Kucera, H.; Francis, WN. Computational analysis of present-day American English. Providence, RI: Brown University Press; 1967.

Kutas M, Federmeier KD. Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). Annual Review of Psychology. 2011 2012.

Lee CL, Federmeier KD. To mind the mind: An event-related potential study of word class and semantic ambiguity. Brain Research. 2006; 1081:191–202. [PubMed: 16516169]

Lee CL, Federmeier KD. Wave-ering: An ERP study of syntactic and semantic context effects of ambiguity resolution for noun/verb homographs. Journal of Memory and Language. 2009; 61:538–555. [PubMed: 20161361]

Lee CL, Federmeier KD. Differential age effects on lexical ambiguity resolution mechanisms. Psychophysiology. 2011; 48:960–972. [PubMed: 21175671]

Lee CL, Federmeier KD. Ambiguity's aftermath: How age differences in resolving lexical ambiguity affect subsequent comprehension. Neuropsychologia. 2012; 50:869–879. [PubMed: 22321956]

Luis CA, Keegan AP, Mullan M. Cross validation of the Montreal Cognitive Assessment in community dwelling older adults residing in the Southeastern US. International Journal of Geriatric Psychiatry. 2009; 24:197–201. [PubMed: 18850670]

MacDonald MC. The interaction of lexical and syntactic ambiguity. Journal of Memory and Language. 1993; 32:692–715.

Meyer AM, Federmeier KD. The effects of context, meaning frequency, and associative strength on semantic selection: Distinct contributions from each cerebral hemisphere. Brain Research. 2007; 1183:91–108. [PubMed: 17936727]

Meyer AM, Federmeier KD. Event-related potentials reveal the effects of aging on meaning selection and revision. Psychophysiology. 2010; 47:673–686. [PubMed: 20210876]

Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia. 1971; 9:97–113. [PubMed: 5146491]

Rayner K. Eye movements in reading and information processing: 20 years of research. Psychological Bulletin. 1998; 124(3):372–422. [PubMed: 9849112]

Rayner K, Castelhano MS, Yang J. Eye movements and the perceptual span in older and younger readers. Psychology and Aging. 2009; 24(3):755–760. [PubMed: 19739933]

Rayner K, Frazier L. Selection mechanisms in reading lexically ambiguous words. Journal of Experimental Psychology: Learning, Memory and Cognition. 1989; 15(5):779–790.

Rayner K, Pacht JM, Duffy SA. Effects of prior encounter and global discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. Journal of Memory and Language. 1994; 33:527–544.

Rayner K, Reichle ED, Stroud MJ, Williams CC, Pollatsek A. The effects of word frequency, word predictability and font difficulty on the eye movements of young and older readers. Psychology and Aging. 2006; 21(3):448–465. [PubMed: 16953709]

Rayner K, Slattery TJ, Drieghe D, Liversedge SP. Eye movements and word skipping during reading: Effects of word length and predictability. Journal of Experimental Psychology: Human Perception and Performance. 2011; 37(2):514–528. [PubMed: 21463086]

Raz N, Lindenberger U, Rodrigue KM, Kennedy KM, Head D, Williamson A, et al. Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. Cerebeal Cortex. 2005; 15:1676–1689.

Reichle ED, Pollatsek A, Rayner K. E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. Cognitive Systems Research. 2006; 7:4–22.

Reingold EM, Yang J, Rayner K. The time course of word frequency and case alternation effects on fixation times in reading: Evidence for lexical control of eye movements. Journal of Experimental Psychology: Human Perception and Performance. 2010; 36(6):1677–1683. [PubMed: 20731513]

Salthouse, TA. Major Issues in Cognitive Aging. New York: Oxford University Press; 2010. Relations between age and cognitive functioning.

Seidenberg MS, Tanenhaus MK, Leiman JM, Bienkowski M. Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. Cognitive Psychology. 1982; 14:489–537.

Shake MC, Stine-Morrow EAL. Age differences in resolving anaphoric expressions during reading. Aging, Neuropsychology and Cognition. 2011; 18(6):678–707.

Stine-Morrow EAL, Gagne DD, Miller LMS, Hertzog C. Self-regulated reading in adulthood. Psychology and Aging. 2008; 23(1):131–153. [PubMed: 18361662]

Stine-Morrow EAL, Noh SR, Shake MC. Age differences in the effects of conceptual integration training on resource allocation in sentence processing. Quarterly Journal of Experimental Psychology. 2010; 63(7):1430–1455.

Stine-Morrow EAL, Soederberg Miller LM, Hertzog C. Aging and self-regulated language processing. Psychological Bulletin. 2006; 132(4):582–606. [PubMed: 16822168]

Stuss DT, Levine B. Adult clinical neuropsychology: Lessons from studies of the frontal lobes. Annual Review of Psychology. 2002; 53:401–433.

Swaab T, Brown C, Hagoort P. Understanding words in sentence contexts: The time course of ambiguity resolution. Brain and Language. 2003; 86:326–343. [PubMed: 12921771]

Tanenhaus MK, Leiman JM, Seidenberg MS. Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. Journal of Verbal Learning and Verbal Behavior. 1979; 18:427–440.

Tucker-Drob EM. Global and domain-specific changes in cognition throughout adulthood. Developmental Psychology. 2011; 47:331–343. [PubMed: 21244145]

Tyler LK, Marslen-Wilson WD. The on-line effects of semantic context on syntactic processing. Journal of Verbal Learning and Verbal Behavior. 1977; 16:683–692.

Van Petten C, Kutas M. Interactions between sentence context and word frequency in event-related brain potentials. Memory & Cognition. 1990; 18(4):380–393. [PubMed: 2381317]

Zacks R, Hasher L. Cognitive gerontology and attentional inhibition: A reply to Burke and McDowd. Journal of Gerontology: Psychological Sciences. 1997; 52B(6):274–283.

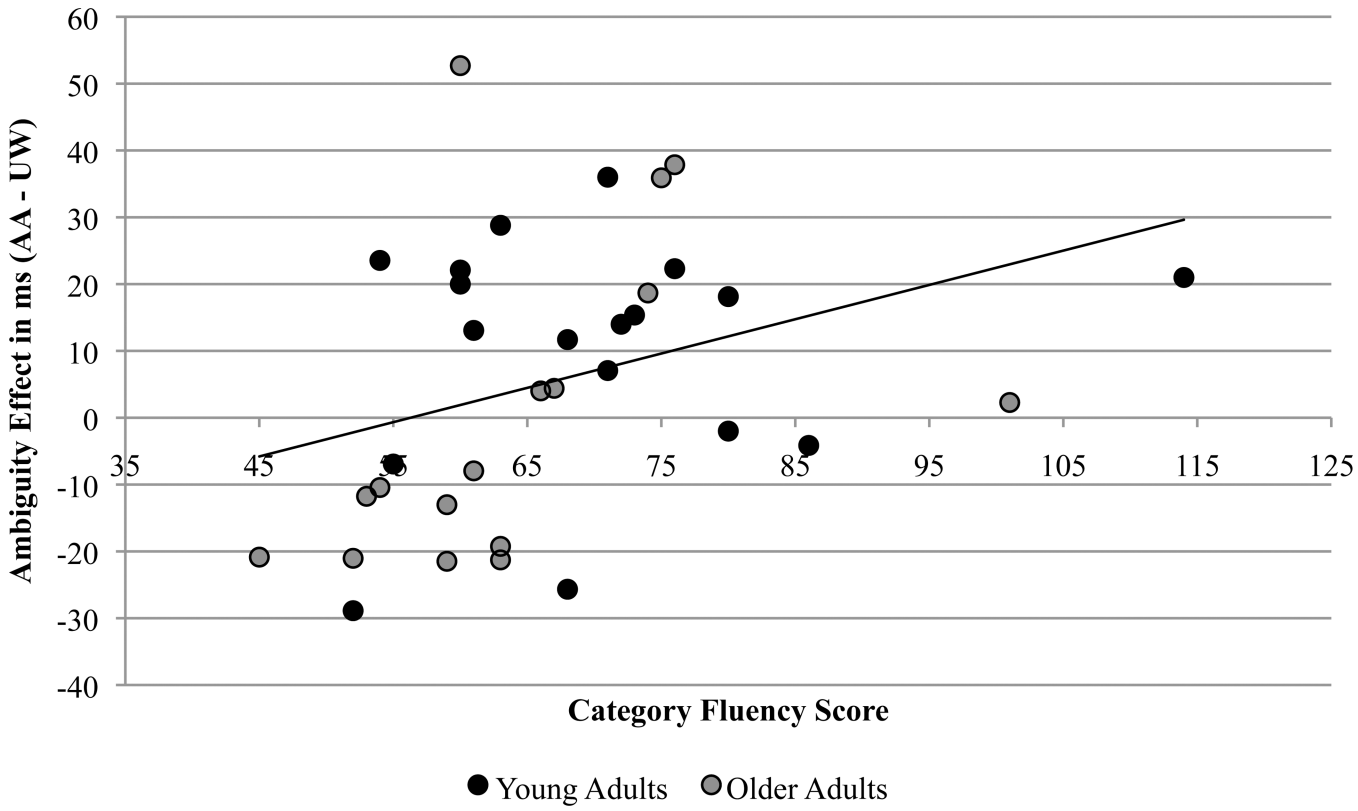

Figure 1.
Scatterplot of the correlation between category verbal fluency and first fixation ambiguity effects (measured by subtracting first fixations to unambiguous words from those to ambiguous words) for both young and older adults.

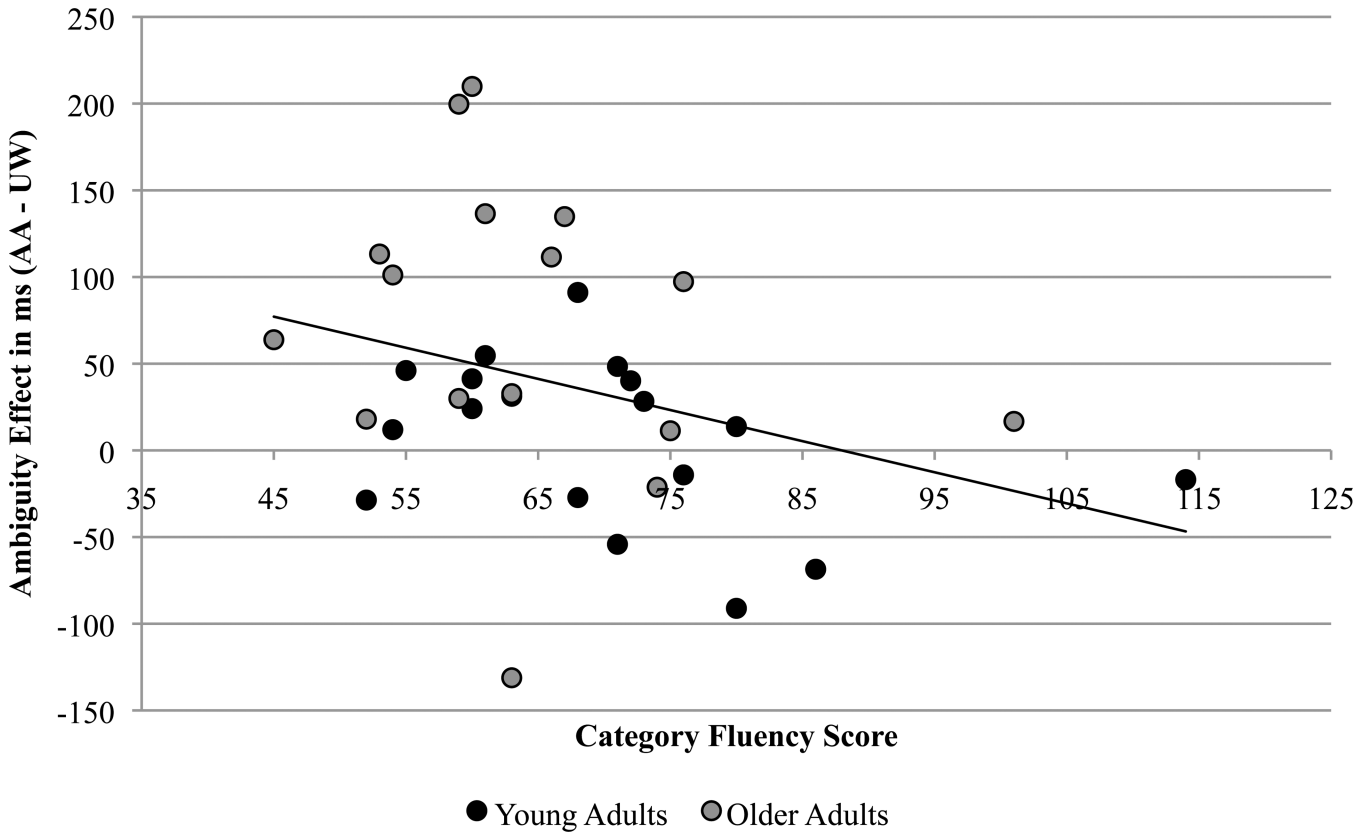

Figure 2.
Scatterplot of the correlation between category fluency and ambiguity effects on rereading times (calculated by subtracting rereading times on unambiguous words from those to ambiguous words) for young and older adults.

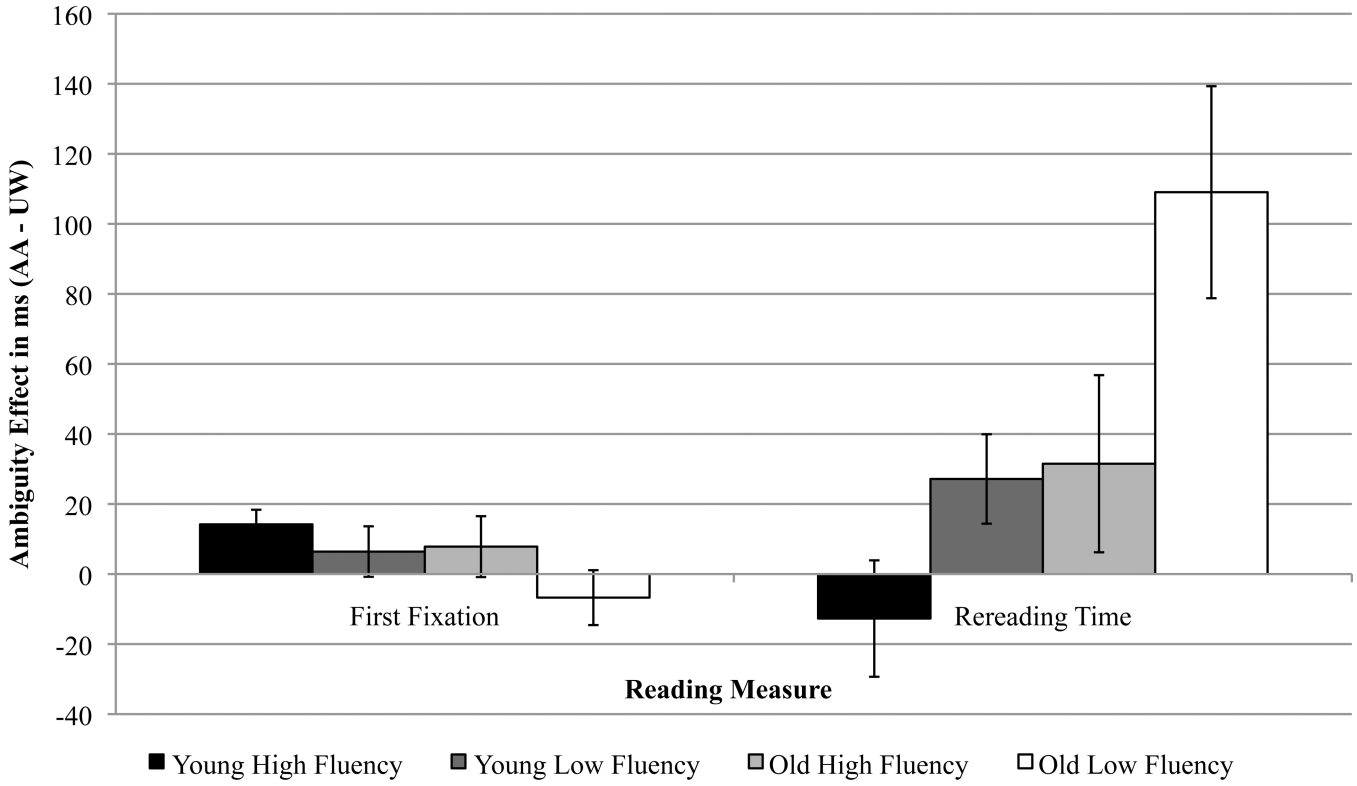

Figure 3.
Ambiguity effects in ms (calculated by subtracting reading times on unambiguous from ambiguous words) on first fixation duration and rereading time, plotted for each age and verbal fluency group separately.

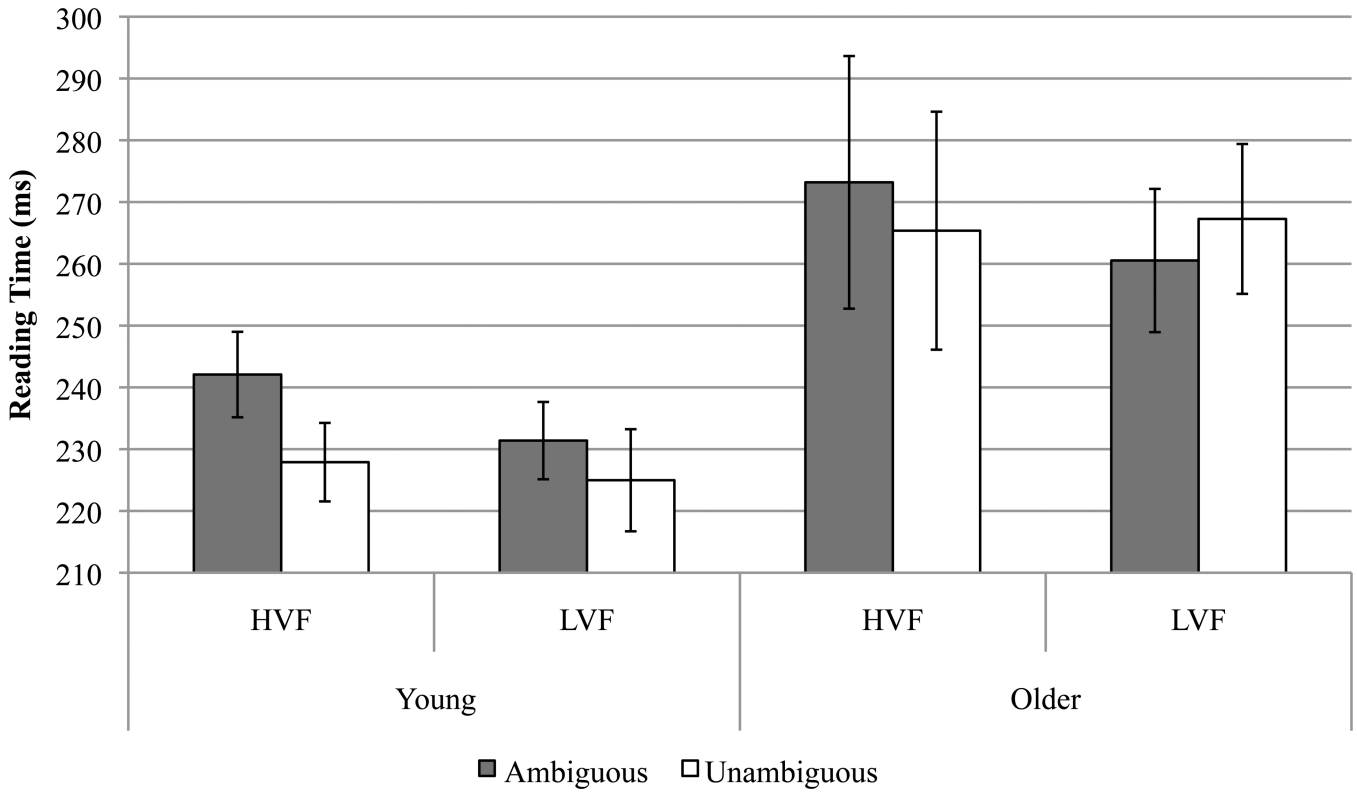

Figure 4.
First fixation durations on the target region, in the syntactic prose contexts only, as a function of both age and verbal fluency group. (HVF = High verbal fluency; LVF = Low verbal fluency)

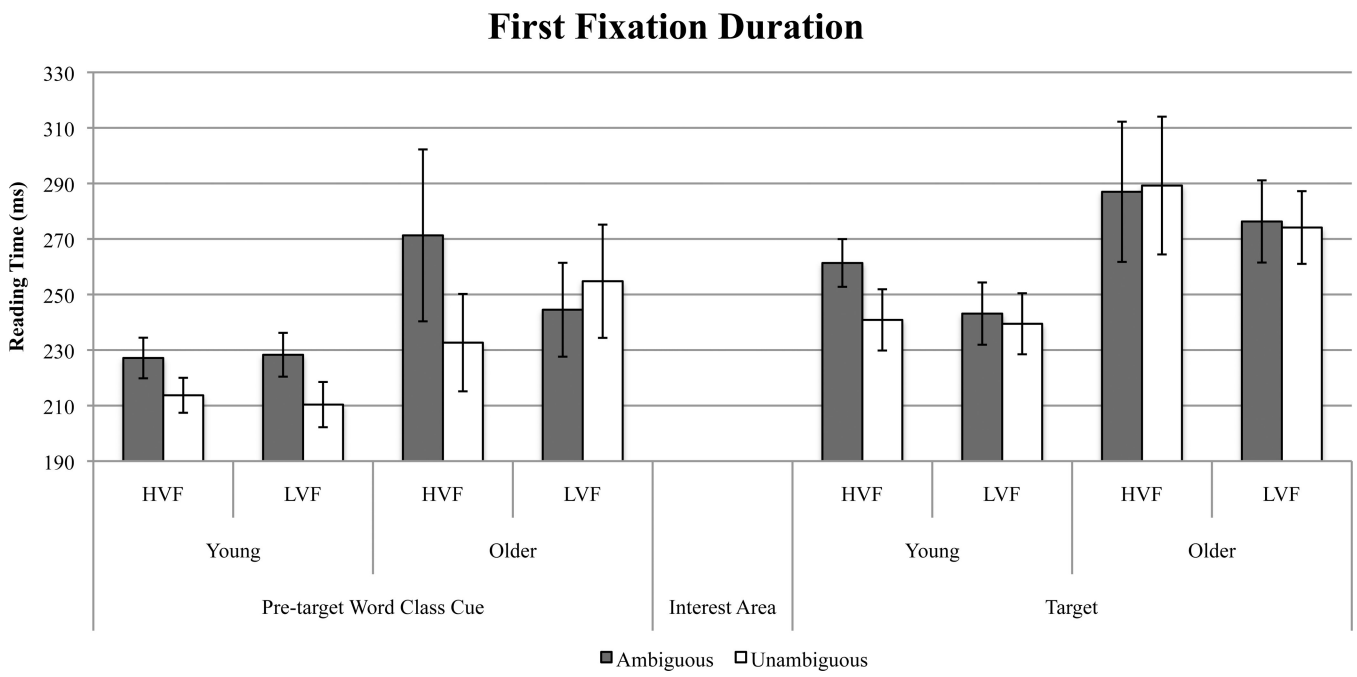

Figure 5.
First fixation durations on the pre-target word class cue and target word considered separately, in the syntactic prose contexts only, as a function of both age and verbal fluency group. (HVF = High verbal fluency; LVF = Low verbal fluency)

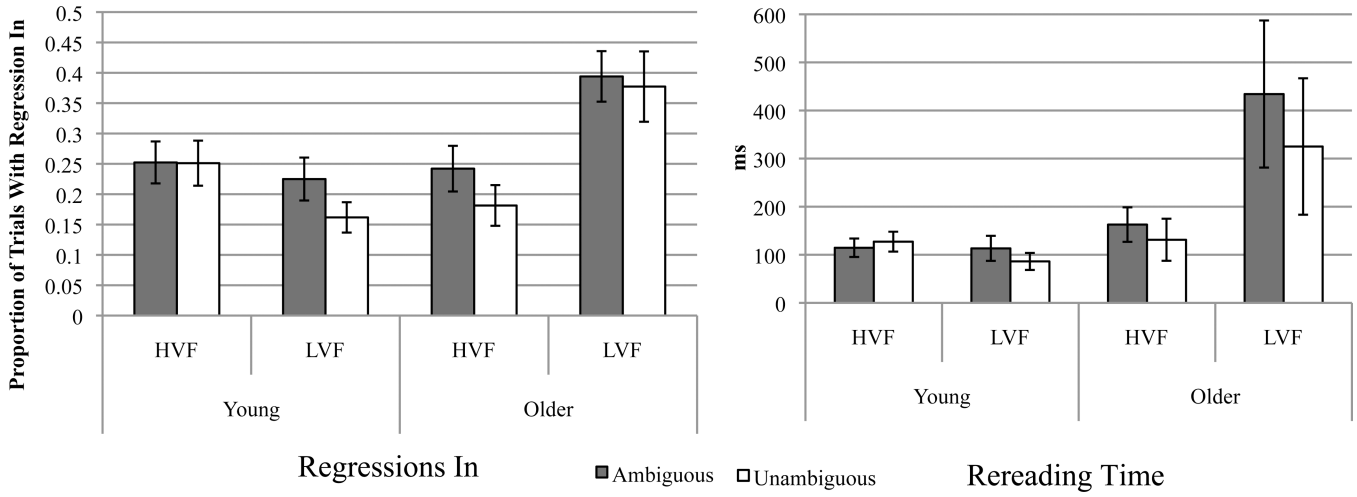

Figure 6.
Later reading measures on the target region in the syntactic prose contexts only, as a function of both age and verbal fluency group. (HVF = High verbal fluency; LVF = Low verbal fluency)

Table 1

Example sentences from critical contexts.

Word Type	Context Type	
	Congruent	**Syntactic Prose**
Ambiguous	On many freeways now, you can just use a pass instead of having to stop to pay the **toll**. This keeps traffic moving quickly	On many books now, you can just understand a room instead of having to walk to find the **toll**. This keeps psychology waiting quickly.
Unambiguous	Last weekend, they went to the theater, bought popcorn, and watched a **movie**. The theatre was crowded since it was Friday night.	Last lot, they tried to the company, chose dorm, and threw a **movie**. The charge was favorite since it was security game.

Table 2

Performance on neuropsychological measures for both younger and older adults. Standard deviations are provided in parentheses.

Measure	Young Adults		Older Adults	
	Mean	Range	Mean	Range
Reading Span				
Set size	2.5 (0.44)	2–3.5	2.5 (.60)	2–4
Total	54.8 (10.72)	30–69	51.2 (13.80)	35–78
Hayling Test				
Congruent Ending	23.5 (0.62)	22–24	23.9 (.25)	23–24
Anomalous Ending	22.1 (1.50)	18–24	21 (4.50)	10–24
Verbal Fluency				
Letter Fluency	45.2 (10.4)	32–67	48.6 (13.5)	26–70
Category Fluency	70.2 (14.5)	52–114	65.2 (13.0)	45–101
Print Exposure (d-prime)				
Author Recognition	1.19 (.54)	.16–2.15	1.37 (.72)	.31–2.89
Magazine Recognition	1.35 (.42)	.61–2.21	1.77 (.54)	.90–2.63

Table 3

D-prime scores (and standard deviations) for word and sentence recognition tasks.

Context	Word Type	Word		Sentence	
		Young	Old	Young	Old
Congruent					
	Ambiguous	2.83 (.29)	2.78 (.27)	2.24 (.56)	2.40 (.42)
	Unambiguous	2.76 (.28)	2.75 (.35)	2.31 (.54)	2.39 (.46)
Syntactic Prose					
	Ambiguous	2.60 (.41)	2.34 (.48)	1.81 (.76)	1.52 (.86)
	Unambiguous	2.57 (.36)	2.32 (.37)	1.59 (.48)	1.50 (.65)

J Exp Psychol Learn Mem Cogn. Author manuscript; available in PMC 2014 November 01.

Table 4

First fixation durations (and standard deviations) for young and older adults

Context	Word Type	Target Region		Target (Only)		Pre-target Word Class Cue	
		Young	Old	Young	Old	Young	Old
Congruent							
	Ambiguous	222 (20)	251 (42)	235 (28)	255 (50)	207 (25)	241 (46)
	Unambiguous	227 (23)	251 (38)	237 (23)	258 (49)	216 (34)	246 (44)
Syntactic Prose							
	Ambiguous	237* (20)	265 (43)	252† (31)	280 (54)	228† (22)	253 (45)
	Unambiguous	226 (22)	263 (43)	240 (32)	273 (57)	212 (21)	254 (44)

Note:

* indicates $p < .05$

† indicates $.05 < p < .08$ (for within-context comparison between ambiguous and unambiguous). *Target(Only)* refers to cases in which the first fixation in the target region was to the target word; *Pre-target Word Class Cue* refers to cases in which the first fixation in the target region was to the pre-target word class cue word.

J Exp Psychol Learn Mem Cogn. Author manuscript; available in PMC 2014 November 01.

Table 5

Remaining reading time measures (and standard deviations) on target region for both young and older adults.

Context	Word Type	Gaze Duration		Regressions In		Rereading Time	
		Young	Old	Young	Old	Young	Old
Congruent							
	Ambiguous	359 (69)	454 (99)	.17 (.11)	.22 (.12)	60* (37)	92 (129)
	Unambiguous	398* (75)	475 (102)	.16 (.11)	.23 (.15)	33 (33)	90 (122)
Syntactic Prose							
	Ambiguous	467 (81)	554 (132)	.24 (.10)	.32 (.13)	114 (67)	289* (289)
	Unambiguous	446 (73)	560 (128)	.21 (.10)	.29 (.16)	107 (60)	219 (287)

Note:

* indicates $p < .05$

† indicates $.05 < p < .08$ (for within-context comparison between ambiguous and unambiguous)

Table 6

Reading time measures (and standard deviations) divided by verbal fluency groups for both young and older adults in syntactic prose context only

Verbal Fluency	Word Type	First Fixation						Regressions In		Rereading Time	
		Target Region		Target: Only		Pre-target word class cue		Target Region		Target Region	
		Yng	Old	Yng	Old	Yng	Old	Yng	Old	Yng	Old
Low Fluency											
	Ambiguous	231 (19)	261 (33)	243 (34)	276 (42)	228 (24)	245 (48)	.22 (.11)	.39 (.12)	113 (78)	434 (432)
	Unambiguous	225 (25)	267 (34)	239 (33)	274 (38)	210 (24)	255 (58)	.16 (.08)	.38 (.16)	86 (53)	325 (401)
High Fluency											
	Ambiguous	242 (21)	273 (58)	261 (26)	287 (71)	227 (22)	271 (88)	.25 (.10)	.24 (.11)	115 (58)	163 (102)
	Unambiguous	228 (19)	265 (55)	241 (33)	289 (70)	214 (19)	233 (50)	.25 (.11)	.18 (.09)	127 (62)	131 (124)

Note: *Yng* denotes young adults. *Target(Only)* refers to cases in which the first fixation in the target region was to the target word; *Pre-target Word Class Cue* refers to cases in which the first fixation in the target region was to the pre-target word class cue word. in the target region was to the pre-target word class cue.

J Exp Psychol Learn Mem Cogn. Author manuscript; available in PMC 2014 November 01.