Brain & Language 115 (2010) 149–161

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Age-related and individual differences in the use of prediction during language comprehension

Kara D. Federmeier[a,*], Marta Kutas[b], Rina Schul[c]

[a] Department of Psychology, Program in Neuroscience, and The Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign IL, United States
[b] Department of Cognitive Science and Neurosciences, Center for Research in Language, and the Kavli Institute for Brain and Mind, University of California, San Diego CA, United States
[c] Department of Psychiatry and Counseling and Psychological Services, University of California, San Diego CA, United States

ARTICLE INFO

Article history:
Accepted 19 July 2010
Available online 21 August 2010

Keywords:
Aging
Language comprehension
Language production
Prediction
Category exemplar generation
Event-related brain potentials
N400
Frontal positivity
Verbal fluency

ABSTRACT

During sentence comprehension, older adults are less likely than younger adults to predict features of likely upcoming words. A pair of experiments assessed whether such differences would extend to tasks with reduced working memory demands and time pressures. In Experiment 1, event-related brain potentials were measured as younger and older adults read short phrases cuing antonyms or category exemplars, followed three seconds later by targets that were either congruent or incongruent and, for congruent category exemplars, of higher or lower typicality. When processing the less expected low typicality targets, younger – but not older – adults elicited a prefrontal positivity (500–900 ms) that has been linked to processing consequences of having predictions disconfirmed. Thus, age-related changes in prediction during comprehension generalize across task circumstances. Analyses of individual differences revealed that older adults with higher category fluency were more likely to show the young-like pattern. Experiment 2 showed that these age-related differences were not due to simple slowing of language production mechanisms, as older adults generated overt responses to the cues as quickly as – and more accurately than – younger adults. However, older adults who were relatively faster to produce category exemplars in Experiment 2 were more likely to have shown predictive processing patterns in Experiment 1. Taken together, the results link prediction during language comprehension to language production mechanisms and suggest that although older adults can produce speeded language output on demand, they are less likely to automatically recruit these mechanisms during comprehension unless top-down circuitry is particularly strong.

1. Introduction

Cognitive change with age is characterized by multiple trajectories (Verhaeghen & Cerella, 2008; Zelinski & Lewis, 2003). There is well documented slowing across a wide range of domains and tasks, including visual search, memory search, and word/nonword (lexical decision) judgments (see, e.g., Salthouse, 1991). Indeed, slowing is evident even in basic sensory and motor processing, as seen in latency measures of event-related brain potentials (ERPs) linked to perception and attention (Iragui, Kutas, Mitchiner, & Hillyard, 1993) and delays in simple reaction times (Fozard, Vercryssen, Reynolds, Hancock, & Quilter, 1994); some have linked cognitive changes directly to these more basic changes in processing speed (e.g., Baltes & Lindenberger, 1997). In addition to generalized slowing, particular cognitive tasks, such as spatial processing, source memory, and tasks that require cognitive control, have been argued to be especially impacted by aging (Spencer & Raz, 1995; Verhaeghen & Cerella, 2008; Verhaeghen, Steitz, Sliwinski, & Cerella, 2003).

On the other hand, on measures that require access to knowledge without significant time pressure, older adults generally perform as well as or even outperform their younger counterparts. For example, older adults show improvements on tests of general knowledge, vocabulary, and reading skill (Ackerman & Rolfhus, 1999; Ronnlund, Nyberg, Backman, & Nilsson, 2005; Uttl, 2002; Verhaeghen, 2003). Knowledge structure, as assessed by word associations, category exemplar generation, and synonym and antonym judgments have also been shown to remain stable with age (e.g., Bowles, Williams, & Poon, 1983; Burke & Peters, 1986; Howard, 1980; Lovelace & Cooley, 1982; reviewed in Light, 1991). These aspects of cognition that are more age-invariant have sometimes been referred to as ‘‘crystallized” (as opposed to ‘‘fluid”) intelligence (Horn & Cattell, 1967).

Age-related changes in brain structure and functioning with age are similarly variable across region. Some of the steepest rates of brain atrophy (in terms of gray matter volume) are observed in

* Corresponding author at: Department of Psychology, UIUC, 603 E. Daniel Street, Champaign, IL 61820, United States.
E-mail address: kfederme@illinois.edu (K.D. Federmeier).

0093-934X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandl.2010.07.006

prefrontal, frontal and parietal areas, including regions that have been associated with cognitive control, working memory, and spatial processing. In contrast, occipital areas and some parts of the temporal lobe have been found to remain much more stable with age (Raz et al., 2005; Resnick, Pham, Kraut, Zonderman, & Davatzikos, 2003). Connectivity changes with age as well, with a number of studies finding evidence in particular for reduced volume and efficacy of (pre)frontal white matter (Double et al., 1996; Pfefferbaum, Adalsteinsson, & Sullivan, 2005). Overall, the pattern is consistent with claims that age-related cognitive decline is most pronounced for functions that rely on frontal lobe circuitry (West, 1996).

Language comprehension, then, lies at an intersection of these age-related cognitive and neural trajectories. Clearly, comprehending an utterance critically requires accessing the kind of word-related information and world knowledge that seem to be augmented with age and experience. On the other hand, real-time language processing carries unavoidable time pressures as well as loads on frontally-mediated processes such as working memory that have been found to be impacted by aging (Light & Anderson, 1985; Park et al., 2002; Wingfield, Stine, Lahar, & Aberdeen, 1988). Perhaps not surprisingly, then, studies of language show mixed results of aging.

Tasks that tap into less resource-intensive (more "automatic") aspects of comprehension, such as lexical/semantic priming, often manifest little age-related change (e.g., Burke, White, & Diaz, 1987; Howard, McAndrews, & Lasaga, 1981; see review in Burke and Shafto (2008)), and hemodynamic imaging studies of these lexical functions have reported qualitatively similar activation patterns for younger and older adults (Madden et al., 1996, 2002). Like younger adults, older adults also show processing advantages for words in congruent sentence contexts (Roe et al., 2000; Stine-Morrow, Miller, & Nevin, 1999), and, in some cases, seem to rely even more heavily than younger adults on context information to, for example, aid word recognition under noise (e.g., Pichora-Fuller, Schneider, & Daneman, 1995). At the same time, however, there are also age-related changes in the availability and quality of that context information. As language tasks tap more heavily into older adults' ability to construct and use message-level meaning, spatial and temporal brain imaging methods alike suggest qualitative shifts (as well as important individual differences) in older adults' appreciation and use of language information, in terms of when information becomes available (Federmeier, van Petten, Schwartz, & Kutas, 2003), which brain areas are recruited (Grossman et al., 2002), and what processing strategies are brought to bear (Federmeier, McLennan, De Ochoa, & Kutas, 2002).

For example, Federmeier et al. (2003) measured the N400 component of the event-related brain potential (ERP) as older and younger adults listened to sentences for comprehension. The N400 reflects brain activity associated with relatively early, implicit aspects of semantic access and is reduced in amplitude for words in the presence of supportive word-, sentence-, and discourse-level context information (for review, see Kutas and Federmeier (2000)). Older adults showed N400 facilitation from the presence of lexically associated words in the sentence contexts that was similar to that measured for younger adults. The lack of any delay on the N400 in this condition, despite latency delays on earlier sensory components, suggests that older adults' increased experience with lexical information may help to counter more generalized slowing of sensory processing. However, effects of sentence-level congruity on the N400 were delayed by more than 200 ms in the healthy older sample. Given the rapid pace of normal language input (3–4 words per second), such delays in the impact of sentence-level context on processing are likely to result in qualitative changes in older adults' ability to use context information to, for example, resolve ambiguity (Dagerman, MacDonald, & Harm, 2006). In subsequent work, Federmeier and Kutas (2005) found that, compared to younger adults, older adults obtain reduced facilitation from strongly constraining contexts and that the ability of older adults to make use of contextual constraints was linked to working memory resources as indexed by reading span. These findings suggest that older adults, perhaps especially those with more limited working memory resources (e.g., Gunter, Jackson, & Mulder, 1995), may have a difficult time building and maintaining message-level representations of sentences with rich information.

Older adults have also been found to be less successful at making use of context information – for example, to prepare for the processing of likely upcoming words. A growing body of literature attests to the fact that the language comprehension system of younger adults uses context information to probabilistically preactivate semantic, morphosyntactic, phonological, and even orthographic features of likely upcoming words. Such predictive preactivation facilitates the semantic processing (as indexed by the N400) of items with predicted features (Federmeier & Kutas, 1999a; Laszlo & Federmeier, 2009). It even impacts the processing of preceding determiners and adjectives as a function of their agreement with predicted nouns (DeLong, Urbach, & Kutas, 2005; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2004). Correspondingly, there are also processing consequences when predictions are disconfirmed. In the ERP signal, such consequences are observed in the form of a frontally-distributed positivity elicited by unexpected but plausible words when these are encountered in the face of a strong prediction for a different item (e.g., "*The groom took the bride's hand and placed the ring on her dresser* (where *finger* is predicted)."; Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007). However, Federmeier et al. (2002) found that older adults as a group were less likely to manifest prediction-related benefits during sentence processing, although a subset of older adults with higher verbal fluencies – i.e., those that were able to more rapidly produce words in response to a letter or category cue – showed the young-like pattern of responses. This pattern of individual differences suggests that the speed with which older adults are able to generate appropriate words is predictive of the processing strategies (e.g., more predictive or more integrative in nature) that they are likely to bring to bear during comprehension.

Many older adults thus seem to use context information differently and perhaps less efficiently, and these age-related differences have been linked to both delays and reductions in the ability to integrate information over time in order to build effective message-level representations and to difficulties in using message-level representations to rapidly shape word processing, as through predictive preactivation. Teasing these factors apart is difficult in sentence processing tasks, since the availability of an integrated message-level representation is critical for context-based prediction. A primary aim of the current pair of studies is to determine whether there are age-related differences in predictive processing under task circumstances that do not impose the kinds of working memory and timing pressures that might accrue during sentence processing.

Therefore, in Experiment 1, we look at older adults' tendency to predict words from simple cues to lexical and world knowledge under conditions that impose little time pressure. In addition, we analyze individual differences to examine what cognitive skills and resources are conducive to the use of predictive processing strategies during language comprehension. In Experiment 2, we then test the hypothesis that prediction during comprehension is related to language production mechanisms by directly comparing patterns during comprehension with patterns seen when older and younger participants overtly generate responses to the same cues.

2. Experiment 1

In Experiment 1, we set out to examine age-related changes in predictive processing under task circumstances that reduce the cognitive loads that sentence processing may entail and that build upon older adults' rich base of lexical and world knowledge. To that end, we presented younger and older adults with phrasal cues for category exemplars (e.g., "An insect") or antonyms (e.g., "The opposite of closed"), followed by targets that were congruent ("ant" or "open") or incongruent ("gate" or "final") with the cues. These are cue types for which older adults readily produce accurate unspeeded responses, and previous work using ERPs to look at the comprehension of targets following cues like these found significant N400 priming for congruous (compared to incongruous) targets in all age groups (despite general age-related changes in N400 amplitude and timing; Kutas & Iragui, 1998).

To specifically probe for predictive processing, we added congruent but low typicality category targets (e.g., "hornet", following "An insect"). The N400 has been previously shown to be sensitive to semantic memory structure as reflected in typicality (Heinze, Muente, & Kutas, 1998), so both groups would be expected to elicit N400s of intermediate amplitude to these items. Critically, however, these low typicality exemplars, while in fact category members, are much less predictable than the high typicality congruent items. If participants use the cues to predict likely targets, then such plausible but unpredictable targets would be expected to elicit the frontal positivity, which, in sentence processing work, has been associated with neurocognitive consequences of having a strong prediction disconfirmed (Federmeier et al., 2007). Thus, we expect young participants to elicit frontal positivity to the low typicality targets as compared with both the high typicality targets and the incongruent ones (as, in sentence processing work, anomalous endings have not been associated with frontal positivity: Federmeier & Kutas, 1999a).

If age-related changes in predictive processing during sentence processing are secondary to difficulty in developing a message-level interpretation, then older adults may show predictive processing effects in this much simpler task, which loads less heavily on working memory and more heavily on world knowledge. If, however, deficits in prediction are more basic and hence more pervasive, then we expect older adults as a group to differ from younger ones in not showing this prediction-related brain activity. However, we might then expect to see individual differences, driven by verbal fluency (as in Federmeier et al., 2002) – perhaps especially by fluency with cued category exemplar generation, given its similarity to the specific knowledge domain tapped by this comprehension task.

2.1. Methods

2.1.1. Participants

Sixteen young adults attending the University of California, San Diego (7 men and 9 women, 18–24 years of age, mean age 20) and 20 San Diego community dwelling older adults (9 men and 11 women, 60–76 years of age, mean age 68) participated in the experiment and were compensated for their time with course credit or cash.[1] All were right-handed (Oldfield, 1971) monolingual English speakers with normal or corrected-to-normal vision and no history of reading difficulties or neurological/psychiatric disorders. Older adults were screened for cognitive impairment using the Mattis Dementia Scale (Mattis, 1976) and were on average more educated than the younger adults (all had completed high school, 10 of the older adults had completed college and 7 had advanced degrees).

2.1.2. Materials

Stimuli consisted of 240 phrasal cues followed by a single word target. 120 *antonym cues* (e.g., "The opposite of alive/float/leader/rare") were paired with expected and incongruent completions. Completions were matched for part of speech (most were adjectives, but nouns and adverbs were also common), length, and word frequency (Francis & Kucera, 1982). Expected antonyms were taken from online sources listing antonyms. These were highly associated with the antonym cues: average lexical association was 41.3% and for 85 of the 120 items the antonym was the top associate for the cue word (Nelson, McEvoy, & Schreiber, 1998). Incongruent antonyms were wholly unassociated with the cues.

In addition, 120 *category cues* (e.g., "A kind of tree", "A kitchen utensil", "A type of dance", "A kind of cheese") were paired with three types of endings: high typicality exemplars, low typicality exemplars, and incongruent exemplars. These exemplar types were matched for part of speech (most were nouns), length, and word frequency. High and low typicality exemplars were taken from existing category production norms (Battig & Montague, 1969; Hunt & Hodge, 1971; McEvoy & Nelson, 1982; Shapiro & Palermo, 1970). High typicality exemplars were generally the most frequently generated response to the category cues (median rank = 1), except in cases in which the most frequent response was not a single word, was also the response to a different category cue in the set, or had frequency characteristics that could not be matched to the other ending types. In those cases a different exemplar of as high a rank as possible was used. Low typicality exemplars were items that were generated in response to the cue, but with substantially lower probability (median rank = 15). For the most part, the targets did not have a strong lexical association with words in the category cues: lexical association between the target and any word in the category cue averaged 8.3% for the high typicality exemplars and 0.3% for the low typicality exemplars (Nelson et al., 1998). Incongruent exemplars were not category members, were never generated in response to the category cues, and had no lexical association with the cues. Table 1 gives examples of the stimuli.

Stimuli were divided into lists such that an individual subject would see each antonym or category cue only once. Half the items of each cue type were paired with incongruent endings and half with congruent endings; for category cues, half of the congruent endings were high typicality exemplars and half were low typicality exemplars.

Within each list, stimulus characteristics were controlled across target type, and across lists, each cue was paired with each of its ending types the same number of times.

2.1.3. Procedures

Participants were tested in a single experimental session conducted in a soundproof, electrically-shielded chamber. They were seated in a comfortable chair 40 in. in front of a monitor and instructed to read the cues and targets for the purpose of making a delayed congruity judgment.

The complete cue phrase was presented in the center of the screen for 2150 ms, followed by a variable inter-stimulus interval (ISI; a variable interval was used to temporally jitter anticipatory ERP responses) of between 750 and 1050 ms. The target was then presented for 1000 ms. Participants were asked not to blink or move their eyes during target presentation. After an ISI of 2000 ms, the cue "Yes or No?" was presented in the center of the screen for 2000 ms. Participants were asked to respond to this cue with a button press indicating if the target was congruent or incongruent with the cue phrase. The next cue phrase came up

[1] More older than younger adults were run in order to better equate power in the two samples, as older adults' ERP effects, for example on the N400 (Kutas & Iragui, 1998), have been shown to decrease systematically with age in paradigms like this.

Table 1
Example stimuli.

Opposite cue	Expected		Incongruent
The opposite of above	below		civil
The opposite of bottom	top		clear
The opposite of dirty	clean		final
The opposite of future	past		stage
The opposite of heaven	hell		chief
The opposite of male	female		green
The opposite of over	under		least
The opposite of rise	fall		name
The opposite of shallow	deep		active
The opposite of victory	defeat		company
Category cue	High typicality	Low typicality	Incongruent
A kind of tree	oak	ash	tin
A part of the human body	leg	throat	keys
A type of reading material	magazine	poem	ration
An insect	ant	hornet	gate
A type of dance	waltz	tap	bait
A type of tax	income	estate	sleeve
A green vegetable	lettuce	cabbage	patient
A part of a house	bedroom	basement	dictator
A cleaning instrument	broom	towel	porch
A water sport	swimming	rowing	pottery

automatically after an ISI of 3000 ms. Participants were given a short break after every 40 trials.

In a separate session, older adult participants returned to the lab and completed a battery of neuropsychological tests. These included verbal fluency (FAS and category: Benton & Hamsher, 1978), the Peabody Picture vocabulary test (Dunn & Dunn, 1997), and measures of short term and working memory (reading span, Daneman & Carpenter, 1980; the self-ordered pointing task, Petrides & Milner, 1982; forward and backward digit span, Wechsler, 1981).

2.1.4. EEG recording and analysis

The electroencephalogram (EEG) was recorded from 26 geodesically-spaced tin electrodes embedded in an Electro-cap, referenced to the left mastoid. A schematic of the electrode arrangement can be seen in Fig. 1. Blinks and eye movements were monitored via electrodes placed on the outer canthus (left electrode serving as reference) and infraorbital ridge of each eye (referenced to the left mastoid). Electrode impedances were kept below 5 kΩ. EEG was processed through Grass amplifiers set at a bandpass of 0.01–100 Hz. EEG was continuously digitized at 250 Hz and stored on hard disk for later analysis.

Data were re-referenced off-line to the algebraic average of the left and right mastoids. Trials contaminated by eye movements, blinks, excessive muscle activity, or amplifier blocking were rejected off-line before averaging. Trial loss averaged 13.3% for younger adults (range 4–25%) and 17.9% for older adults (range 0–34%). ERPs were computed for epochs extending from 100 ms before stimulus onset to 920 ms after stimulus onset. Averages of artifact-free ERP trials were calculated for each type of target word (expected antonym target, ANT-EX; incongruent antonym target, ANT-IN; high typicality category target, CAT-HI; low typicality category target, CAT-LO; incongruent category target, CAT-IN) after subtraction of the 100 ms pre-stimulus baseline. Only trials for which participants gave a correct button press response were included in the averages; each condition in each participant had a minimum of 15 trials. All *p*-values are reported after Epsilon correction (Huynh–Felt) for repeated measures with greater than one degree of freedom. Interactions with electrode are reported only when these are of theoretical significance.

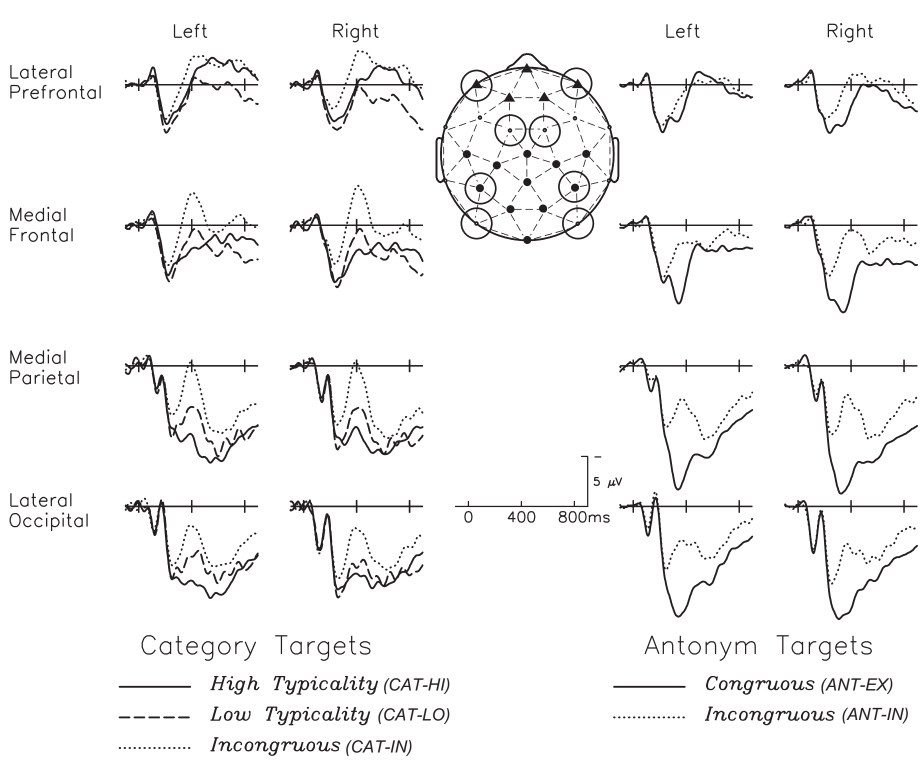

Fig. 1. Younger adults' ERP responses to category targets (at left) and antonym targets (at right) are plotted for a representative sample of electrodes across the head (circled electrode locations shown on head icon). Negative voltage is plotted up in this and all subsequent figures. N400 responses at medial centro-posterior electrode sites (filled circles on head icon; sites include Midline Central, Parietal, and Occipital electrodes, and Left and Right Medial Central, Dorsal Central, Dorsal Parietal, and Medial Occipital pairs) are larger for incongruent category and antonym targets (dotted lines) than for wholly expected ones (solid lines), and are intermediate for low typicality category targets (dashed line). Low typicality targets also elicited a sustained positivity (500–900 ms) at prefrontal electrode sites (filled triangles on head icon; sites include the Midline Prefrontal electrode and Left and Right Medial and Lateral Prefrontal pairs).

2.2. Results

2.2.1. Behavior

A delayed congruency judgment task was used to ensure that participants were attending to the stimuli and appreciating the relationship between the cues and targets. The behavioral results confirm this, as accuracy was near ceiling for both groups, with younger adults averaging 97.8% correct and older adults averaging 98.3% correct. In all conditions, older adults' accuracy was numerically identical to or higher than that for younger adults, in line with findings pointing to preserved or augmented knowledge stores with age. Patterns across conditions were the same for the two groups as well. In both groups, responses were less accurate to low typicality (CAT-LO) targets (young: 91.0%; old: 93.5%), whose fit to the category was more marginal by design, than to high typicality targets (young: 98.2%; old: 99.7%), which did not differ from incongruent targets (young: 97.5%; old: 99.3%) [young: CAT-LO vs. CAT-HI $t = 3.99$, $p < 0.05$, two-tailed; CAT-HI vs. CAT-IN $t = 1.51$, $p = 0.15$, two-tailed; old: CAT-LO vs. CAT-HI $t = 5.25$, $p < 0.05$, two-tailed; CAT-HI vs. CAT-IN $t = 0.96$, $p = 0.35$, two-tailed]. For antonym cues, in both groups there was a small but significant accuracy advantage for incongruent (young: 99.4%; old: 99.4%) than expected (young: 97.5%; old: 97.5%) targets [young: $t = 2.89$, $p < 0.05$, two-tailed; old: $t = 2.53$, $p < 0.05$, two-tailed].

2.2.2. ERPs

Figs. 1 and 2 show each group's ERPs at a representative selection of channels for antonym and category cues. In all conditions, sensory components are followed by a broadly-distributed negativity (N400), largest at centro-posterior sites. Younger and older adults' waveforms are characterized by general age-related morphological differences that have been described in prior work, including reductions in the amplitude of the P2 component (e.g., Federmeier et al., 2003) and reductions and delays in the N400 component and N400 effect (e.g., Kutas & Iragui, 1998). However, condition-related effects on the N400 show a similar pattern across groups: N400 amplitudes are more negative for incongruent targets (ANT-IN, CAT-IN) compared to expected ones (ANT-EX, CAT-HI). N400 responses to CAT-LO targets are intermediate between CAT-HI and CAT-IN. In the younger adults but not the older adults, CAT-LO targets are also characterized by a later-occurring positivity (500–900 ms) over prefrontal electrode sites. Older adults showed a left-lateralized frontal positivity to incongruous items (CAT-IN and ANT-IN), especially prominent in the antonym condition.

2.2.3. Group comparisons

2.2.3.1. N400. N400 latency and amplitude were assessed at the 11 medial centro-posterior channels where such responses are characteristically most prominent (e.g., Kutas & Hillyard, 1982); sites are marked with filled circles in Fig. 1. Peak latency of the N400 effect – i.e., the difference between each incongruent target type (ANT-IN, CAT-IN) and its highly expected counterpart (ANT-EX, CAT-HI), measured on waveforms that were bandpass filtered from 0.2 to 5 Hz – was 399 ms for young adults and 452 ms for older adults. A repeated measures Analysis of Variance (ANOVA) with Group as a between subjects variable and two levels of Condition (Antonym and Category) and 11 levels of Electrode Site as within subjects variables revealed a main effect of Group [$F(1, 34) = 21.71$, $p < 0.05$]. The age-related delay in N400 effect latency – approximately 50 ms in this sample – replicates prior observations using similar stimuli (Kutas & Iragui, 1998). There was also a main effect of Condition, with earlier peaking N400 effects in the Antonym (young: 379 ms; old: 437 ms) than in the Category (young:

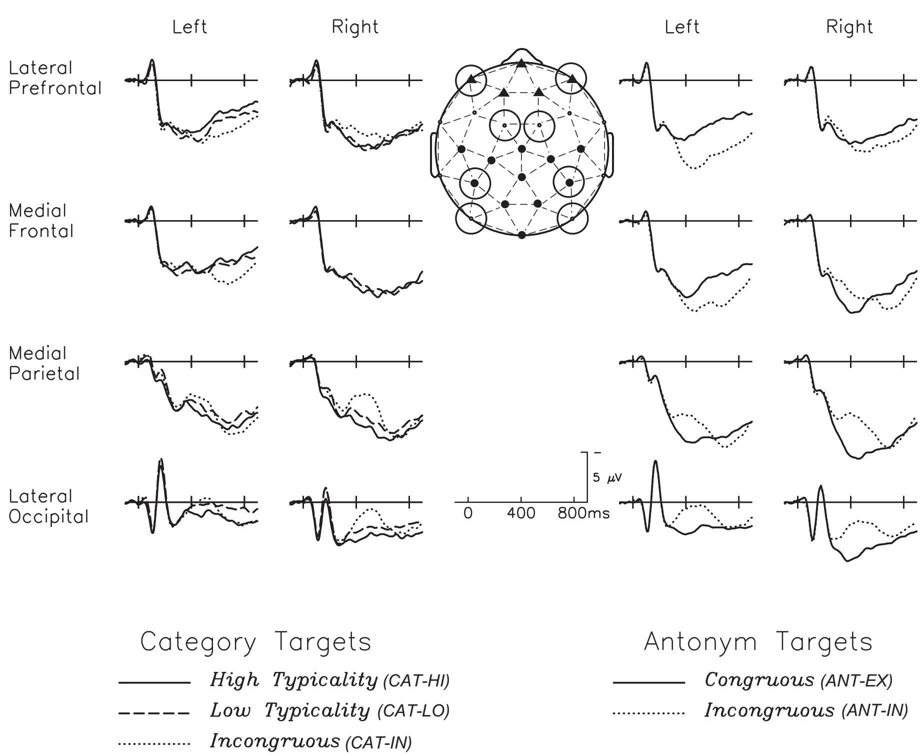

Fig. 2. Older adults' ERP responses to category targets (at left) and antonym targets (at right) are plotted for a representative sample of electrodes across the head (circled electrode locations shown on head icon). Replicating the pattern see for younger adults, N400 responses at medial centro-posterior electrode sites (filled circles on head icon; sites include Midline Central, Parietal, and Occipital electrodes, and Left and Right Medial Central, Dorsal Central, Dorsal Parietal, and Medial Occipital pairs) are larger for incongruent category and antonym targets (dotted lines) than for wholly expected ones (solid lines), and are intermediate for low typicality category targets (dashed line). The frontal positivity seen in the younger adult data for low typicality targets was absent in older adults as a group. Older adults also elicited a left lateralized positivity to incongruent targets, especially in the antonym condition.

420 ms; old: 466 ms) condition [$F(1, 34) = 10.11$; $p < 0.05$]; see Figs. 1 and 2. Group did not interact with Condition [$F(1, 34) = 0.27$; $p = 0.61$].

N400 mean amplitudes were then assessed in a 200 ms window centered on the effect peak for each group (i.e., 300–500 ms for younger adults and 350–550 ms for older adults). For antonym cues, an ANOVA with Group as a between subjects variable and two levels of Target Type (ANT-EX and ANT-IN) and 11 levels of Electrode Site as within subjects variables revealed no main effect of Group [$F(1, 34) = 0.56$; $p = 0.46$]. There was a main effect of Target Type, with more negative N400s to ANT-IN than ANT-EX targets [$F(1, 34) = 128.95$; $p < 0.05$], and this interacted with Group [$F(1, 34) = 16.39$; $p < 0.05$], reflecting the fact that N400 responses to ANT-IN items were fairly similar in amplitude across groups (young: 3.8 uV; old: 4.5 uV), whereas congruity-based reductions in N400 amplitude were greater for younger adults (young: 10.1 uV; old: 7.5 uV). Follow-up comparisons were then done within each group separately. For younger adults an ANOVA with two levels of Target Type (ANT-EX and ANT-IN) and 11 levels of Electrode Site revealed a significant effect of Target Type, with more negative N400s to ANT-IN than ANT-EX targets [$F(1, 15) = 82.50$; $p < 0.05$]. The same comparison for older adults also revealed a significant effect of Target Type [$F(1, 19) = 39.13$; $p < 0.05$]. Thus, both groups showed N400 priming for expected antonym targets.

For category cues, an ANOVA with Group as a between subjects variable and three levels of Target Type (CAT-HI, CAT-LO, and CAT-IN) and 11 levels of Electrode Site as within subjects variables again revealed no main effect of Group [$F(1, 34) = 0.13$; $p = 0.72$]. There was a main effect of Target Type [$F(2, 68) = 37.97$, $p < 0.05$, $\varepsilon = 0.85$] and a Group by Target Type interaction [$F(2, 68) = 7.91$, $p < 0.05$, $\varepsilon = 0.85$], with younger adults manifesting a fairly evenly graded pattern (CAT-IN: 1.2 uV; CAT-LO: 4.5 uV; CAT-HI: 6.6 uV) but older adults showing less differentiation between CAT-HI and CAT-LO conditions (CAT-IN: 2.6 uV; CAT-LO: 4.0 uV; CAT-HI: 4.6 uV). Again, follow-up comparisons were done within each group separately. For younger adults, an ANOVA with 3 levels of Target Type and 11 levels of Electrode Site revealed a significant effect of Target Type [$F(2, 30) = 19.30$, $p < 0.05$, $\varepsilon = 0.77$]. Pairwise comparisons revealed that N400 amplitudes were more positive to CAT-HI than CAT-LO targets [$F(1, 15) = 4.54$, $p < 0.05$] and were more positive to CAT-LO than CAT-IN targets [$F(1, 15) = 38.63$, $p < 0.05$]. A direct comparison of the amplitude of the overall congruency effect in the two conditions (ANT-IN minus ANT-EX vs. CAT-IN minus CAT-HI) revealed no significant difference in the size of the congruency effect [$F(1, 15) = 0.87$; $p = 0.37$]. For older adults, the same analysis also revealed a significant effect of Target Type [$F(2, 38) = 20.67$, $p < 0.05$, $\varepsilon = 1.0$]. Pairwise comparisons revealed that N400 amplitudes were more positive to CAT-HI (4.6 uV) than CAT-LO (4.0 uV) targets [$F(1, 19) = 4.58$, $p < 0.05$] and were more positive to CAT-LO than CAT-IN (2.6 uV) targets [$F(1, 19) = 17.66$, $p < 0.05$]. Thus, older adults also showed N400 facilitation for congruent as compared with incongruent category cues, and differentiated low from high typicality targets. A direct comparison of the amplitude of the overall congruency effect in the two conditions (ANT-IN minus ANT-EX vs. CAT-IN minus CAT-HI) revealed a statistically marginal tendency for larger effect sizes in the antonym than in the category condition [$F(1, 15) = 3.58$; $p = 0.07$].

2.2.3.2. Frontal positivity. As predicted, younger adults elicited a frontal positivity to the plausible but unexpected low typicality items, whereas this effect was not apparent in the older adults' waveforms. For younger adults, the positivity onset around 500 ms and continued through the end of the epoch. The effect was analyzed between 500 and 900 ms over the five prefrontal channels (marked with filled triangles in Fig. 1) where such effects have been found to be largest in prior work (Federmeier et al., 2007). An ANOVA with Group as a between subjects variable and 3 levels of Target Type and 5 levels of Electrode Site as within subjects variables revealed a main effect of Group [$F(1, 34) = 12.04$, $p < 0.05$]. There was also a main effect of Target Type [$F(2, 68) = 8.52$, $p < 0.05$, $\varepsilon = 1.0$], and a Group by Target Type interaction [$F(2, 68) = 7.91$, $p < 0.05$, $\varepsilon = 1.0$].

For younger adults, an ANOVA with 3 levels of Target Type and 5 levels of Electrode Site revealed a significant effect of Target Type [$F(2, 30) = 9.22$, $p < 0.05$, $\varepsilon = 1.0$]. Responses were more positive to CAT-LO targets (2.0 uV) than to CAT-HI targets (−0.5 uV) [$F(1, 15) = 9.06$, $p < 0.05$] or CAT-IN targets (−1.1 uV), which were not different from CAT-HI targets [$F(1, 15) = 0.77$, $p = 0.39$]. In contrast, for older adults the same analysis revealed no main effect of Target Type [$F(2, 38) = 1.25$; $p = 0.30$, $\varepsilon = 0.90$], and the pairwise difference between responses to CAT-HI (5.01 uV) and CAT-LO (5.50 uV) targets was not significant [$F(1, 19) = 1.98$; $p = 0.18$]. Indeed, in this group, it was responses to CAT-IN targets that were numerically most positive (5.64 uV).

2.2.4. Individual differences among older adults

To examine whether, as in Federmeier et al. (2002), a subset of older adults might be showing the frontal positivity effect, we conducted regression analysis to look at individual differences in the size of the frontal positivity effect, measured as the difference in mean amplitude (500–900 ms post-stimulus-onset) over the five prefrontal channels between the CAT-LO targets and the average of the CAT-HI and CAT-IN targets. The tendency to elicit more positive responses to the CAT-LO targets was not predicted by vocabulary [$R^2 = 0.00$, $F(1, 18) = 0.00$, $p = 0.99$]. A multiple regression using all of the short-term/working memory measures did not reveal a significant correlation [$R^2 = 0.08$, $F(3, 16) = 0.44$, $p = 0.73$], and none of the partial correlations for the individual short-term/working memory measures was significant (all p-values >0.30). A multiple regression using both of the verbal fluency measures was also not significant [$R^2 = 0.20$, $F(2, 17) = 2.19$, $p = 0.14$], but there was a significant independent contribution from category fluency [beta = 0.45, $t = 2.03$, $p < 0.05$, one-tailed]. Individuals with higher category fluency were more likely to show the young-like effect pattern; Fig. 3 shows a comparison of the effect pattern elicited by older adults in the top and bottom thirds of the group for category fluency and Fig. 4 (top) shows a scatterplot of the correlation. Mean category fluency (animals, fruits and vegetables, and first names) was 60.4 (range 40–72), which is comparable to previously published averages for educated older adults using similar methods (Tombaugh, Kozak, & Rees, 1999b), and prior observations in our own work (Federmeier et al., 2002).

2.3. Discussion

Young adults' N400 responses were similarly sensitive to categorical and antonym relationships, despite the differences between them: the categorical relationships were based on high levels of feature overlap but low levels of lexical association whereas the antonym relationships were characterized by important featural differences and high levels of lexical association. Mean N400 effects for these two types of semantic relations were indistinguishable in size, although the antonym effect peaked earlier – an interesting observation given that N400 latencies have been found largely invariant to psychological variables of this kind (see also Kutas & Iragui, 1998). For the category cues, typicality modulated the N400 response, with low typicality items eliciting N400 responses intermediate in amplitude between high typicality and wholly incongruous targets (cf, Heinze et al., 1998).

As expected based on patterns seen in sentence processing research (Federmeier et al., 2007), young adults also elicited a

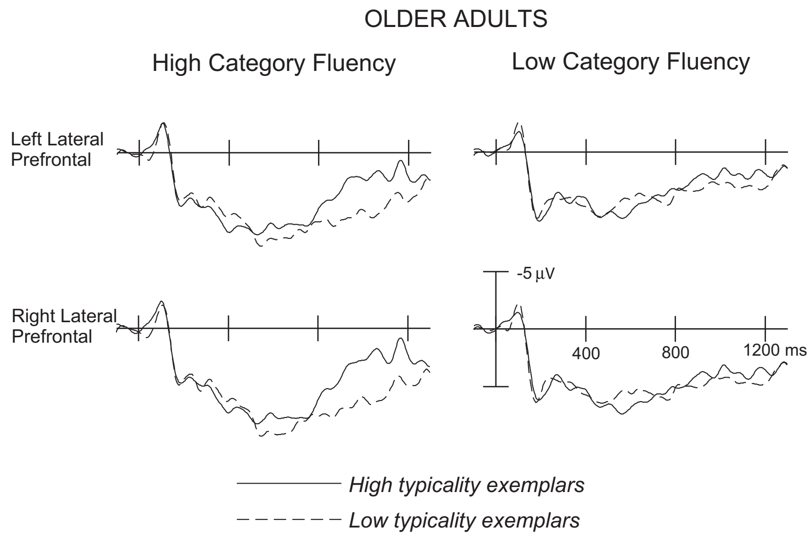

Fig. 3. ERP responses at the left and right lateral prefrontal electrodes are shown for the subset of older adults with the highest category fluency scores in the group ($N = 5$, at left) and with the lowest category fluency scores ($N = 6$, at right). Older adults with higher category fluency scores were more likely to show the sustained frontal positivity to low typicality exemplars, as seen for younger adults as a group.

sustained frontal positivity when processing the less predictable – but still plausible – low typicality items. The timecourse (500–900 ms) and distribution (maximal over prefrontal electrode sites) of this positivity are the same as that seen for plausible but unexpected words completing sentence frames. Consistent with the pattern observed in the present study, this positivity has been associated with processing elicited when predictions are disconfirmed. The appearance of the positivity in this design, and its pattern across conditions, extends our understanding in two ways.

First, the present results make clear that the frontal positivity is not specific to sentence processing but can appear even in simplified designs such as this one, in which prediction is made from a phrasal cue and in which the prediction-inducing cue and the predicted target are not part of a single linguistic unit. Second, and critically, the pattern of results across conditions demonstrates that the frontal positivity is not a simple index of unexpectedness or mismatch, since it appeared selectively for low typicality exemplars. Frontal positivity was *not* present for the wholly incongruous completions (to either category or antonym cues), even though these words are presumably more surprising and involve a higher level of mismatch. Patterns across sentence processing studies had suggested this, with unexpected but plausible endings eliciting the effect (Federmeier et al., 2007), but unexpected *implausible* endings failing to elicit the effect, even when related to the expected completion (Federmeier & Kutas, 1999a). In the current study, this pattern is shown for the first time within a single set of subjects and a single stimulus set.

Overall, then, the ERP data show that when young adults read phrasal cues in the context of making acceptability judgments, they predict likely targets. Such predictions serve to preactivate semantic, lexical, and orthographic information associated with likely upcoming words (DeLong et al., 2005; Federmeier & Kutas, 1999a; Laszlo & Federmeier, 2009; Van Berkum et al., 2005; Wicha et al., 2004), facilitating processing when those expected targets are encountered. However, when predictions are incorrect – but the input is plausible and thus must be integrated and appreciated as such – additional processing is involved, presumably to suppress or revise the initial prediction and/or to learn from the prediction error.

Given that younger adults do seem to routinely predict during language comprehension tasks, the primary goal of the present experiment was to examine whether healthy older adults would also engage in predictive processing in this case. Prior work in sentence processing studies has found that older adults as a group are less able or less likely to predict than their younger counterparts, although analyses of individual differences revealed that some older adults, particularly those with high verbal fluency, do predict (Federmeier et al., 2002). What has not been clear is whether these age-related changes in the use of predictive processing strategies are secondary to difficulties in building a message-level representation from rich context information (e.g., Federmeier & Kutas, 2005) and/or making that information available in time to impact processing during language comprehension at its normal rate (e.g., Federmeier et al., 2003). In the present study, therefore, we examined whether older adults make predictions during language processing tasks that place little demand on the need to maintain information in working memory or update memory over time. Instead, the task used simple cues to tap into the kind of knowledge demonstrated to accumulate over the life-span, and gave older adults ample time (approximately three seconds) to process those cues before the target was presented.

Older adults' N400 responses manifested the general delays and amplitude reductions that have been reported in past work (Federmeier et al., 2003; Gunter, Jackson, & Mulder, 1992), including designs similar to this one (Kutas & Iragui, 1998). Importantly, however, the pattern of N400 sensitivity to experimental variables was unchanged. Like younger adults, older adults showed significant N400 congruency effects for both category based relations and antonym based relations, and, as in younger adults, the antonym congruency effect peaked earlier (in older adults there was also a tendency for this effect to be larger than the category effect, although that difference did not reach statistical significance). Older adults also elicited N400 responses of intermediate amplitude to low typicality category exemplars. The online results thus replicate off-line results showing preservation with age of the basic structure of semantic memory, here in terms of both category (including typicality) and antonymy relations, and extend those findings by showing that this structure affects processing within the first 200–300 ms after word onset.

For both category and antonym targets, older adults also elicited a left-lateralized frontal difference between congruous and incongruous completions (more positive to incongruous). This pattern, with the same apparent scalp distribution, is also evident in Kutas and Iragui's (1998) data (see Fig. 1 in that paper) using the

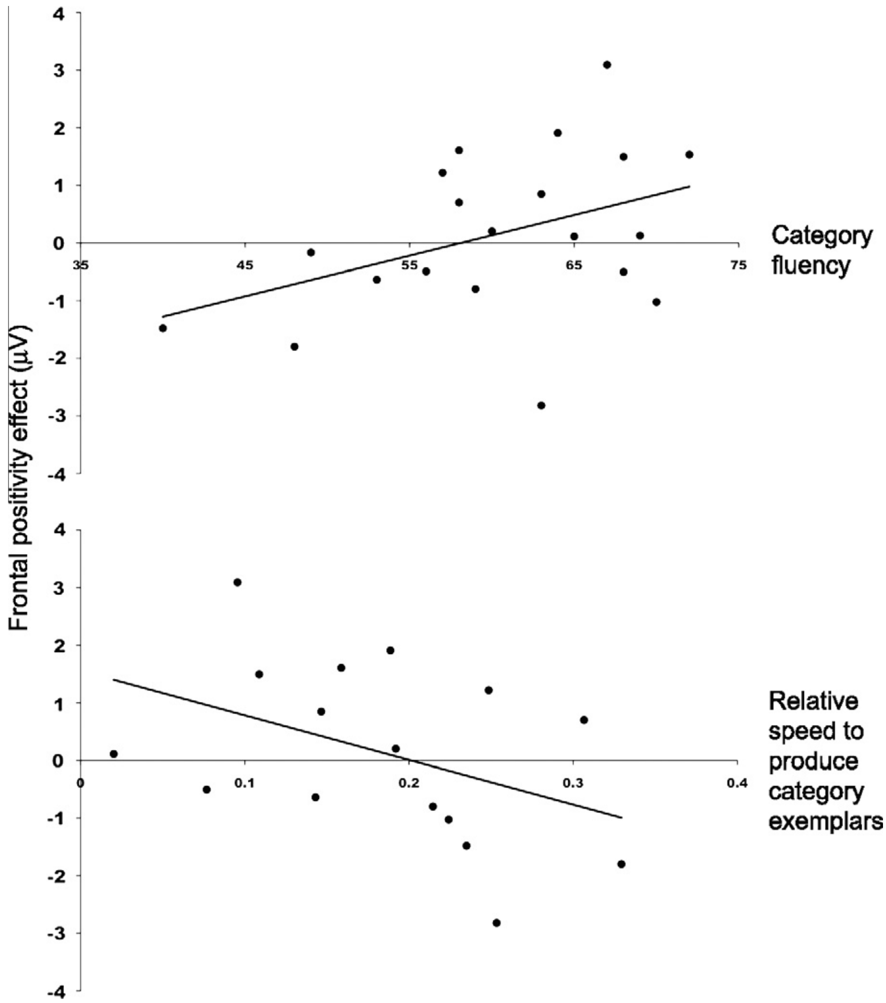

Fig. 4. At top is shown a scatterplot (with trendline) of the relationship between, on the *x*-axis, individual older adults' category fluency scores (total number of items produced across one minute trials for the three categories of animals, fruits and vegetables, and names) and, on the *y*-axis, the size, in microvolts, of their frontal positivity effects (mean amplitude difference, 500–900 ms post-stimulus-onset, over the five prefrontal electrode sites between the CAT-LO condition and the average of the CAT-HI and CAT-IN conditions). At bottom is shown, for those older adult participants who completed both Experiments 1 and 2, a scatterplot (with trendline) of the relationship between, on the *x*-axis, normalized response time differences in the language production task of Experiment 2 (average RT to produce correct responses to category cues minus average RT to produce correct responses to an antonym cues, divided by average overall correct RT) and, on the *y*-axis, the size, in microvolts, of their frontal positivity effects from Experiment 1. The tendency for older adults to show a frontal positivity effect was positively correlated with the number of correct category exemplars they could generate in a fixed time period and negatively correlated with their speed to produce targets to the category cues (relative to their response speed for the more lexically-constrained antonym cues).

same type of stimuli as in the present study. In particular, although in that paper the congruent and incongruent conditions are well-aligned over frontal channels for young adults in their 20s, responses to incongruent targets are more positive over left frontal channels beginning even for the group in their 30s and with a pronounced difference for the groups in their 50s and 60s. Across these two studies, the more positive-going left, frontal ERP response is specific to the incongruent conditions (not seen, for example, for the low typicality category targets in the present study) – and, therefore, for conditions that require a "no" response. In this data set, accuracy was too high to allow an examination of incorrect responses to determine if the pattern was driven more by whether the stimulus was incongruous or the executed response was a "no". Thus, the functional significance of this ERP response remains to be elucidated.

Important for the focus of the present study is the fact that older adults as a group notably failed to elicit the frontal positivity to CAT-LO targets that was observed in the younger participants and that has been associated with processing consequences arising from the use of prediction during comprehension. Thus, age-related changes in the tendency to engage predictive processing mechanisms are not limited to sentence processing, and are therefore unlikely to simply be secondary to demands imposed by the need to maintain and rapidly integrate information when processing a sentence context. In the current task, the cues were simple pointers to world knowledge and participants had ample time to process the cue before apprehending the target. These conditions would seem very conducive for allowing older adults to generate predictions and get them ready in time ... and yet they did not. The age-related changes observed in preactivation during sentence processing thus generalize to very different task circumstances and may therefore reflect something fairly fundamental about how older, as compared with younger, adults make use of language input online.

Although older adults as a group did not elicit the frontal positivity, analyses of individual differences revealed that a subset of older adults did show the young-like pattern associated with predictive processing, with enhanced frontal positivity selective for the low typicality targets. The tendency to show this pattern was not predicted by short-term/working memory measures or vocabulary size, but was correlated with category fluency (ability to rapidly generate appropriate exemplars in response to a category label cue). This accords with findings in sentence processing tasks, in which verbal fluency was found to be the best correlate of the tendency to show prediction-related effects on the N400 (Federmeier et al., 2002). However, whereas in the sentence processing study, letter and category fluency were jointly correlated with prediction-related effects, in the present task looking at prediction of category-related information, only category fluency (and *not* letter fluency) was correlated with the tendency to exhibit the frontal positivity. This suggests that it is the ability to generate information specifically relevant for the particular task – and not some more general aspect of fluency or speed – that promotes the tendency or ability to predict during comprehension.

The fact that predictive processing has been associated with verbal fluency across several types of stimulus and task environments lends support to the hypothesis that predictive processing in language comprehension taps into processes involved in generating language output ... i.e., language production mechanisms (Federmeier et al., 2007). Yet we know of no studies that directly compare results from language comprehension tasks that provide measures of predictive processing with those from overt language production tasks using the same materials and subject populations. Therefore, to examine the prediction-production link more directly, in Experiment 2 we used the same category and antonym cues from Experiment 1 in a speeded language production task.

3. Experiment 2

To examine the possible links between language production mechanisms and predictive processing during language comprehension, in this experiment we presented the antonym and category cues from Experiment 1 in a speeded generation task. Younger and older adults' speed (as measured by voice onset times) and accuracy to respond to each cue type were examined. As large a subset as possible of the older adults from Experiment 1 were brought back so that language production measures could be directly correlated with ERP responses.

Although work has shown age-preservation in the ability to generate knowledge-based information such as category exemplars in off-line tasks (see review in Light, 1991), much less is known about how older adults respond to these cues under time pressure. Verbal fluency measures have been found to decline slightly with age, and more so for category fluency than for letter fluency (Tombaugh, Kozak, & Rees, 1999a). Thus, generation tasks may be subject to some degree of age-related slowing. Yet, when only a single response is required to a given cue, response times in some tasks (such as verb generation) have not been found to differ significantly across age groups (e.g., Persson et al., 2004).

If age-related differences in the tendency to use predictive processing mechanisms during comprehension stem from generalized slowing in language production mechanisms, then we would predict that older adults as a group would respond less quickly than younger adults to the cues in Experiment 2, and perhaps show responses delayed beyond about three seconds (when the targets in Experiment 1 appeared). This would suggest that the failure to use predictive processing mechanisms during comprehension might arise simply because older adults are unable to generate the necessary information in time for it to shape comprehension. If so, the subset of older adults who did show predictive processing patterns may be those who are able to overtly produce responses in time.

Alternatively, if older and younger adults' response times are similar – and particularly if both groups can, on average, generate responses to the cues fast enough to make clear that they could have generated timely predictions in Experiment 1 – then a simple slowing account would seem insufficient to explain the age-related comprehension differences in Experiment 1. Instead, this outcome would suggest that older adults are less likely to recruit relatively intact language production mechanisms during comprehension – or that the recruitment of these processes is less efficacious. Given the observed correlation between the frontal positivity and category fluency in Experiment 1, however, we might still expect to find correlations between the ERP results and some aspect of older adults' responses to the cues in Experiment 2.

3.1. Methods

3.1.1. Participants

Young adult participants ($N = 22$, 10 men and 12 women, 18–21 years of age, mean age 19) were recruited from the same population of UCSD undergraduates as that used for Experiment 1 and compensated for their time with course credit or cash. Sixteen of the older adults (8 men and 8 women) who participated in Experiment 1 were able and willing to return to participate in Experiment 2. Test sessions were on average 5 months apart (range 3 months to 10 months). An additional 6 older adult participants (1 man, 5 women) were recruited from the same population of San Diego community dwelling older adults. As a group, the 22 older adult participants had a mean age of 69 (range 60–77 years of age) and were on average more educated than the younger adults (all had completed high school, 12 had completed college, and 9 had advanced degrees). They were compensated for their time with cash. All participants were right-handed (Oldfield, 1971) monolingual English speakers with normal or corrected-to-normal vision and no history of reading difficulties or neurological/psychiatric disorders. Older adults were screened for cognitive impairment using the Mattis Dementia Scale (Mattis, 1976).

3.1.2. Materials and procedures

Stimuli consisted of the 240 phrasal cues (120 *antonym cues and* 120 *category cues*) from Experiment 1. Antonym and category cues were intermixed and presented in random order to each participant.

Participants were tested individually in a quiet room. They received instructions before a brief practice session. Cue stimuli were presented visually on a CRT using PsyScope software. Participants were told to read each cue and to generate a one word response as quickly as possible without compromising accuracy. Cues remained on the screen for ten seconds or until a response was registered. Voice responses were collected using a microphone connected to a CMU button box, designed to log millisecond-precision response times via the Macintosh modem port. A failure to respond was automatically marked in the data file if the image disappeared prior to the participant's response; on average, younger adults failed to give a response on 6 out of the 240 trials (range 0–20) and older adults on 4 (range 0–10). All other responses were hand recorded by an experimenter during testing, and voice key failures (irrelevant noise or failure to perceive sound) were coded, resulting in 3% data loss for younger adults and 4% for older adults. For verification purposes, an external tie-microphone recorded verbal responses to a magnetic tape recorder. As in Experiment 1, participants were given a short break after every 40 trials.

3.1.3. Data coding

Responses to category cues were counted as correct if they appeared in the category norms from which the stimuli were derived (Battig & Montague, 1969; Hunt & Hodge, 1971; McEvoy & Nelson, 1982; Shapiro & Palermo, 1970). Responses to antonym cues were counted as corrected if they were the expected lexical item from Experiment 1 or a reasonable synonym of it (e.g., *The opposite of polite: rude, impolite, discourteous; The opposite of static: dynamic,* *active, fluid; The opposite of boring: interesting, exciting, amusing*). Additional items removed for response time analyses but not considered errors (4% of all responses for both younger and older participants) were responses that were semantically appropriate but the wrong part of speech (e.g., *A water sport: swim; The opposite of cowardly: hero*) or, in the case of opposites, that were along the correct semantic dimension but differed from the expected response in scale (e.g., *The opposite of pleasure: agony; The opposite of cool: hot; The opposite of future: present*). Examples of responses judged to be clear errors were: *An island: Florida; A green vegetable: carrot; A wind instrument: windmill; The opposite of heavy: soft; The opposite of little: small; The opposite of wrinkled: pure*.

3.2. Results

3.2.1. Accuracy

Table 2 shows accuracy data for antonym and category cues for younger and older adults. Overall accuracy was high: 89% for both younger and older adults. An ANOVA with two levels of Group (younger and older adults) and two levels of Cue Type (antonyms and categories) revealed no main effect of Group [$F(1, 42) = 0.01$, $p = 0.91$]. There was a main effect of Cue Type [$F(1, 42) = 188.76$, $p < 0.05$] and a Group by Cue Type interaction [$F(1, 42) = 8.93$, $p < 0.05$]. Pairwise comparisons following up on this interaction revealed that older adults were more accurate than young adults in responding to category cues [$F(1, 42) = 4.60$, $p < 0.05$], whereas the groups did not differ in accuracy to respond to antonym cues [$F(1, 42) = 1.29$, $p = 0.26$]. The difference in accuracy on the category cues was not due to younger adults producing a larger number of incorrect responses [$F(1, 42) = 0.17$, $p = 0.68$] but rather to their greater tendency to omit responses [$F(1, 42) = 4.65$, $p < 0.05$] (and thus was not a speed/accuracy tradeoff).

3.2.2. Response times

Response times (voice onsets) to antonym and category cues for both younger and older adults are given in Table 2. An ANOVA with two levels of Group (younger and older adults) and two levels of Cue Type (antonyms and categories) revealed no main effect of Group [$F(1, 42) = 1.39$, $p = 0.24$] and no interaction of Group with Cue Type [$F(1, 42) = 0.01$, $p = 0.91$]. There was a main effect of Cue Type [$F(1, 42) = 129.21$, $p < 0.05$], with faster responses to antonym cues (1603 ms) than to category cues (1977 ms).

3.2.3. Individual differences

For that subset of older adults ($N = 16$) who participated in both experiments, regression analyses were used to examine whether overt responses to the cues were predictive of the tendency to show the frontal positivity effect (again measured as the difference in mean amplitude, 500–900 ms post-stimulus-onset, over the five prefrontal channels between the CAT-LO targets and the average of the CAT-HI and CAT-IN targets), which is indicative of predictive processing during comprehension. The tendency to elicit the frontal positivity effect was not predicted by response accuracy across the cues [$R^2 = 0.07$, $F(2, 13) = 0.52$, $p = 0.61$] nor by accuracy to either condition independently (both ps > 0.3). Frontal ERP responses were also not predicted by reaction times taken together [$R^2 = 0.21$, $F(2, 13) = 0.44$, $p = 0.21$], but the partial correlation of category response times just missed significance [beta = −0.84, $t = 1.67$, $p = 0.06$, one-tailed], showing the expected pattern that faster responses to category cues were associated with greater ERP frontal positive effects. There was also a marginal positive correlation between antonym cue reaction times and the frontal positive effect [beta = 0.93, $t = 1.86$, $p = 0.08$, two-tailed]; that is, opposite to the effect for category cues, faster responses to antonym cues predicted reduced frontal positive effects in the ERPs. This pattern suggests that the tendency to show the frontal ERP effect may be better predicted by the relative – rather than absolute – speed of responding to the category cues (all participants, both younger and older, were faster on average to respond to the more lexically-constrained antonym cues). In other words, overall response times to the category cues reflect a combination of general response speed (which may not be predictive of the frontal ERP effect) and speed to specifically access category information. To look at speed to respond to the category items with more general response speed factored out, the frontal ERP effect was regressed with the average response speed difference between the category and antonym cue conditions, normalized by each participant's overall average response time. This revealed a significant correlation [$R^2 = 0.18$, $t = 1.73$, $p = 0.05$, one-tailed]; a scatterplot of the relationship is shown in Fig. 4 (bottom). The average response speed difference was itself significantly correlated with category fluency [$R^2 = 0.28$, $t = 2.34$, $p < 0.05$, one-tailed] but not with letter (FAS) fluency [$R^2 = 0.05$, $t = 0.80$, $p < 0.22$, one-tailed].

Table 2
Accuracy (# correct out of 120) and response times (milliseconds to voice onset) for younger and older adults for the two cue types. Standard deviations in parenthesis.

	Antonym cues		Category cues	
	Accuracy	Response times	Accuracy	Response times
Younger adults	103 (5)	1550 (254)	111 (4)	1920 (342)
Older adults	102 (6)	1656 (285)	113 (3)	2033 (404)

3.3. Discussion

Both younger and older adults responded more quickly to the antonym cues, for which correct responses were more highly associated with the cue and more lexically constrained. The greater constraint of the antonym cues also resulted in both groups obtaining fewer correct trials for this condition compared to the category condition, although accuracy in both conditions was high. Strikingly, older adults were not significantly slower to respond to either the category or antonym cues. Furthermore, for the category cues, older adults were slightly – but significantly – more accurate than younger adults, as the younger group was more likely to fail to generate responses to these items. Thus, there was no evidence for age-related decline on the speeded generation task, and some evidence for small but reliable age-related benefits.

Notably, both groups began producing responses within, on average, 1.5–2 s after cue presentation – well in advance of when the comprehension targets were presented in Experiment 1. Thus, when the task demands called for overt generation, older adults seemed as able as younger ones to generate responses in a timely (and accurate) manner. The hypothesis that older adults in Experiment 1 failed to show prediction-related effects because of generalized slowing in their ability to generate responses to the cue items is therefore disconfirmed by the data patterns from Experiment 2.

Nevertheless, aspects of older adults' production behavior were correlated with individual differences in predictive processing effects during the comprehension task. For the subset of older adults that we were able to bring back for Experiment 2, the partial correlation of category response times with the frontal positive effect just missed significance (and was a "medium" sized effect, $f^2 = 0.21$). And, when overall individual differences in response speed were taken into account, the relative speed to generate category exemplars was significantly correlated with the tendency to show the frontal positivity, as well as with category (and not FAS) fluency. Thus, older adults' facility with producing responses to the specific category cues in our stimulus set, as well as their more general ability to produce category exemplars on demand, as in the verbal fluency task, is predictive of their tendency to show young-like patterns of ERP responses in comprehension tasks. Thus, even though the age-related shift in comprehension processes does not arise directly from slowing of production mechanisms, it is the case that relatively faster production of (and correspondingly higher levels of fluency for) the kind of information called upon by the comprehension task mitigates against age-related change.

4. General discussion

Young adults show prediction-related effects in language comprehension under a wide range of task circumstances, from verification tasks like that in Experiment 1 to listening or reading sentences for comprehension (DeLong et al., 2005; Federmeier & Kutas, 1999a; Federmeier et al., 2002; Laszlo & Federmeier, 2009; Van Berkum et al., 2005; Wicha et al., 2004). Prediction affords facilitation when an item containing expected features is encountered (Federmeier & Kutas, 1999a; Federmeier et al., 2002; Laszlo & Federmeier, 2009), but has consequences when unexpected items are encountered instead (Federmeier et al., 2007). In particular, a frontally-distributed positivity occurring between 500 and 900 ms post-stimulus onset has been associated with the processing of stimuli that disconfirm a likely prediction. Extending patterns seen across studies in the sentence processing literature, the results from the current experiments show that this processing consequence is specific to unexpected items that are also plausible, suggesting the effect reflects something about the need to assimilate the new information into the current representation, rather than a more general effect of unexpectedness or mismatch.

Older adults as a group seem either less likely to engage predictive processing mechanisms during comprehension or less effective at doing so. On average, they fail to show the kinds of facilitative effects of prediction seen for younger comprehenders (Federmeier et al., 2002) and, in the present study, also fail to show the kind of processing consequences seen for young adults when expectations are violated (as in the frontal positivity). In sentence processing tasks, failure to predict might have arisen as a consequence of older adults finding it more difficult to build a sentence message-level representation from which to make a prediction (e.g., Federmeier & Kutas, 2005), perhaps because of working memory limitations. However, the results of Experiment 1 showed that older adults also fail to show predictive processing during comprehension when working memory demands are relatively minimized and when tasks instead rely primarily on the kind of knowledge shown to be preserved or augmented with age.

A second hypothesis is that age-related slowing in the mechanisms needed to generate language information on demand may result in predictions not becoming available in time to impact comprehension. However, Experiment 2 revealed no evidence that older adults are actually slower than young adults at producing category exemplars or antonyms on demand. Furthermore, in Experiment 1, older adults were given three seconds to process the cue before the target was presented, and Experiment 2 showed that it took them, on average, only two seconds to begin producing an overt response to the same cues. Thus, older adults seemingly had sufficient time to generate expectations about the targets in Experiment 1, but nevertheless did not seem to be doing so, at least at the group level.

Interestingly, then, even when older adults are not less fluent than young adults (Federmeier et al., 2002) and are not slower to produce responses to language cues (as in the present experiment), they are still less likely to manifest evidence of using generative strategies during comprehension. This pattern is reminiscent of findings in other domains suggesting that older adults may be less likely to routinely engage top-down mechanisms to, for example, promote memory for verbal material (e.g., Craik & McDowd, 1987; Logan, Sanders, Snyder, Morris, & Buckner, 2002), or avoid attentional capture (Whiting, Madden, & Babcock, 2007) even though, in at least some cases, they seem able to successfully employ these processes when task demands or explicit instructions cue them to do so. One hypothesis, then, is that age-related deterioration of frontal lobe pathways reduces the likelihood that top-down mechanisms will be automatically engaged by incoming stimuli. In other words, when young adults comprehend language, top-down (language production related) circuitry may routinely be involved (at least in the left cerebral hemisphere, Federmeier, 2007). Older adults may seem to show preserved functioning of these mechanisms when tasks explicitly call upon them, as during overt production. However, with age, these top-down circuits may be less reliably engaged under conditions of passive stimulation where control is internally (rather than externally) orchestrated, as during normal reading or listening for comprehension. The result, then, is that language comprehension comes to be more dominated by feed-forward mechanisms.

Importantly, however, we also find individual differences among older comprehenders, such that a subset of older adults shows the young-like pattern. For prediction-related effects, these individual differences have been consistently most strongly related to language production measures such as verbal fluency, strengthening the proposed link between prediction during comprehension and language production mechanisms. In the present experiment, the tendency for older adults to show prediction-related effects when comprehending category cue targets was specifically correlated with measures of category fluency and speed of cued category exemplar generation. Thus the individual differences are not just a simple function of overall fluency or speed – consistent with the fact that there were no overall age-related differences in production speed. On the hypothesis outlined above, then, strong connectivity between the frontal lobes and other cortical regions involved in processing/storing a given type of information supports high levels of fluency for that information type (and, indeed, verbal fluency tasks are associated with both frontal and temporal lobe functioning: e.g., Gourovitch et al., 2000). In turn, this strong connectivity makes it more likely that bottom-up activation of that information will engage frontally-mediated top-down processing mechanisms, like those involved in prediction.

Irrespective of the precise nature of the neural or cognitive changes responsible for the observed age-related and individual differences, however, this pair of experiments contributes to a growing body of work suggesting that language comprehension involves multiple mechanisms, which differ in the extent to which they are primarily stimulus-driven as opposed to shaped by top-down, context-based expectancies (e.g., Federmeier, 2007; Federmeier & Kutas, 1999b). For many people, aging seems to involve a shift from comprehension mechanisms that are strongly shaped by expectancies to those that are more stimulus-driven and feed-forward, with concomitant implications for how efficient, error-prone, flexible, and adaptive language processing will be. However, just as there are multiple age-related trajectories across cognitive domains, there are also multiple trajectories within a given domain across people. Uncovering the factors that influence the nature or degree of age-related neurocognitive changes in language processing promises to provide us with the greatest opportunity to maintain effective ways of gaining information from the environment and maintaining social ties across the life-span.

Acknowledgments

This work was supported by NIA Grant AG026308 to K.D.F. and NIA Grant AG08313 and NICHD Grant HD22614 to M.K. We thank Charlene Lee for help with data processing.

References

Ackerman, P. L., & Rolfhus, E. L. (1999). The locus of adult intelligence: Knowledge, abilities, and nonability traits. *Psychology & Aging, 14*(2), 314–330.

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology & Aging, 12*, 12–21.

Battig, W. F., & Montague, W. E. (1969). Category norms of verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology, 80*(3, Pt. 2), 1–46.

Benton, A. L., & Hamsher, K. (1978). *Multilingual aphasia examination manual*. Iowa City, IA: University of Iowa.

Bowles, N. L., Williams, D., & Poon, L. W. (1983). On the use of word association norms in aging research. *Experimental Aging Research, 9*(3), 175–177.

Burke, D. M., & Peters, L. (1986). Word associations in old age: Evidence for consistency in semantic encoding during adulthood. *Psychology & Aging, 1*(4), 283–291.

Burke, D. M., & Shafto, M. (2008). Language and aging. In F. I. M. Craik & T. A. Salthouse (Eds.). *Handbook of aging and cognition* (Vol. 3, pp. 373–444). New York: Psychology Press.

Burke, D. M., White, H., & Diaz, D. L. (1987). Semantic priming in young and older adults: Evidence for age constancy in automatic and attentional processes. *Journal of Experimental Psychology: Human Perception & Performance, 13*(1), 79–88.

Craik, F. I. M., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory & Cognition, 13*(3), 474–479.

Dagerman, K. S., MacDonald, M. C., & Harm, M. W. (2006). Aging and the use of context in ambiguity resolution. *Cognitive Science, 30*, 311–345.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior, 19*(4), 450–466.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117–1121.

Double, K. L., Halliday, G. M., Kril, J. J., Harasty, J. A., Cullen, K., Brooks, W. S., et al. (1996). Topography of brain atrophy during normal aging and Alzheimer's disease. *Neurobiology of Aging, 17*(4), 513–521.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test* (3rd ed. (PPVT-III)). Circle Pines, MN: American Guidance Service.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*(4), 491–505.

Federmeier, K. D., & Kutas, M. (1999a). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language, 41*(4), 469–495.

Federmeier, K. D., & Kutas, M. (1999b). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research, 8*(3), 373–392.

Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology, 42*(2), 133–141.

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology, 39*(2), 133–146.

Federmeier, K. D., van Petten, C., Schwartz, T. J., & Kutas, M. (2003). Sounds, words, sentences: Age-related changes across levels of language processing. *Psychology & Aging, 18*(4), 858–872.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research, 1146*, 75–84.

Fozard, J. L., Vercryssen, M., Reynolds, S. L., Hancock, P. A., & Quilter, R. E. (1994). Age differences and changes in reaction time: The Baltimore Longitudinal Study of Aging. *Journal of Gerontology, 49*(4), P179–189.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage*. Boston: Houghton Mifflin Company.

Gourovitch, M. L., Kirkby, B. S., Goldberg, T. E., Weinberger, D. R., Gold, J. M., Esposito, G., et al. (2000). A comparison of rCBF patterns during letter and semantic fluency. *Neuropsychology, 14*(3), 353–360.

Grossman, M., Cooke, A., DeVita, C., Chen, W., Moore, P., & Detre, J. (2002). Sentence processing strategies in healthy seniors with poor comprehension: An fMRI study. *Brain and Language, 80*(3), 296–313.

Gunter, T. C., Jackson, J. L., & Mulder, G. (1992). An electrophysiological study of semantic processing in young and middle-aged academics. *Psychophysiology, 29*(1), 38–54.

Gunter, T. C., Jackson, J. L., & Mulder, G. (1995). Language, memory, and aging: A electrophysiological exploration of the N400 during reading of memory-demanding sentences. *Psychophysiology, 32*(3), 215–229.

Heinze, H.-J., Muente, T.-F., & Kutas, M. (1998). Context effects in a category verification task as assessed by event-related brain potential (ERP) measures. *Biological Psychology, 47*(2), 121–135.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica (Amsterdam), 26*(2), 107–129.

Howard, D. V. (1980). Category norms: A comparison of the Battig and Montague (1969) norms with the responses of adults between the ages of 20 and 80. *Journal of Gerontology, 35*(2), 225–231.

Howard, D. V., McAndrews, M. P., & Lasaga, M. I. (1981). Semantic priming of lexical decisions in young and old adults. *Journal of Gerontology, 36*(6), 707–714.

Hunt, K. P., & Hodge, M. H. (1971). Category-item frequency and category-name meaningfulness (m′): Taxonomic norms for 84 categories. *Psychonomic Monograph Supplements, 4*(6), 97–121.

Iragui, V. J., Kutas, M., Mitchiner, M. R., & Hillyard, S. A. (1993). Effects of aging on event-related brain potentials and reaction times in an auditory oddball task. *Psychophysiology, 30*(1), 10–22.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*(12), 463–470.

Kutas, M., & Hillyard, S. A. (1982). The lateral distribution of event-related potentials during sentence processing. *Neuropsychologia, 20*, 579–590.

Kutas, M., & Iragui, V. (1998). The N400 in a semantic categorization task across 6 decades. *Electroencephalography & Clinical Neurophysiology: Evoked Potentials, 108*(5), 456–471.

Laszlo, S., & Federmeier, K. D. (2009). A beautiful day in the neighborhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory and Language, 61*, 326–338.

Light, L. L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Review of Psychology, 42*, 333–376.

Light, L. L., & Anderson, P. A. (1985). Working-memory capacity, age, and memory for discourse. *Journal of Gerontology, 40*(6), 737–747.

Logan, J. M., Sanders, A. L., Snyder, A. Z., Morris, J. C., & Buckner, R. L. (2002). Under-recruitment and nonselective recruitment: Dissociable neural mechanisms associated with aging. *Neuron, 33*(5), 827–840.

Lovelace, E. A., & Cooley, S. (1982). Free associations of older adults to single words and conceptually related word triads. *Journal of Gerontology, 37*(4), 432–437.

Madden, D. J., Langley, L. K., Denny, L. L., Turkington, T. G., Provenzale, J. M., & Hawk, T. C. (2002). Adult age differences in visual word identification: Functional neuroanatomy by positron emission tomography. *Brain and Cognition, 49*(3), 297–321.

Madden, D. J., Turkington, T. G., Coleman, R. E., Provenzale, J. M., DeGrado, T. R., & Hoffman, J. M. (1996). Adult age differences in regional cerebral blood flow during visual world identification: Evidence from H215O PET. *Neuroimage, 3*(2), 127–142.

Mattis, S. (1976). Mental status examination for organic mental syndrome in the elderly patient. In R. Bellack & B. Karasu (Eds.), *Geriatric psychiatry* (pp. 77–121). New York: Grune & Stratton.

McEvoy, C. L., & Nelson, D. L. (1982). Category name and instance norms for 106 categories of various sizes. *American Journal of Psychology, 95*(4), 581–634.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97–113.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology & Aging, 17*(2), 299–320.

Persson, J., Sylvester, C. Y., Nelson, J. K., Welsh, K. M., Jonides, J., & Reuter-Lorenz, P. A. (2004). Selection requirements during verb generation: Differential recruitment in older and younger adults. *Neuroimage, 23*(4), 1382–1390.

Petrides, M., & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia, 20*(3), 249–262.

Pfefferbaum, A., Adalsteinsson, E., & Sullivan, E. V. (2005). Frontal circuitry degradation marks healthy adult aging: Evidence from diffusion tensor imaging. *Neuroimage, 26*(3), 891–899.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America, 97*(1), 593–608.

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., & Williamson, A. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex, 15*(11), 1676–1689.

Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., & Davatzikos, C. (2003). Longitudinal magnetic resonance imaging studies of older adults: A shrinking brain. *Journal of Neuroscience, 23*(8), 3295–3301.

Roe, K., Jahn-Samilo, J., Juarez, L., Mickel, N., Royer, I., & Bates, E. (2000). Contextual effects on word production: A lifespan study. *Memory & Cognition, 28*(5), 756–765.

Ronnlund, M., Nyberg, L., Backman, L., & Nilsson, L. G. (2005). Stability, growth, and decline in adult life span development of declarative memory: Cross-sectional and longitudinal data from a population-based study. *Psychology & Aging, 20*(1), 3–18.

Salthouse, T. A. (1991). *Theoretical perspectives of cognitive aging*. Hillsdale, NJ: Lawrence Elbaum Associates.

Shapiro, S. I., & Palermo, D. S. (1970). Conceptual organization and class membership: Normative data for representatives of 100 categories. *Psychonomic Monograph Supplements, 3*(11), 107–127.

Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology & Aging, 10*(4), 527–539.

Stine-Morrow, E. A. L., Miller, L. M. S., & Nevin, J. A. (1999). The effects of context and feedback on age differences in spoken word recognition. *Journals of Gerontology Series B – Psychological Sciences & Social Sciences, 54B*(2), 125–134.

Tombaugh, T. N., Kozak, J., & Rees, L. (1999a). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology, 14*(2), 167–177.

Tombaugh, T. N., Kozak, J., & Rees, L. (1999b). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology, 14*(2), 167–177.

Uttl, B. (2002). North American Adult Reading Test: Age norms, reliability, and validity. *Journal of Clinical and Experimental Neuropsychology, 24*(8), 1123–1137.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31*(3), 443–467.

Verhaeghen, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology & Aging, 18*(2), 332–339.

Verhaeghen, P., & Cerella, J. (2008). Everything we know about aging and response times: A meta-analytic integration. In S. M. Hofer & D. F. Alwin (Eds.), *The handbook of cognitive aging: Interdisciplinary perspectives* (pp. 134–150). Thousand Oaks: Sage Publications.

Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., & Cerella, J. (2003). Aging and dual-task performance: A meta-analysis. *Psychology & Aging, 18*(3), 443–460.

Wechsler, D. (1981). *Wechsler adult intelligence scale – Revised manual*. New York: Psychological Corporation.

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin, 120*(2), 272–292.

Whiting, W. L., Madden, D. J., & Babcock, K. J. (2007). Overriding age differences in attentional capture with top-down processing. *Psychology & Aging, 22*(2), 223–232.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience, 16*(7), 1272–1288.

Wingfield, A., Stine, E. A. L., Lahar, C. J., & Aberdeen, J. S. (1988). Does the capacity of working memory change with age? *Experimental Aging Research, 14*(2), 103–107.

Zelinski, E. M., & Lewis, K. L. (2003). Adult age differences in multiple cognitive functions: Differentiation, dedifferentiation, or process-specific change? *Psychology & Aging, 18*(4), 727–745.