[This document contains the author's accepted manuscript. For the publisher's version, see the link in the header of this document.]

Paper citation:

Kemper, S., & Liu, C. – J. (2007). Eye movements of young and older adults during reading. *Psychology and Aging*, *22*, 84-94. PMID: 17385986

Abstract:

The eye movements of young and older adults were tracked as they read sentences varying in syntactic complexity. Contrary to the findings of Waters and Caplan (2001), age group differences in sentence processing were apparent but only for the most complex sentences, subject-object relative clause sentences in Experiment 1 and cleft-object and subject-object sentences missing "that" complementizers in Experiment 2. These findings suggest that eye tracking may be more sensitive to age group differences in on-line sentence processing than the auditory moving windows paradigm, that there is a threshold in processing complexity for age group differences to appear, and that working memory limitations do affect on-line sentence processing.

Text of paper:

Eye Movements of Young and Older Adults during Reading

Susan Kemper and Chiung-ju Liu University of Kansas

Address correspondence to:

Susan Kemper Gerontology Center, 3090 Dole HDC 1000 Sunnyside Ave. University of Kansas Lawrence, KS 66045 Telephone: 785 864-4131 Email: <u>SKEMPER@KU.EDU</u>

Eye Movements of Young and Older Adults in Sentence Reading

Caplan and Waters (1999) have argued that syntactic processing and other interpretive processes rely on a specialized processing system with a separate sentence-interpretation resource, unrelated to traditional span measures of working memory. The Caplan and Waters' theory (1999) predicts similar patterns of on-line processing for all readers since interpretive processes are buffered from working memory limitations. Waters and Caplan (1996, 1997, 2001) have directly examined the hypothesis that working memory limitations affect older adults' ability to process complex sentences. These studies have used the auditory moving windows paradigm (Ferreira, Henderson, Anes, Weeks, & McFarlane, 1996). This technique allows the listener to start and stop the presentation of sentence and permits the analysis of phrase-by-phrase listening times, analogous to visual moving windows paradigms which permit the analysis of word-by-word or phrase-by-phrase reading times. The studies by Caplan and Waters typically examine the processing of subject- and object-relative clause constructions, such as those below:

Object Subject Relative Clause:

The dancer found the music_{i,j} that (t_j) delighted the director.

Subject Object Relative Clause:

The music_{i, i} that the dancer found (t_i) (t_i) delighted the director.

The object subject relative clause construction imposes few processing demands on the reader or the listener, the object of the main clause, (t_i) , is also the subject of the embedded relative clause, (t_j) . The subject object relative clause construction challenges the reader or listener to assign the correct syntactic relations, the subject of the main clause, (t_j) , must also be interpreted as the object of the embedded clause, (t_i) .

Waters and Caplan (2001) compared how young and older readers allocate listening times to critical phrases of relative clause sentences. Despite age-group differences in working memory, listening times were distributed similarity by young and older listeners. All paused longer when they heard the embedded verb in the object relative clause sentences than when they heard the corresponding verb in the subject relative clause version; this additional time was attributed to the extra processing required to recover the direct object of the embedded verb. They found no evidence that differences in age or working memory lead to different processing strategies, supporting their theory.

Waters and Caplan's choice of the auditory moving window paradigm over other, more widely accepted techniques such as word-by-word reading paradigms or eye tracking paradigms is problematic. They defend the auditory moving window paradigm as "not obviously less natural" (Waters & Caplan, 2001, p. 130) than other techniques. However, it may conflict with the findings of Wingfield and his colleagues who compared young and older adults' segmentation strategies, preferred presentation times, and allocation of processing time during listening and reading tasks. Wingfield et al. (1989, 1999) showed that older adults prefer slower

speech rates but also smaller segments than young adults. Stine-Morrow et al. (1995) showed that older adults ignore clause, phrase, and sentence boundaries. Waters and Caplan segment the sentences so that they can compare listening times for words or phrases used in different constructions. Some segments are single words, some noun phrases, some a complementizer plus a noun phrase. Hence, participants do not control the length of segments or the location of segment boundaries, only the interval between the presentation of one segment and the next. It may be that this imposed segmentation conflicts with older adults' natural segmentation strategies, obscuring any difference in the remaining processing parameter, time, due to age or working memory. A task that permits participants to control both segmentation and presentation may be more sensitive to individual differences in syntactic processing than the auditory moving window paradigm.

A recent study by Kemper, Crow, and Kemtes (2004) using eye-tracking methodology reexamined these issues. Eye-tracking is a more naturalistic task that imposes few restrictions on readers; they are free to skip words or phrases, read ahead and glance backwards, and re-read entire segments. Using this technology, Kemper et al. examined three aspects of reading: first fixations to key phrases, regressions to earlier phrases, and the total time key phrases were fixated. They examined reduced relative clause sentences such as those below:

Reduced Relative Clause Sentence:

Several angry workers warned about the low wages decided to file complaints. Main Clause Sentence:

Several angry workers warned about the low wages during the holiday season. Focused Reduced Relative Clause Sentence:

Only angry workers warned about the low wages decided to file complaints.

Kemper, Crow, and Kemtes (2004) found partial support for Waters and Caplan's theory: young and older adults' first pass fixations were alike and both groups showed a clear "garden path" effect: a peak in fixation time at the second verb in reduced relative clause sentences but not at the verb in main clause sentences. This garden path effect suggests that all readers initially interpret the first verb as the main verb and must reanalyze it when they encounter the second verb in the reduced relative clause sentence. However, Kemper et al. also observed an increase in regressions and in regression path times for older readers for the reduced relative clause sentences, suggesting that older adults were unable to correctly parse these sentences. Further, low span readers, identified by their scores on a battery of working memory tests, also produced more regressions and an increase in regression path times for reduced relative clause sentences, suggesting that they were unable to correctly parse the sentences. The results from the eye-tracking analysis of the focused reduced relative clauses sentences also posed problems for Caplan and Water's theory: high span readers initially allocated additional processing time the first noun phrase and then were able to avoid the "garden path" because the focus operator "only" led them to correctly interpret the first verb phrase as a reduced relative clause.

Thus this eye-tracking study poses some challenges to Waters and Caplan's theory by revealing age group and span group differences in reading. Eye-tracking may be more sensitive to individual differences in language processing that the auditory-moving window paradigm and

may reveal subtle differences in processing strategies that other techniques miss. The present experiment used eye-tracking to compare young and older adults' processing of unambiguous clauses differing in the locus of embedding and the form of the embedded sentences that are similar to the sentence paradigms used by Waters and Caplan (2001). According to Caplan and Waters theory (1999), if sentence processing is unrelated to traditional span measure of working memory, older and young adults should show similar processing patterns regardless of working memory capacity.

Experiment 1 Method

Participants

Twenty-nine young and 39 older adults were tested in the present study. Young participants were recruited by signs posted on campus or by word of mouth through referrals. Older adults were recruited by phone solicitation from a panel of past research participants and through referrals by participants. All were native English speakers. The eye tracking system could not record the eye-movements of 4 young participants and 11 older participants due to technical problems caused by their eyeglasses or contact lens. One young participant and four older participants made more than 20% errors on the on-line processing test (see below) and were excluded from further analysis. As a result, 24 young and 24 older adults were included in the final analysis. The mean age for young adults was 20.5 years (SD = 3.1) and for older adults was 72.8 years (SD = 5.9). The mean years of education for young adults was 13.8 years (SD = 1.4) and for older adults was 15.6 years (SD = 2.6), F(1, 46) = 9.32, p = .004. The Digits Forward and Digits Backward tests (Wechsler, 1958) and the Daneman and Carpenter (1980) reading span tests were used to measure working memory capacity. The data were presented in Table 1. Young and older adults did not differ significantly on either digit span test but older adults had significantly lower reading spans than young adults; a composite (Loehlin, 1992) formed from these variables using confirmatory factor analysis did differ significantly between groups, F(1, 46) = 23.30, p < .001. Shipley's (1940) vocabulary test was used to measure vocabulary ability. Older adults had higher scores than young adults. Materials

Sentence stimuli were constructed following Waters and Caplan (2001) by varying the location and type of an embedded clause. They consisted of pairs of cleft object(CO) and cleft subject(CS) sentences and pairs of subject-object(SO) and object-subject(OS) sentences. See Table 2 for example sentences. There were 12 pairs of each type of sentence. Sentence pairs were created by switching the position, hence syntactic role, of the subjects and objects from CO to CS sentences and from SO to OS sentences. The sentences were segmented into critical regions following Waters and Caplan. For CO and CS sentences, the critical regions corresponded to the introductory phrase "it was," the first noun phrase (NP1), the verb (V), and

the second noun phrase (NP2). Note that NP1 of CS sentences is lexically identical to NP2 of the CO sentences and functions as the subject of both; NP2 of CS sentences corresponds to NP1 of CO sentences. For OS and SO sentences, the critical regions consisted of the first noun phrase (NP1), the first verb (V1), the second noun phrase (NP2), the second verb (V2), and the third noun phrase (NP3). The subject of the main clause, NP1, of OS sentences corresponds to NP2 of SO sentences and the subject of the embedded clause, NP2, of OS sentences maps onto NP1 of SO sentences.

Two lists of experimental sentences were constructed by assigning the members of each pair of sentences to different lists. In addition, there were 72 filler sentences in each list for a total of 120 sentences. There were 18 practice sentences followed by four blocks of 30 sentences. The eye tracker was re-calibrated between blocks. Participants were randomly assigned to sentence lists. One-half of the experimental sentences and one-half of the fillers were followed by probe questions. The probe questions for the experimental sentences required the participant to correctly identify the subject or object of the embedded verb. *Task and Procedure*

Participants were first given the battery of working memory tests and the vocabulary test. They were then seated before the eyetracker computer monitor. Participants sat in an adjustable chair with a head rest. They wore reading glasses if they normally did so. The chair could be raised or lowered to accommodate to bi- or tri-focal glasses. The participants also wore a visor with a small magnetic sensor attached. Each trial consisted of a fixation point centered on a blank screen for 500 msec followed automatically by the presentation of a sentence. The participants controlled presentation by pressing the mouse when they had completed reading the sentence. The sentences were presented in a 17 in flat panel computer screen at a viewing distance of 16 in. The fixation point a stimulus items were presented in white (125.5 lux) on a black background (0.03 lux) to maximize pupil size. Text was presented in Arial typeface with a mean size for individual letters of 0.57° visual angle. The participants held a computer mouse in their preferred hand which was used to control sentence presentation. Participants answered the probe questions aloud and their responses were recorded by the experimenter.

An Applied Sciences Laboratories eye tracker (Model 504) with a magnetic headtracker was used to record eye movements. Eye movements were sampled 60 times per sec with an accuracy rating of 0.5° visual angle. This translates to approximately 0.5 to 1 cm accuracy at 16 in. The headtracker noted displacements of the sensor attached to the readers' visor relative to a base unit and corrected the record of eye movements for head movements. Head movements were sampled 100 times per sec with an accuracy of 0.03° at 12 in. Stimuli were presented using GazeTracker software (Lankford, 2001) which also analyzed the eye movement data. The eyetracker was calibrated at the start of each session and between blocks for each participant. One

microcomputer controlled the eye tracker; it was interfaced with a second microcomputer running the GazeTracker software for presentation and analysis.

Two fixation measures were computed for each critical region of the sentences: the duration of the first pass fixations to the region and the regression path time for the region. In addition, first pass regressions leftward to a previously fixated region following a first pass fixation were also identified. Fixations were defined as a minimum of two sampled eye positions occurring with a fixation diameter of 30 pixels with a minimum duration of 100 msec. First pass fixation duration was defined as the summed duration of all fixations to a region beginning with the first fixation to the region and ending with first fixation rightward or leftward outside of the region. Regression path time included fixations from the first fixation to a region until the first fixation rightward outside of the region; it included re-fixations resulting from leftward regressions to a critical region. To facilitate comparison of the critical regions across pairs of CS/CO and SO/OS sentences, fixations to the "that" complementizers were excluded from analysis; hence, all comparisons examined first pass fixations, regression path times, and regressions to lexically identical segments that differed in grammatical role as a function of sentence type. Participants with a high error rate, defined as 20% or greater errors on the probe questions, were excluded from the analysis. For the remaining participants, accuracy rates were uniformly high, averaging over 90% correct, and did not differ with age group or sentence type.

Following the reading task, the participants were given a sentence acceptability judgment task modeled after that of Waters and Caplan (2001). EPRIME (Schneider, Eschman, & Zuccolotto, 2002) was used to collect the acceptability judgments and decision times. CS, CO, OS, and SO sentences were tested. One-half were meaningful sentences and one-half were meaningless ones. See Table 3 for examples. Twelve examples of each type of sentence as well as 24 fillers were tested for a total of 72 sentences. The sentences were randomly presented. Participants were instructed to read each sentence and to decide if the sentence was meaningful or not. Reaction times, from the onset of the sentence until the participant pressed a response key, were recorded along with the meaningfulness judgment.

Results

Results of the analysis of eye fixation patterns are first presented followed by the analysis of the meaningfulness judgment task. All fixations were analyzed with square root transformations to normalize distributions. The results are organized by sentence type. Separate analyses were conducted for each critical region as specified below. Lower order main effects of age group or sentence type that are subsumed by significant age group by sentence type interactions are not reported.

CO versus CS sentences

First pass fixation times and regression path times for CO and CS sentences were analyzed with 2 (age group) by 2 (sentence type) ANOVAs for each of 4 critical regions: the Introductions, the

comparison of NP1 from CS sentences with NP2 from CO sentences, the Verbs, and the comparison of NP2 from CS sentences with NP1 from CO sentences. First pass regressions were analyzed only for the Introductions and NP1_{cs} versus NP2_{co} comparison since there were no leftward regressions to the final critical region of each sentence.

First-pass fixations. The results are summarized in Figure 1. The sentence type main effect was significant for the comparison of NP1 of CS sentences (M = 473 ms, SD = 152 ms) to NP2 of CO sentences (M = 551 ms, SD = 135 ms), F(1, 46) = 13.191, p < .001, $\eta^2 = .223$, F2(1, 22) = 8.692, p < .001, $\eta^2 = .566$, and for the comparison of NP2 of CS sentences (M = 442 ms, SD = 99 ms) to NP1 of CO sentences (M = 612 ms, SD = 75 ms), F1(1, 46) = 32.327, p < .001, $\eta^2 = .413$, F2(1, 22) = 25.256, p < .001, $\eta^2 = .847$.

Regressions. The results are summarized in Figures 2. The sentence type main effect was significant for the number of leftward regressions to NP1 of CS sentences (M = 1.18, SD = .45) compared to NP2 of CO sentences (M = .51, SD = .33), F1(1, 46) = 75.949, p < .001, $\eta^2 = .623$, F2(1, 22) = 34.575 p < .001, $\eta^2 = .611$.

Regression path times. The results are summarized in Figure 3. The sentence type main effect was significant for the comparison of NP1 of CS sentences to NP2 of CO sentences, *F1*(1, 46) = 37.689, p < .001, $\eta^2 = .450$, *F2*(1, 22) = 36.607 p < .001, $\eta^2 = .433$, and for the Verb comparison, *F1*(1, 46) = 50.214, p < .001, $\eta^2 = .522$, *F2*(1, 22) = 21.409 p < .001, $\eta^2 = .493$. Both age groups had a longer regression path times to NP2 of CO sentences (M = 1403 ms, SD = 237 ms) compared to NP1 of CS sentences (M = 1086 ms, SD = 239 ms) and to the V of CO sentences (M = 819 ms, SD = 308 ms) compared to the V of CS sentences (M = 560 ms, SD = 207 ms), both $t(47) \ge 7.11$, p < .001. In addition, there was a significant age group by sentence type interaction for the comparison of NP2 of CS sentences to NP1 of CO sentences, *F1*(1, 46) = 4.990, p < .030, $\eta^2 = .098$, *F2*(1, 22) = 8.692, p < .001, $\eta^2 = .566$. Both groups had longer regression path times to NP2 of CS sentences but this difference was greater for older adults (NP1_{co}: M = 1217 ms, SD = 496 ms; NP2_{cs}: M = 730 ms, SD = 281) than for young adults (NP1_{co}: M = 954 ms, SD = 219 ms; NP2_{cs}: M = 657 ms, SD = 187). OS versus SO sentences

First pass fixation times and regression path times for OS and SO sentences were analyzed with a 2 (age group) by 2 (sentence type) ANOVAs for each of 5 critical regions NP1 of OS sentences compared to NP2 of SO sentences, Verb 1, NP2 of OS sentences compared to NP1 of SO sentences, Verb 2, and NP3. First pass regressions were analyzed for 4 critical regions since there were no leftward regressions to NP3.

First-pass fixations. The results are summarized in Figure 4. The sentence type main effects were significant for the comparison of NP2 of OS sentences to NP1 of SO sentences, *F1*(1, 46) = 12.365, *p* < .001, η^2 = .212, *F2*(1, 22) = 3.637, *p* = .063, η^2 = .073. and for Verb 2, *F1*(1, 46) = 10.351, *p* = .002, η^2 = .184, *F2*(1, 22) < 1.00. Both age groups had longer first pass fixations to NP2 (*M* = 482 ms, *SD* = 150 ms)

of SO sentences compared to NP1 (M = 382 ms, SD = 131 ms) of OS sentences, and to V2 (M = 454 ms, SD = 140 ms) of SO sentences compared to V2 (M = 329 ms, SD = 145 ms) of OS sentences.

Regressions. The results are summarized in Figure 5. There were significant age group by sentence type interactions for leftward regressions for the comparison of NP1 of OS sentences to NP2 of SO sentences, F(1, 46) = 6.016, p = .018, $\eta^2 = .116$; F2(1, 22) = 7.082, p =,014, $\eta^2 = .244$, and for the comparison of NP2 of OS sentences to NP1 of SO sentences, F(1, 46)= 4.295, p = .044, $\eta^2 = .085$; F2(1, 22) < 1.0. Both groups made more regressions to NP2 of SO sentences than to NP1 of OS sentences but this difference was greater for older adults (NP2_{SO}: M = 1.22, SD = .24; NP1_{OS}: M = .58, SD = .28) than for young adults (NP2_{SO}: M = .88, SD = .24; NP1_{OS}: M = .38, SD = .22). Likewise, both groups made more regressions to NP1 of SO sentences than to NP2 of OS sentences but this difference was greater for older adults (NP1_{SO}: M = 1.60, SD = .28; NP2_{OS}: M = .49, SD = .20) than for young adults (NP1_{SO}: M = 1.15, SD = .23; NP2_{OS}: M = .46, SD = .20). In addition, the sentence type main effect was significant for Verb1, F1(3,44) = 77.997, p < .001, $\eta^2 = .610$; F2(3,20) = 19.002, p < .001, $\eta^2 = .740$. Both groups made more regressions to V1 (M= .88, SD= .23) of OS sentences than to V1 of SO sentences (M= .33, SD= .23).

Regression path times. The results are summarized in Figure 6. The age group by sentence type interaction was significant for the comparison of NP1 of OS sentences to NP2 of SO sentences, F(1, 46) = 21.457, p = .018, $\eta^2 = .318$; F2(1, 22) = 25.984, p < .001, $\eta^2 = .542$, and for the comparison of NP2 of OS sentences to NP1 of SO sentences, F(1, 46) = 6.830, p = .020, $\eta^2 = .113$; F2(1, 22) = 3,562, p = .072, $\eta^2 = .438$. Older adults had longer regression path times to NP2 of SO sentences (M = 1754 ms, SD = 191 ms) than to NP1 of OS sentences (M = 767 ms, SD = 230), as did young adults (NP2_{SO}: M = 1144 ms, SD = 150 ms; NP1_{OS}: M = 596 ms, SD = 120); however, this difference due to sentence type was greater for older adults (difference: M = 987 ms, SD = 241) than for young adults (difference: M = 548 ms, SD = 175 ms) than to NP2 of OS sentences (M = 782 ms, SD = 260 ms) whereas regression path times for young adults did not vary with sentence type (NP1 of SO: M = 611 ms, SD = 130 ms; NP2 of OS: M = 631 ms, SD = 210 ms).

Summary

The results from the first pass fixations, leftward regressions, and regression path times indicate that young and older adults have similar processing strategies for reading CS, CO and OS relative clause sentences. Both age groups allocated additional time to decoding the subjects and objects of CO sentences than to those of CS sentences. Although young and older adults' first pass fixations to SO relative clause sentences were similar, their patterns of leftward regressions and regression path times were different. Older adults made many more

regressions to both the main clause subject (NP1) and the embedded clause subject (NP2) of SO sentences than did young adults. Consequently, older adults' regression path times for both the main clause subject and the relative clause subject of SO sentences were longer than young adults'.

Off-line sentence processing

Table 4 presents accuracy rates and reaction times for young and older adults for the off-line acceptability judgment task. Data were analyzed with a 2 (age group) X 2 (sentence acceptability) X 4 (sentence type) analysis of variance.

Accuracy rate. There were no age differences, F(1, 46) = 1.34, p > .05. Both age groups had high accuracy rates ($M_Y = 96\%$, $SD_Y = 10$; $M_0 = 97\%$, $SD_0 = 7$). There was a significant the two-way interaction of sentence type by sentence acceptability, F(3, 138) = 3.31, p < .05, $\eta^2 =$.07. Accuracy rates for acceptable CS, OS, and SO sentences were equivalent and higher than those for acceptable CO sentences; there were no differences in accuracy rates for the four types of unacceptable sentences.

Reaction times. Older adults had longer reaction times (M = 4332 ms, SD = 1761 ms) than young adults (M = 3471 ms, SD = ms), F(1, 46) = 4.94, p < .05. There was a significant two-way interaction of sentence type by sentence acceptability, F(3, 138) = 9.33, p < .05, $\eta^2 = .17$. Reaction times for acceptable sentences were ordered: CS < CO and SO = OS; reaction times for unacceptable sentences were ordered: CS < CO and OS < SO.

Discussion

This study used eye tracking to compare young and older adults' processing of unambiguous object-relative sentences and subject-relative sentences which differed in the locus of embedding and the form of the embedded sentences. Young and older adults showed similar patterns of the first pass fixation times, regression path times, and leftward regressions to critical regions for both types of cleft sentences and for object-subject relative clause sentences. However, older adults generally needed more time to process subject-object relative clause sentences than young adults; they made more regressions back to both the main clause subject and the embedded clause subject than did young adults and, consequently, their regression path times for these critical regions were longer.

These findings directly address Waters and Caplan's hypothesis (2001) that working memory and sentence processing are unrelated. They also indicate that age group differences, reflecting differences in working memory, arise for some, but not all types of complex sentences. Whereas fixation patterns of young and older adults were similar for both CS and CO sentences and OS sentences, SO sentences gave rise to marked age group differences in regressions and regression path times. CS and OS sentences can be parsed as two sequential clauses: the main clause is followed by an embedded clause signaled by a "that" complementizer which is indexed to the preceding noun phrase. CO sentences are somewhat more challenging to parse since the cleft object also serves as the object of the embedded clause and must be temporarily buffered while the embedded clause is processed. SO sentences impose yet greater demands for processing since the subject of the main clause must

also be assigned as the object the embedded clause; further, the embedded clause interrupts the main clause, so that the main clause subject must be temporarily buffered if it is to be correctly linked with its verb. It may be that there is a threshold for processing complexity such that differences due to age group, and by inference, working memory span, are not apparent until this threshold is surpassed. What is apparent is that there are differences in the size of the temporary buffer required for syntactic analysis of SO sentences, mirroring age differences in working memory as measured by traditional span measures. Compared to young adults, older adults, with smaller syntactic processing buffers, must make more regressions and allocate additional processing time to establishing the main clause subject and relative clause subject of SO sentences.

These results also indicate that not all methodologies for studying syntactic analysis are sensitive to age group differences. Auditory moving window listening times may not be sensitive to age group differences in syntactic processing because they do not reflect the active re-analysis of critical regions revealed by regressions during reading. The auditory moving window paradigm requires participants to press a button to hear a sentence phrase by phrase. (Note, in Ferreira et al. (1996), sentences were presented word by word which may increase the sensitivity of the technique.) On the other hand, the current study used an eye-tracking technique to measure on-line sentence processing. Participants were free to skip a phrase or re-read one or more phrases. Syntactic interpretation during the on-line reading task may be more immediate, postponed only when readers are unable to make an immediate assignment of syntactic roles after an initial fixation or one or more leftward regressions.

Furthermore, before participants in the Waters and Caplan study engaged in the listening task, they were given a reading span test, modeled after that of Daneman & Carpenter (1980); participants read acceptable and unacceptable sentences, made an acceptability judgment and then attempted to remember the final word of a block of sentences. A similar acceptability judgment task was used in conjunction with the auditory moving windows paradigm to assess on-line sentence processing. Hence, prior exposure to this task, which is relatively uncommon, may have enabled the participants to develop ad hoc processing strategies. Comprehension was not directly tested, rather the ability of participants to detect semantic anomalies, e.g., "It was the girl that the food nourished," which Waters and Caplan point out can be "based on fairly accessible, general semantic features" (p. 132). Participants in the current study were given separate reading span and acceptability judgment tests, the acceptability judgment task was administered after the on-line reading task, and sentence comprehension was directly tested.

Waters and Caplan (2001) included both acceptable and unacceptable sentences in the listening task; participants listened to the sentences and then made an acceptability judgment. However, we included only acceptable sentences in the on-line reading test and tested comprehension by probing for the correct identification of the subject or object of the embedded verb. Again, mixing acceptable and unacceptable sentences together during the listening task may have lead participants to develop an ad hoc processing strategy that contributed to the sentence-final processing peaks. The acceptability

judgment task, in conjunction with the implicit training provided by the prior span test, may have obscured age group differences in the allocation of listening times.

Kemper et al. (2004) found age group and span group differences in regressions and total fixation durations for reduced relative clause sentences containing temporary ambiguities. The cleft sentences and relative clause sentences used in the present experiment were unambiguous. It may be that age group and span group differences arise only for extremely demanding processing tasks , such as subject-object sentences and reduced relative clause sentences containing temporary ambiguities. One way to increase processing demands for the present types of sentences is to delete the optional "that" complementizers from CO and SO sentences. The complementizers signal the onset of the embedded clause and trigger the formation of a trace, indexed to the prior noun phrase, which is assigned as the object of the embedded clause verb in CO and SO sentences. (In English, complementizers are obligatory for CS and OS sentences.) Increasing the processing difficulty of CO and SO sentences by deleting the complementizers should give rise to age group differences in fixation patterns for CO sentences for SO sentences.

Experiment 2

A second experiment was conducted to compare eye fixation patterns by young and older adults to CO and SO sentences marked by "that" complementizers and temporarily ambiguous versions without "that" complementizers. Deleting the complementizers was expected to increase processing difficulty and, perhaps, exacerbate age group differences in fixation patterns. In CO sentences, the complementizer signals that the prior noun phrase must be temporarily buffered until required as the object of the embedded clause verb. In SO sentences, the prior noun phrase also must be buffered until it is required as the object of the embedded clause and later as the subject of the main clause. In the absence of a complementizer, the reader must detect a gap in the sentence phrase structure and find a noun phrase to fill that gap; finding the noun may require re-analysis of the sentence in order to locate a suitable candidate.

Method

Participants

Thirty young adults and 32 older adults were recruited from the same sources used in Experiment I. Excessive eye tracking failures and other technical problems resulted in excluding 5 young adults and 8 older adults. One young adult was also excluded due to excessive errors (greater than 20%) on the on-line processing task. As a result, 24 young adults and 24 older adults were included in the final analysis. The mean age for young adults was 19.79 years (*SD* = 3.3) and for older adults was 76.13 years (*SD* = 6.3). The mean years of education for young adults was 12.9 years (*SD* = 1.2) and for older adults was 15.3 years (*SD* = 2.6), *F*(1, 46) = 16.683, *p* < .001. Further information about the participants is presented in Table 1. Young and older adults did not differ significantly on the digits forward test but older adults had significantly lower digits backwards spans and reading spans than young adults; a composite (Loehlin, 1992) formed from these variables using confirmatory factor analysis did differ significantly between

groups, F(1, 46) = 8.593, p = .005. Shipley's (1940) vocabulary test was used to measure vocabulary ability. Older adults had higher scores than young adults. *Materials.*

The 12 CO and 12 SO sentences prepared for Experiment I were used as experimental items. Two versions of each CO and SO sentence were created by deleting the "that" complementizer from one version. Two lists of experimental sentences were constructed by assigning the members of each pair of sentences to different lists. In addition, there were 12 CS, 12 OS, and 72 other types of sentences used as fillers in each list for a total of 120 sentences. There were 18 practice sentences followed by four blocks of 30 sentences. Participants were randomly assigned to sentence lists. The eye tracker was re-calibrated between blocks. One-half of the experimental CO and SO sentences and one-half of the fillers were followed by probe questions. The probe questions for the experimental sentences required the participant to correctly identify the subject or object of the embedded verb. *Task and Procedure.*

The task and procedure were identical to those of Experiment I. First pass fixations and the regression path time were computed for each critical region as well as first pass regressions to previous regions. With the exception of one young adult dropped from the analysis, comprehension accuracy rates were uniformly high, averaging over 92%, and did not vary with age group or sentence type. On the sentence acceptability judgment task, CO and SO sentences with and without "that" complementizers were tested along with CS and OS sentences and a variety of filler sentences.

Results

Results of the analysis of eye fixation patterns are first presented followed by the analysis of the meaningfulness judgment task. All the fixations were analyzed with square root transformations to normalize distributions. The results are organized by sentence type. Separate analyses were conducted for each critical region as specified below. Lower order main effects of age group or sentence type that are subsumed by significant age group by sentence type interactions are not reported.

CO versus "that-less" CO sentences

First pass fixation times and regression path times for CO and "that-less" CO sentences were analyzed with a 2 (age group) by 2 (sentence type) ANOVA for 4 critical regions: the Introduction, NP1, NP2, and the V. First pass regressions were analyzed for 3 critical regions since there were no leftward regressions to the final NP.

First-pass fixations. The results are summarized in Figure 7. The sentence type main effect was significant for NP2, F1(1,46) = 5.039, p = .030, $\eta^2 = .099$; F2(1, 22) = 5.046, p = .035, $\eta^2 = .180$, and the V, F1(1,46) = 83.851, p < .001, $\eta^2 = .851$; F2(1,22) = 51.197, p < .001, $\eta^2 = .885$. Both age groups had a longer first path fixations to NP2 of "that-less" CO sentences (M = 611 ms, SD = 191 ms) compared to NP2 of CO sentences (M = 442 ms, SD = 130 ms). Both age groups had a longer first path fixations to the V of "that-less" CO sentences (M = 406 ms, SD = 84 ms) compared to the V of CO sentences (M = 332 ms, SD = 99 ms).

Regressions. The results are summarized in Figure 8. The sentence type main effect was significant for the Introduction, F1(1,46) = 65.946, p < .001, $\eta^2 = .589$; F2(1,22) = 36.559, p < .001, $\eta^2 = .624$. The sentence type by age group interaction was significant for NP1, F1(1, 46) = 56.915, p < .001, $\eta^2 = .717$; F2(1,22) = 28.143, p < .001, $\eta^2 = .728$, and NP2, F1(1,46) = 149.770, p < .001, $\eta^2 = .868$; F2(1,22) = 95.679, p < .001, $\eta^2 = .901$. Deleting the "that" complementizer had a greater effect on older adults than on young adults: older adults' regressions increased from an average of 1 to 1.7 per sentence to NP1 and from an average of 1.5 to 2.4 per sentence for NP2 when the complementizer was deleted than whereas young adults' regressions increased increased from .8 to 1.1 per sentence to NP1 and were unchanged to NP2.

Regression path times. The results are summarized in Figure 8. The age group by sentence type interaction was significant for regression path times NP2, F1(1, 46) = 1.301, p > .05; F2(1, 22) = 1.225, p > .05, and the V, F1(1,46) = 19.364, p < .001, $\eta^2 = .246$; F2(1,22) = 5.386, p = .030, $\eta^2 = .197$. Deleting the complementizer from CO sentences lead to longer regression path times for NP2 and the V; however, the increase in older adults' regression path times to NP2 (from 1300 ms to 1670 ms) and the V (from 1264 ms to 1333 ms) due to deleting the complementizer was greater than that for young adults' (from 1264 ms to 1333 ms and from 668 ms to 703 ms, respectively).

SO versus "that-less" SO sentences

First pass fixation times and regression path times for SO and "that-less" SO sentences were analyzed with a 2 (age group) by 2 (sentence type) ANOVAs for 5 critical regions, NP, V1, NP2, V2, and NP3. First pass regressions were analyzed for 4 critical regions since there were no leftward regressions to the final critical region.

First-pass fixations. The results are summarized in Figure 10. None of the main effects or interactions for age group or sentence type were significant for any critical region.

Regressions. The results are summarized in Figure 11. The age group by sentence type interactions were significant for NP1, F(1, 46) = 15.835, p < .001, $\eta^2 = .245$; F2(1, 22) = 17/332, p < .001, $\eta^2 = .441$, for V1, F1(1,46) = 21.052, p < .001, $\eta^2 = .314$; F2(1,22) = 17.045, p < .001, $\eta^2 = .437$, and NP2, F1(1,46) = 60.569, p < .001, $\eta^2 = .805$; F2(1,22) = 47.146, p < .001, $\eta^2 = .876$. Deleting the "that" complementizer led to more regressions, particularly to V1. However, this effect of deleting the complementizer was greater for older adults than for young adults. Older adults' regressions increased from an average of 1.7 per sentence to 2.3 to NP1, from .4 to 1.9 per sentence to V1, and from 1.6 to 2.1 per sentence to NP2 when the complementizer was deleted. In contrast, young adults' regressions to NP1 were unchanged and their regressions increased from .3 to 1.3 per sentence to V1 and from 1.2 to 1.3 to NP2 when the complementizer was deleted.

Regression path times. The results are summarized in Figure 8. The age group by sentence type interaction was significant for regression path times for NP1, F1(1,46) = 22.644,

 $p < .001, \eta^2 = .330; F2(1,22) = 6.856, p < .001, \eta^2 = .197, \text{ for V1}, F1(1, 46) = 4.611, p = .037, \eta^2 = .090; F2(1, 22) = 3.281, p = .084, \eta^2 = .130, \text{ and NP2}, F1(1,46) = 30.709, p < .001, \eta^2 = .741; F2(1,22) = 20.973, p < .001, \eta^2 = .815.$ The increase in regression path times due to deleting the "that" complementizer was greater for older adults for NP1 (from 556 ms to 993 ms), V1 (from 380 ms to 871 ms), and NP2 (from 1023 ms to 1780 ms) than for young adults (NP1: from 501 ms to 605 ms; V1: from 330 ms to 590 ms; NP2: from 645 ms to 1031 ms).

Summary. Deleting the complementizers increased the difficulty of both CO and SO sentences, particularly for older adults. Older adults made many more leftward regressions and had longer regression path times to critical regions of the "that-less" CO sentences than to the versions containing complementizers as they attempted to resolve the temporary ambiguities created by the missing complementizers. Deleting the complementizers also exacerbated the age group differences observed in Experiment 1 for SO sentences, resulting in further increases in regressions and further slowing regression path times.

Off-line sentence processing.

The results are summarized in Table 5. Accuracy rates and reaction times for young and older adults for the off-line sentence acceptability judgment task were analyzed with a 2 (age group) x 2 (sentence acceptability) x 2 (sentence type) x 2 (complementizer) ANOVA. There were no significant main effects or interactions for the accuracy scores; accuracy rates were high, averaging 91% for both young and older adults. There was a significant age group by complementizer interaction for the reaction times, F1(1,46) = 35.972, p < .001, $\eta^2 = .770$; F2(1,22) = 61.198, p < .001, $\eta^2 = .928$. Older adults required an additional second to respond to "that-less" sentences (M = 5708, SD = 1263) than to sentences with complementizers (M = 4726, SD = 1518) whereas young adults' reaction times (M = 3590, SD = 1111) were unaffected by deleting the complementizers.

Discussion and General Conclusions

These results confirm two implications of Experiment 1: First, there is a threshold of sentence complexity such that age group differences in syntactic processing are not apparent until this threshold is crossed. In Experiment 1 and in this Experiment, no age group differences in sentence processing were observed for CO sentences in which the embedded clause is marked by the complementizer "that." Deleting this complementizer affected both young and older adults; however, deleting the complementizer had a greater impact on older adults, leading to the appearance of age group differences in syntactic processing. Older adults, with smaller working memory buffers, made significantly more leftward regressions in order to interpret "that-less" CO sentences. As a result, their' regression path times for the second noun phrase and verb were longer for "that-less" CO sentences than for CO sentences with complementizers. The complementizer marks the presence of a relative clause and signals that

the preceding noun phrase must be temporarily retained until required as either the subject or object of the embedded verb. Deleting the complementizer meant that older adults, with smaller working memory buffers, were unable to retain this noun phrase and therefore older adults were forced to engage in additional processing when they encountered the second noun phrase and embedded verb. The increase in leftward regressions and longer regression path times for these phrases reflects older adults' additional processing as they attempted to correctly interpret the noun phrases as subject and object of the embedded verb..

Experiment 2 also confirmed a second implication from Experiment 1. Deleting the complementizers from SO sentences exacerbated the age group differences observed in Experiment 1 for this type of sentence. SO sentences, even those with complementizers, are difficult to interpret because the first noun phrase must be interpreted as both the object of the first verb and the subject of the second verb. The complementizer marks the embedded relative clause and signals that that preceding noun phrase must be retained until required as both object and subject. Deleting the complementizer increases processing difficulty, resulting in many more regressions and longer regression path times for both young and older adults. Older adults were particularly challenged to interpret "that-less" SO sentences, reflected in the further increase in their leftward regressions and regression path times as they re-read earlier parts of the sentence.

Eye tracking provides a visual trace of the strategies used by young and older adults as they process complex sentences. They reveal that older adults must engaged in more reprocessing, as revealed by leftward regressions and longer regression path times, than young adults, for some, but not all types of complex sentences. There appears to be a threshold for the appearance of age differences in syntactic processing. At least two indicators of this threshold have been traced: temporary syntactic ambiguities created by deleting complementizers and subject-embedded clauses.

The auditory moving windows paradigm and other techniques for monitoring on-line reading or listening, may not be sensitive to such age group differences in processing strategies. These techniques may force older adults to adopt artificial processing strategies to cope with the imposed segmentation, restricted opportunity for play-back, implicit pressure to respond in a timely fashion, or monitor sentences for semantic anomalies. Under more naturalistic conditions when they are listening for comprehension, older adults may seek to avoid processing problems by cuing speakers to adopt syntactic simplifications or to provide paraphrases or repetitions of complex sentences. When they are unable to do so, older adults' comprehension of complex sentences may break down whenever they encounter temporary syntactic ambiguities, missing complementizers, or subject-embedded clauses.

Acknowledgements

This research was supported by grant # RO1 AG009952 from the National Institute on Aging and by the Roy A. Roberts endowment at the University of Kansas. Correspondence concerning this article may be addressed to Susan Kemper, Gerontology Center, 3090 Dole HDC, 1000 Sunnyside, University of Kansas, Lawrence, KS 66045. Electronic mail may be addressed to SKEMPER@KU.EDU.

References

- Caplan, D., & Waters, G. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*, 114-126.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Ability, 19, 450-466.
- Ferreira, F., Henderson, J. M., Anes, M. D., Weeks, P. A., & McFarlane, D. K. (1996). Effects of lexical frequency and syntactic complexity in spoken-language comprehension: Evidence from the auditory moving-window technique. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 324-335.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68*, 1-76.
- Kemper, S., Crow, A., Kemtes, K. (2004). Eye fixation patterns of high and low span young and older adults: Down the garden path and back again. *Psychology and Aging, 19,* 157-170.
- Loehlin, J. C. (1992). *Latent variable models: An introduction to factor, path, and structural analysis*. Hillsdale, NJ: Erlbaum.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tool Inc.
- Stine, E. A. L., Cheung, H., & Henderson, D. (1995). Adult age differences in the on-line processing of new concepts in discourse. *Aging and Cognition*, *2*, 1-18.
- Warren, T., & Gibson, E. (2002). The influence of referential processing on syntactic complexity. *Cognition, 85*, 79-112.
- Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *Quarterly Journal of Experimental Psychology, 49A*, 51-79.
- Waters, G. S., & Caplan, D. (2001). Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology and Aging*, *16*,128-144.
- Wechsler, D. (1958). *The measurement and appraisal of adult intelligence*. Baltimore: Williams & Wilkins.
- Wingfield, A., & Ducharme, J. L. (1999). Effects of age and passage difficulty on listening-rate preferences for time-altered speech. *Journals of Gerontology: Psychological Sciences and Social Sciences, 54B,* 199-202.
- Wingfield, A., Lahar, C. J., & Stine, E. L. (1989). Age and decision strategies in running memory for speech: Effects of prosody and linguistic structure. *Journal of Gerontology: Psychological Sciences, 44,* 106-113.

Table 1

Means and Standard Deviations (in parentheses) for Working Memory and Vocabulary Tests for the Participants in Experiments I and II.

| | Experiment I | | | Experiment II | | | | |
|----------------|-----------------|-----------------|----------|---------------|-----------------|-----------------|----------|--------|
| | Young adults | Older adults | F(1, 46) | p | Young adults | Older adults | F(1, 46) | p |
| Digit Forward | 8.96 (2.05) | 7.83 (2.49) | 2.91 | = .095 | 8.96 (2.56) | 8.75 (2.03) | .174 | = .678 |
| Digit Backward | 7.46 (2.25) | 7.04 (2.46) | .37 | = .544 | 6.96 (1.96) | 5.29 (1.45) | 6.185 | = .017 |
| Reading Span | 3.77 (0.75) | 3.25 (0.75) | 5.76 | = .020 | 3.71 (0.99) | 2.85 (0.45) | 14.540 | < .001 |
| Vocabulary | 32.41 (4.41) | 35.88 (3.05) | 16.56 | < .001 | 30.83 (3.93) | 35.67 (2.53) | 25.654 | < .001 |

Table 2

Critical Regions and Corresponding Grammatical Roles for Pairs of CS and CO Sentences and Pairs of OS and SO Sentences.

| Sentence type | Sentence example | | | | | |
|------------------|---|--|--|--|--|--|
| CS | It was / the tailor / that altered / the suit coat. | | | | | |
| Critical Region | Intro/NP1/V/NP2 | | | | | |
| Grammatical role | Intro/SUBJ / V / OBJ | | | | | |
| СО | It was / the suit coat / that the tailor / altered. | | | | | |
| Critical Region | Intro / NP1 / NP2 / V | | | | | |
| Grammatical role | Intro / OBJ / × SUBJ / V | | | | | |
| OS | The dancer / found / the music / that delighted / the director. | | | | | |
| Critical Region | NP1 / V1 / NP2 / V2 / NP3 | | | | | |
| Grammatical role | SUBJ1 /VERB1/ SUBJ2 / × VERB2 / OBJ | | | | | |
| SO | The music / that the dancer / found / delighted / the director. | | | | | |
| Critical Region | NP1 / NP2 / V1 / V2 / NP3 | | | | | |
| Grammatical role | SUBJ1 /× SUBJ2 /VERB2/ VERB1 / OBJ | | | | | |

Note. CS - cleft subject. CO - cleft object. OS - object subject. SO - subject object. Intro - introduction. NP1 - first noun phrase. NP2 - second noun phrase. NP3 - third noun phrase. V1 - first verb. V2 - second verb. SUBJ-subject. SUBJ1 - subject of the main clause. SUBJ2 - subject of the embedded clause. VERB1 - main clause verb. VERB2 - embedded clause verb. OBJ -object. ×-"that" complementizer excluded from analysis of fixation data.

Table 3

| Sentence type | Sentence example | | |
|---------------|---|--|--|
| CS - A | It was the boy that walked the dog. | | |
| CS - U | It was the food that saw the dog. | | |
| CO - A | It was the dog that the boy walked. | | |
| CO - U | It was the dog that the food saw. | | |
| OS - A | The lady saw the child that broke the toy. | | |
| OS - U | The baby spit the mother that read the story. | | |
| SO - A | The child that the lady saw broke the toy. | | |
| SO -U | The mother that the baby spit read the story. | | |

Example Sentences used in the Off-line Processing Task.

Note. CS-cleft subject. CO-cleft object. OS-object subject. SO-subject object. A-acceptable. U-unacceptable.

Table 4

Means and Standard Deviations (in parentheses) for Accuracy Rates and Reaction Times in the Off-line Task for Experiment I.

| | СО | CS | OS | SO | |
|---------------------|-------------|-------------|-------------|-------------|--|
| Accuracy rates (%) | | | | | |
| Acceptable | | | | | |
| Young adults | 93 (16) | 95 (14) | 96 (7) | 99 (5) | |
| Older adults | 91 (16) | 99 (5) | 94 (8) | 99 (5) | |
| | | | | | |
| | | | | | |
| Unacceptable | | | | | |
| Young adults | 98 (6) | 96 (9) | 95 (9) | 94 (13) | |
| Older adults | 97 (6) | 99 (3) | 98 (6) | 99 (5) | |
| Reaction times (ms) | | | | | |
| Acceptable | | | | | |
| Young adults | 3536 (1451) | 2449 (824) | 4053 (2160) | 3507 (1600) | |
| Older adults | 4163 (1324) | 3028 (880) | 4364 (1421) | 4382 (1501) | |
| | | | | | |
| | | | | | |
| Unacceptable | | | | | |
| Young adults | 3346 (1245) | 2827 (1126) | 3375 (1608) | 4279 (1440) | |
| Older adults | 4195 (2187) | 3343 (1313) | 5177 (2571) | 6005 (2896) | |

Note. CO = cleft object; CS = cleft subject; OS = object subject; SO = subject object.

Table 5

Means and Standard Deviations (in parentheses) for Accuracy Rates and Reaction Times for the Off-line Task in Experiment II. .

| | (| 0 | SO | | |
|------------------------------|----------------------------|---------------------------|----------------------------|----------------------------|--|
| | With "that" | Without "that" | With "that" | Without "that" | |
| Accuracy rates (%) | | | | | |
| Acceptable | | | | | |
| Young adults Older adults | 93 (16) 91 (16) | 85 (14) 86 (5) | 94 (7) 94 (8) | 89 (5) 86 (5) | |
| Unacceptable | | | | | |
| Young adults Older adults | 95 (6) 97 (6) | 92 (9) 89 (3) | 95 (9) 95 (6) | 91 (13) 89 (5) | |
| Reaction times (ms) | | | | | |
| Acceptable | | | | | |
| Young adults Older adults | 3585 (1391) 4858 (1541) | 3439 (717) 5835 (753) | 3987 (1149) 4861 (1225) | 3615 (1256) 5764 (1638) | |
| Unacceptable | | | | | |
| Young adults Older adults | 3556 (975) 4395 (1874) | 3462 (874) 5413 (1214) | 3503 (1215) 4792 (1432) | 3570 (1310) 5815 (1446) | |

Note. CO = cleft object; SO = subject object.

Figure Captions

- Figure 1. First Pass Fixation Times (and SEs) for Cleft Subject (upper panel) and Cleft Object (lower panel) sentences.
- Figure 2. Leftward Regressions (and SEs) for Cleft Subject (upper panel) and Cleft Object (lower panel) sentences.
- Figure 3. Regression Path Fixation Times (and SEs) for Cleft Subject (upper panel) and Cleft Object (lower panel) sentences.
- Figure 4. First Pass Fixation Times (and SEs) for Object Subject (upper panel) and Subject Object (lower panel) sentences.
- Figure 5. Leftward Regressions (and SEs) for Object Subject (upper panel) and Subject Object (lower panel) sentences.
- Figure 6. Regression Path Fixation Times (and SEs) for Object Subject (upper panel) and Subject Object (lower panel) sentences.
- Figure 7. First Pass Fixation Times (and SEs) for Cleft Subject (upper panel) and "that-less" Cleft Subject (lower panel) sentences.
- Figure 8. Leftward Regressions (and SEs) for Cleft Subject (upper panel) and "that-less" Cleft Subject (lower panel) sentences.
- Figure 9. Regression Path Fixation Times (and SEs) for Cleft Subject (upper panel) and "that-less" Cleft Subject (lower panel) sentences.
- Figure 11. First Pass Fixation Times (and SEs) for Subject Object (upper panel) and "that-less" Subject Object (lower panel) sentences.
- Figure 11. Leftward Regressions (and SEs) for Subject Object (upper panel) and "that-less" Subject Object (lower panel) sentences.
- Figure 12. Regression Path Fixation Times (and SEs) for Subject Object (upper panel) and "that-less" Subject Object (lower panel) sentences.



Critical region











28





30













