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

Paper citation:

Abstract:

Two experiments have been completed using experimental techniques to study language production under controlled conditions. In Experiment 1, young and older adults were given 2, 3 or 4 words and asked to compose a sentence. Older adults' responses were similar to those of young adults when given 2 or 3 words. When given 4 words, the older adults made more errors and their responses were shorter and less elaborate than the young adults'. In Experiment 2, simple intransitive verbs (*smiled*), transitive verbs (*replaced*), and complement-taking verbs (*expected*) were contrasted. Older adults' responses were similar to those of young adults given intransitive and transitive verbs. Given complement-taking verbs, young adults produced complex sentences whereas the older adults produced simpler, less complex sentences and they made many errors. Both experiments found that older adults respond more slowly than young.

Text of paper:

Age Differences in Sentence Production

Susan Kemper, Ruth Herman, and Cindy Lian
University of Kansas

Address correspondence to:

Susan Kemper

Gerontology Center

3090 Dole

1000 Sunnyside

University of Kansas

Lawrence, KS 66045

Telephone: 785 864-4130

Fax: 785 864-2666

Email: Skemper@ku.edu

Age Differences in Sentence Production

Studies of elicited speech have shown that older adults produce shorter utterances using simpler syntactic structures and reduced propositional content than young adults (Kemper, Thompson, & Marquis, 2001).   However, uncontrolled age-related differences in discourse pragmatics may contribute to these findings (James, Burke, Austin, & Hulme, 1998). To date, few studies have examined these issues of aging and speech production using experimental methods that control for, e.g., lexical choice, discourse style, or the reliance on non-linguistic gestures or paralinguistic devices to convey meaning. Experimental methods that constrain these aspects of speech production have recently gained acceptance (Bock, 1996). Constrained production tasks require the speaker to formulate and utter a sentence using words or phrases presented on a computer screen or in response to stimulus pictures. Sentence formulation time can be assessed as well as various aspects of the utterance such as syntactic form and prosodic structure (Dell & O'Seaghdha, 1992; F. Ferreira, 1991; 1994, F. Ferreira & Swets, 2002; V. Ferreira, 1996; V. Ferreira & Dell, 2000; Lindsley, 1975; Roelofs, 1998; Stallings, MacDonald, & O'Seaghdha, 1998; Wheeldon & Lahiri, 1997).

Two previous studies using constrained production tasks have suggested that sentence production is well-preserved in older adults: Davidson, Zacks, & Ferreira (1996) reported that younger and older adults produce similar patterns of responses and disfluencies on a task requiring participants to complete a sentence stem with either a dative construction (I told….a story to the manager) or a double object construction (I told…the manager a story). Spieler and Griffin (2001) used a picture description task. They suggest older adults adopt a more "deliberate" response style in that older adults take longer to begin and complete sentences but show similar effects of picture codability and name frequency.

The constrained production task used in the present experiments required participants to formulate a sentence using a set of words presented on a computer screen. The words were displayed until the participant spoke, forcing participants to plan their utterance using the entire list of target words. Most prior studies of constrained sentence production have focused on response latency as an indication of sentence planning. The present experiments also compare the length, grammatical complexity, and propositional content of the responses. Experiment 1 varied the number of words presented to the participants. The length, grammatical complexity, and propositional content of the participants' responses were expected to vary with the number of words presented.   Older adults were expected to produce shorter, less complex, and less informative sentences than young adults, replicating findings from the analysis of elicited speech samples.

Experiment 1
Method

 Participants

Thirty young adults, 18 to 28 years of age (*M* = 2.30, *SD* = 1.4), and 30 older adults, 70 to 80 years of age (*M* = 74.2, *SD* = 5.4), participated. The young adults were recruited by solicitations on campus and paid $10 for participating. The older adults were recruited from a registry of previous research participants and paid a modest honorarium for participating. All participants were screened for

hearing acuity and those with clinically significant hearing loss were excluded from participation. A hearing loss was defined as (i) a greater than 40 dB hearing loss at 500, 1000, 2000, or 4000 using pure tone audiometrics or (ii) self-report of 6 or more problems on the Hearing Handicap Inventory (Ventry & Weinstein, 1982).

The participants were given a battery of cognitive tests including the Shipley (1940) vocabulary test, the Digits Forward, Digits Backwards, and Digit Symbol tests (Wechsler, 1958), the Daneman and Carpenter (1980) Reading Span test, and a Stroop test requiring participants to name the color of blocks of X's printed in colored inks or to name the color of color words printed in contrasting colored inks, e.g., RED printed in blue ink. Table 1 summarizes the performance of the participants on these tests. An alpha level of .05 was set for these and all subsequent t and F tests.

Materials

The word stimuli consisted of three sets of words, matched for word frequency using the Francis and Kucera norms (Francis & Kucera, 1982). The sets included 12 human characters, e.g., *pianist, nurse, witch*; 12 locations, e.g., *canyon, meadow, cliff*; and 12 objects, e.g., *rope, photograph, luggage*. The words were chosen to minimize semantic associations between and within sets, using the Battig and Montague (1969) and Howard (1980) norms.

Procedure

EPRIME (Schneider, Eschman, & Zuccolotto, 2002) was used to present the stimuli and collect responses. Participants were seated in front of a computer workstation equipped with a voice-activated response box connected to a microphone. They were instructed to "produce a sentence, as quickly as possible, using the words presented on the computer screen". They were instructed to "use all of the words presented" and encouraged to "add other words to make a complete, grammatical sentence." After a block of 10 practice trials, two blocks of 54 experimental trials were administered. Each trial consisted of a fixation point presented for 2 seconds followed by the presentation of 2, 3, or 4 words in a vertical column. The words remained on the screen until the participant spoke. As soon as the response box detected a vocal response, the words were removed from the computer screen. Response latencies were recorded from the onset of the response. The participant's response was audio-recorded and later transcribed. The participant initiated each trial by pressing a computer key.

The words were randomly selected and randomly ordered for presentation such that each set included at least one human character. Each participant was tested on 36 different 2-word combinations, 36 different 3-word combinations, and 36 different 4-word combinations.

Coding

Each response was initially classified as a response error or a valid response. Response errors were subcategorized as (a) non-fluent responses, trials on which the response box was triggered by a cough or a non-lexical response such as "hum;" (b) false starts, trials on which participants started fluently but then repeated one or more words; (c) responses which were sentence fragments, missing one or more obligatory constituents; and (d) memory errors, responses which failed to use all of the stimulus words presented. Sentence fragments and memory errors were assumed to reflect a breakdown in sentence planning whereas non-fluent responses and false starts were assumed to reflect articulation problems. Multiple errors could occur on a single-trial.

Using the procedures described by Turner and Greene (1977), each valid response was decomposed into its constituent propositions, which represent basic ideas and the relations between them. The number of propositions expressed in a sentence is a measure of how informative it is. DLevel or Developmental Level is an index of grammatical complexity, based on a scale originally developed by Rosenberg and Abbeduto (1987). Each valid response was assigned to 1 of 8 levels of complexity: (0) Simple, one-clause sentences; (1) Complex sentences with embedded infinitival complements, (2) Complex sentences with wh- predicate complements, conjoined clauses, and compound subjects, (3) Complex sentences with relative clauses modifying the object noun phrase or with predicate noun phrase complements, (4) Complex sentences with gerundive complements or comparative constructions, (5) Complex sentences with relative clauses modifying the subject noun phrase, subject noun phrase complements, and subject nominalizations, (6) Complex sentences with subordinate clauses, and (7) Complex sentences with multiple forms of embedding and subordination. To provide a more detailed comparison of the responses, DSS or Developmental Sentence scoring (Lee, 1974) was applied to each valid response. Eight different categories of grammatical forms are scored for each sentence: indefinite pronouns, personal pronouns, main verbs, secondary (embedded verbs), conjunctions, negatives, and two types of questions. Within each category, variants are assigned different points to reflect the developmental order of appearance in children's speech. A total score is derived for each sentence by summing the points for each category plus 1 point if the sentence is fully grammatical. Sentence length in words was also determined. Inter-coder reliability was assessed for each level of coding. Reliability averaged better than 90% for all levels of coding. Coded examples of the participants' responses are presented in Table 2.

Results

Response errors and valid responses were analyzed separately.

Errors

Table 3 summarizes the error results. Young adults made one or more errors on 10% of the trials whereas older adults made one or more errors on 16% of the trials. False starts occurred on less than 1% of all trials and were not analyzed.

Non-Fluent responses. The percentage of trials on which the participant failed to make a fluent response was analyzed with a 2 age group by 3 set size (2, 3, or 4 words presented) ANOVA. The main effect of age group was significant, $F(1, 58) = 7.11$, $p = .01$, $\eta^2 = .11$. The main effect of set size was significant, $F(2, 57) = 50.52$, $p < .01$, $\eta^2 = .47$, as was the age group by set size interaction, $F(2, 57) = 11.43$, $p < .01$, $\eta^2 = .16$. Older adults produced more non-fluent responses than young adults and non-fluent responses by older adults increased with set size.

Sentence Fragments. The percentage of trials on which the participant produced a sentence fragment missing one or more obligatory constituents was analyzed with a 2 age group by 3 set size (2, 3, or 4 words presented) ANOVA. The main effect of age group was significant, $F(1, 58) = 6.92$, $p = .01$, $\eta^2 = .11$. The main effect of set size was significant, $F(2, 57) = 30.59$, $p < .00$, $\eta^2 = .35$, as was the age group by set size interaction, $F(2, 57) = 10.07$, $p < .01$, $\eta^2 = .15$. Older adults produced more sentence fragments than young adults. As set size increased, older adults produced more fragments.

Memory errors. The percentage of trials on which the participant failed to use all words presented was analyzed with a 2 age group by 3 set size (2, 3, or 4 words presented) ANOVA. The main

effect of age group was significant, $F(1, 58) = 13.63$, $p < .01$, $\eta^2 = .19$. The main effect of set size was significant, $F(2, 57) = 41.15$, $p < .01$, $\eta^2 = .42$, as was the age group by set size interaction, $F(2, 57) = 11.88$, $p < .01$, $\eta^2 = .17$. Older adults committed memory errors more often than did young adults, particularly when set size = 4.

Valid Responses

All valid responses were subjected to analysis of their length, propositional content, and grammatical complexity as well as response latency. The responses were classified by set size (2, 3, or 4 words presented). Initially, word order was included as a design factor, contrasting trials on which a human characters was the first word presented with locative-first and object-first trials, reflecting the strong bias in English for animate-first sentences (Bock, Loebell, & Morey, 1992; McDonald, Bock, & Kelly, 1993). The word order factor was not significant in any of the analyses reported below nor did it interact with set size or age group.

Sentence length. The overall main effect of age group was not significant for the sentence length measure, $F(1, 58) < 1.0$; sentence length differed as a function of set size, $F(2, 57) = 187.79$, $p < .01$, $\eta^2 = .77$, and the set size by age group interaction was significant, $F(2, 57) = 4.30$, $p = .02$, $\eta^2 = .07$. Figure 1a summarizes these findings. The length of young adults' responses increased monotonically with set size. Older adults' responses were similar in length to young adults for set size = 2 or 3; young adults produced longer responses than the older adults when set size = 4.

Propositional content. The overall main effect of age group was not significant for the propositional content measure, $F(1, 58) < 1.0$; however, propositional content did differ as a function of set size, $F(2, 57) = 24.74$, $p < .01$, $\eta^2 = .30$, and the set size by age group interaction was significant, $F(2, 57) = 4.26$, $p = .03$, $\eta^2 = .07$. Figure 1b summarizes these findings. Young adults tended to produce more propositions as set size increased; older adults tended to limit their responses to 2 or 3 propositions. Older adults produced as many propositions as young adults with set size = 2 or 3; young adults produced more propositions with set size = 4.

DLevel. The overall main effect of age group was not significant for the DLevel measure, $F(1, 58) < 1.0$. DLevel did differ as a function of set size, $F(2, 57) = 32.89$, $p < .01$, $\eta^2 = .36$, and the set size by age group interaction was significant, $F(2, 57) = 4.85$, $p < .01$, $\eta^2 = .08$. Figure 1c summarizes these findings. DLevel scores for young and older adults were similar when set size = 2 or 3; older adults produced sentences with lower DLevel scores when set size = 4.

DSS. The overall main effect of age group was not significant for the DSS measure, $F(1, 58) < 1.0$. DSS did vary as a function of set size, $F(2, 57) = 53.04$, $p < .01$, $\eta^2 = .72$, and the set size by age group interaction was significant, $F(2, 57) = 6.30$, $p < .01$, $\eta^2 = .10$. Figure 1d summarizes these findings. DSS scores for young adults increased monotonically with set size. DSS scores for young and older adults were similar when set size = 2 or 3; older adults produced sentences with lower DSS scores when set size = 4.

Response Latency. The latency to produce a fluent sentence using all words did vary with age group, $F(1, 58) = 12.33$, $p < .01$, $\eta^2 = .17$. Response latency increased with set size, $F(2, 57) = 146.556$, $p < .01$, $\eta^2 = .72$, and the set size by age group interaction was significant, $F(2, 57) = 7.32$, $p < .01$, $\eta^2 = .11$.

Figure 1e summarizes these findings. Older adults responded more slowly than young adults, the latency to respond increased with set size, and this increase was greater for older adults than for young adults.

Discussion

Since the stimulus display terminated as soon as the participant began to speak, the task imposed a memory load on the participants as they planned their response. This memory load affected sentence planning and articulation by older adults: both non-fluent responses and memory errors increased with the number of words presented. In contrast, young adults had little difficulty retaining all of the stimulus words and set size did not affect sentence planning or execution.

The length, grammatical complexity, and propositional density of the valid responses by older adults provide evidence of an effect of memory load on sentence production. Even on trials on which the older adults were able to retain and use all of the words presented in a fluent response, their responses when set size = 4 were shorter, less complex, and less informative than those produced by young adults. However, when set size = 2 or 3, older adults were able to produce sentences that matched those of young adults' in length, grammatical complexity, and content.

Response latencies for both young and older adults increased with set size. This pattern suggests that the speakers attempted to pre-plan their utterances before speaking; hence, as set size increased, they preplanned more segments or longer segments incorporating additional words.

Experiment 2

Experiment 2 was designed to investigate whether a linguistic manipulation of verb type would affect sentence production by young and older adults. Verb choice affects sentence planning processes by imposing restrictions on the use of direct objects or predicate complements (Holmes, 1982). Three types of verbs were contrasted: simple intransitive verbs, e.g., *smiled*, *jumped*, transitive verbs preferentially taking noun phrase direct objects, e.g., *called*, *replaced*, and complement-taking verbs which optionally can take noun phrase direct objects but which more commonly occur with sentence complements, e.g., *wished*, *guessed*. Producing a sentence with a complement-taking verb requires the speaker to either (i) formulate two clauses, a main clause and a complement such as a that-clause or wh-clause or (ii) use the verb with a direct object or adjunct phrase. For example, the complement-taking verb *wished* commonly occurs with a that-complement, e.g., *wished she was married*; it can be used with an adjunct, e.g., *wished for a husband*, but this usage is less common (Ferreira & Dell, 2000). The prediction was that older adults would have more difficulty producing multi-clause sentences using the complement-taking verbs than young adults, consistent with observational studies that indicate an age-related decline in the use of multi-clause sentences (Kemper, et al., 2001).

Method

Participants

Thirty young adults, 18 to 28 years of age (*M* = 19.8, *SD* = 2.1), and 30 older adults, 70 to 80 years of age (*M* = 75.1, *SD* = 4.8), were recruited and tested using similar procedures to those used in Experiment 1, including screening for significant hearing loss and possible dementia. Characteristics of the participants are given in Table 1 An additional group of 4 young adults and 6 older adults was tested.

These participants produced error responses on more than 25% of the trials and their data was excluded from analysis.

Materials

The word stimuli consisted of five sets of words, matched for word frequency using the Francis and Kucera norms (1982). The sets included 12 human characters, e.g., *pianist, nurse, witch*; 12 locations, e.g., *canyon, meadow, cliff*; and 18 verbs. The verbs included 6 intransitive verbs, e.g., *laughed, smiled, looked*, which typically do not occur with direct objects; 6 transitive verbs, e.g., *copied, examined, replaced*, which typically require direct objects; and 6 complement-taking verbs, e.g., *guessed, suggested, realized*, which preferentially require sentential complements (Ferreira & Dell, 2000). All were regular past tense verbs ending in *–ed*.

Procedure

EPRIME was used to present the stimuli and collect responses using a similar procedure to that followed in Experiment 1. After a block of 10 practice trials, two blocks of 54 experimental trials were administered. Each trial consisted of a presentation of 1 agent plus a verb and optionally a location. The words were randomly selected from each stimulus set and randomly ordered for presentation. Each participant was tested on 18 2-word combinations with each type of verb and 18 agent-locative-verb combinations with each type of verb.

Coding

The participants' responses were transcribed and coded using the same procedures followed in Experiment 1 to classify response errors and valid responses. Response errors included non-fluent responses, false starts, memory errors, and responses which were sentence fragments. In addition a new error response category was used: substitutions. On some trials, participants substituted a different form of the verb, substituting, e.g., progressives such as *laughing*, or nominals such as *laughter*, for the correct form, *laughed*. Valid responses were coded for the number of words in the sentence, the number of propositions, DLevel, and DSS. Inter-coder reliability was assessed for each level of coding. Reliability averaged better than 90% for all levels of coding. Coded examples of the participants' responses are presented in Table 2.

Results

Response errors and valid responses were analyzed separately.

Errors

Table 4 summaries the error findings. Young adults made one or more errors on 5% of the trials whereas older adults made one or more errors on 12% of the trials. False starts occurred on less than 1% of all trials and were not analyzed. Substitutions of adjectives and other parts of speech for the verbs occurred on less than 1% of responses from young adults and 4% of responses from older adults and were not analyzed.

Non-Fluent responses. The percentage of trials on which the participant failed to make a fluent response was analyzed with a 2 age group by 2 set size (2 or 3 words presented) by 3 verb type (intransitive, transitive, complement-taking) ANOVA.

The main effect of age group was significant, $F(1,58) = 11.38$, $p < .01$, $\eta^2 = .16$. The main effect of set size was significant, $F(2,57) = 134.02$, $p < .01$, $\eta^2 = .69$, as was that for verb type, $F(2,57) = 13.23$, $p < .01$, $\eta^2 = .19$. In addition, the set size by age group interaction, $F(2,57) = 20.63$, $p < .01$, $\eta^2 = .26$, and the three-way interaction of set size, verb type, and age group, $F(2,57) = 5.12$ $p < .01$, $\eta^2 = .08$, were significant. Older adults produced more non-fluent responses overall than young adults. For older adults non-fluent responses increased with verb type, but only when an agent and location were presented along with the verb. Older adults produced non-fluent responses on 17% of the trials when 2 words were presented along with a complement-taking verb.

Sentence fragments. The percentage of trials on which the participant produced a sentence fragment missing one or more obligatory constituents was analyzed with a 2 age group by 2 set size (2 or 3 words presented) by 3 verb type (intransitive, transitive, complement-taking) ANOVA.

The main effect of age group was significant, $F(1,58) = 4.14$, $p = .05$, $\eta^2 = .07$. All main effects and interactions were significant, including the three-way interaction of set size, verb type, and group, $F(2,57) = 3.13$, $p = .05$, $\eta^2 = .05$. Older adults produced more fragments than young adults, especially when set size = 3. Fragments varied with verb type, particularly for older adults who produced fragments on 12% of the trials when an agent and a location were presented along with a complement-taking verb.

Memory errors. The percentage of trials on which the participant failed to use all words presented was analyzed with a 2 age group by 2 set size (2 or 3 words presented) by 3 verb type (intransitive, transitive, complement-taking) ANOVA.

The main effect of age group was significant, $F(1,58) = 23.05$, $p < .01$, $\eta^2 = .28$. All main effects and interactions were significant, including the three-way interaction of set size, verb type, and age group, $F(2,57) = 4.04$, $p = .02$, $\eta^2 = .06$. Young adults failed to use all of the words presented on 3% of the trials. Older adults failed to use all of the words presented on 15% of the trials, and such memory errors increased with set size, particularly when complement-taking verbs were presented.

Valid Responses

All valid responses were subjected to analysis of their length, propositional content, and grammatical complexity as well as response latency. The responses were classified by set size (2 or 3 words presented) and verb type (intransitive, transitive, complement-taking). Initially, word order was included as a design factor, contrasting trials on which a human character was the first word presented with locative-first and object-first trials. The word order factor was not significant in any of the analyses reported below nor did it interact with set size, verb type, or age group.

Sentence length. The overall main effect of age group was not significant for the sentence length measure, $F(1,58) < 1.53$, $p = .22$, $\eta^2 = .01$; sentence length did differ as a function of set size, $F(2,57) = 204.32$, $p < .01$, $\eta^2 = .78$, and verb type, $F(2,57) = 26.38$, $p < .01$, $\eta^2 = .31$. Figure 2a summarizes these findings. The participants produced longer sentences when they were given two words in addition to a verb to incorporate into a sentence than when they were given only an agent. They produced longer

sentences when they were given complement-taking verbs than when they were given intransitive or transitive verbs.

Propositional content. The overall main effect of age group was not significant for the propositional content measure, $F(1, 59) < 1.0$; however, propositional content did differ as a function of set size, $F(2, 57) = 48.08$, $p < .01$, $\eta^2 = .45$, and the set size by age group interaction was significant, $F(2, 57) = 25.26$, $p < .01$, $\eta^2 = .30$. Propositional content varied with verb type, $F(2, 57) = 31.52$, $p < .01$, $\eta^2 = .35$, and the verb type by age group interaction was significant, $F(2, 57) = 21.63$, $p < .01$, $\eta^2 = .27$. In addition, the three-way interaction of verb type, set size, and age group was significant, $F(2, 57) = 4.08$, $p = .02$, $\eta^2 = .07$. Figure 2b summarizes these findings. Young adults produced more propositions for complement-taking verbs than for intransitive and transitive verbs, particularly when they were given two words in addition to the verb. Older adults limited their sentences to approximately 3 propositions, regardless of verb type or set size.

DLevel. Although the overall main effect of age group was not significant for the DLevel measure, $F(1, 58) < 1.0$, DLevel did differ as a function of set size, $F(2, 57) = 26.26$, $p < .01$, $\eta^2 = .31$, and verb type, $F(2, 57) = 167.79$, $p < .01$, $\eta^2 = .74$. Figure 2c summarizes these findings. DLevel scores for all participants increased with set size and with verb type. When intransitive or transitive verbs were presented, DLevel scores averaged around 1, indicating the participants produced mostly simple, 1-clause sentences. DLevel scores for complement-taking verbs were higher, averaging 3 to 4 points, indicating that the participants were in fact producing sentences with predicate complements.

DSS. Again, the overall main effect of age group was not significant for the DSS measure, $F(1, 58) = 1.90$, $p = .17$, $\eta^2 = .01$. DSS did vary as a function of set size, $F(2, 57) = 110.36$, $p < .01$, $\eta^2 = .66$, but the set size by age group interaction was not significant, $F(2, 57) = 1.50$, $p = .22$, $\eta^2 = .01$. DSS varied with verb type, $F(2, 57) = 13.38$, $p < .01$, $\eta^2 = .19$, the verb type by set size interaction was significant, $F(2, 57) = 7.98$, $p = .01$, $\eta^2 = .12$, and the verb type by age group interaction was significant, $F(2, 57) = 8.76$, $p < .01$, $\eta^2 = .13$. The three-way interaction was significant, $F(2, 57) = 13.37$, $p < .01$, $\eta^2 = .19$. Figure 2d summarizes these findings. DSS scores for young adults increased with set size, particularly when complement-taking verbs were presented. Older adults' responses tended to be limited to 7 to 10 DSS total points, even when complement-taking verbs were presented.

Response Latency. The latency to produce a fluent sentence using all words did vary with age group, $F(1, 58) = 5.51$, $p = .02$, $\eta^2 = .09$. Response latency increased with set size, $F(2, 57) = 95.76$, $p < .01$, $\eta^2 = .62$, and the set size by age group interaction was significant, $F(2, 57) = 5.82$, $p = .02$, $\eta^2 = .09$. Response latency increased with verb type, $F(2, 57) = 18.54$, $p < .01$, $\eta^2 = .24$, and the verb type by age group interaction was significant, $F(2, 57) = 5.06$, $p < .01$, $\eta^2 = .08$. The three-way interaction was not significant, $F(2, 57) = 1.53$, $p = .22$, $\eta^2 = .01$. Figure 2e summarizes these findings. Older adults responded more slowly than young adults; the latency to respond increased with the number of words to be incorporated into the sentence and this increase was greater for older adults than for young adults. Response latency was longer when complement-taking verbs were presented than when intransitive and transitive verbs were presented, particularly for older adults when set size = 3.

Discussion

Experiment 2 provides evidence that the linguistic manipulation of verb type affects constrained sentence production by older adults. They produced more non-fluent responses and memory errors when complement-taking verbs were specified than when intransitive or transitive verbs were specified, suggesting that the linguistic manipulation affected both sentence planning and execution. Older adults' limited to 3 propositions and 7 to 10 DSS points per sentence using complement-taking verbs compared to the sentences with 4 to 5 propositions and 10 to 15 DSS points for young adults. Whereas young adults were able to incorporate 1 or 2 words into a sentence using a complement-taking verb, older adults were able to do so on less than 70% of the trials. And they did so by limiting the propositional content and syntactic complexity of their responses.

Conclusion

Previous studies using constrained production tasks have found basic syntactic processing to be well-preserved in older adults (Davidson, et al., 1996; Spieler & Griffin, 2001). The current findings, in part, are consistent with these studies. Error rates and measures of syntactic complexity, sentence length, and propositional content were similar for older and young adults when they were given 2 or 3 words to use in a sentence in Experiment 1 or when verbs were limited to intransitive or transitive forms in Experiment 2.

These experiments using a constrained production task demonstrate two manipulations that affect older adults' sentence production: (i) Directly manipulating memory load by increasing the number of words to be used in a sentence and (ii) Manipulating the linguistic characteristics of the words presented while holding memory load constant. In Experiment I, when they were given 4 words to use in a sentence, older adults' produced sentences that were shorter, simpler, and less informative than those produced by young adults. In Experiment 2, older adults' responses using a complement-taking verb shorter, simpler, and less informative than those produced by young adults.

These results support earlier findings, derived from the analysis of elicIted speech samples, that aging affects sentence production in response to explicit manipulation of memory load or the manipulation of key linguistic factors such as verb type. Verb type, particularly complement-taking verbs, appear to impose an implicit memory load on older adults, forcing them to use shorter, simpler sentences. Other linguistic manipulations, such as the dative/double object alternation examined by Davidson et al. (1996), e.g., *I told a story to the manager/the manager a story*, may not increase memory load and, hence, not differentially affect young and older adults' ability to plan and produce fluent, grammatical sentences. Although constrained production tasks impose artificial requirements on sentence production, they provide a degree of experimental control over pragmatic factors that is lacking in the analysis of spontaneous or elicited speech samples. By examining how aging affects the time course of sentence production under controlled conditions, it may be possible to distinguish other cognitive and linguistic factors that affect older adults' ability to generate long, fluent, complex, and informative sentences.

References

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

Bock, K. (1996). Language production: Methods and methodologies. *Psychonomic Bulletin & Review, 3*, 395-421.

Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review, 99*, 150-171.

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

Davidson, D., Zacks, R. T., & Ferreira, F. (1996). *The formulation of English datives: Syntactic choice and aging*. Presented at the Cognitive Aging Conference, Atlanta, GA.

Dell, G. S., & O'Seaghdha, P. (1992). Stages of lexical access in language production. *Cognition, 42*, 287-314.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language, 30*, 210-233.

Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language, 33*, 210-233.

Ferreira, F., & Swets, B. (2002). How incremental is language production: Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language, 46,* 57-84

Ferreira, V. (1996). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language, 35*, 724-755.

Ferreira, V., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology, 40*, 296-340.

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

Holmes, M. (1982). Sentence planning units: Implications for the speaker's representation of meaningful relations underlying sentences. In J. Bresan (Ed.), The mental representation of grammatical relations (pp. 797-827). Cambridge, MA: MIT Press.

Howard, D. (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, 225-231.

James, L. E., Burke, D. M., Austin, A., & Hulme, E. (1998). Production and perception of "verbosity" in younger and older adults. *Psychology and Aging, 13*, 355-368.

Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging*, 16, 312-322.

Kemper, S., Thompson, M., & Marquis, J. (2001). Longitudinal change in language production: Effects of aging and dementia on grammatical complexity and propositional content. *Psychology and Aging, 16,* 600-614*.*

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

Kemper, S., Herman, R., & Lian, C. (2003). Age differences in sentence production. *Journals of Gerontology: Psychological Sciences, 58,* 260-269. Publisher's official version: http://dx.doi.org/10.1093/geronb/58.5.P260. Open Access version: http://kuscholarworks.ku.edu/dspace/.

Lee, L. (1974). *Developmental sentence analysis*. Evanston, IL: Northwestern University Press.

Lindsley, J. R. (1975). Producing simple utterances: How far ahead do we plan? *Cognitive Psychology, 7*, 1-19.

McDonald, J. L., Bock, K., & Kelly, M. H. (1993). Word and word order: Semantic, phonological, and metrical determinates of serial position. *Cognitive Psychology, 25*, 188-230.

Pfeiffer, E. (1975). A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. *Journal of the American Geriatrics Society, 23*, 433-441.

Roelofs, A. (1998). Rightward incrementality in encoding simple phrasal forms in speech production: Verb-particle combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 904-921.

Rosenberg, S., & Abbeduto, L. (1987). Indicators of linguistic competence in the peer group conversational behavior of mildly retarded adults. *Applied Psycholinguistics, 8*, 19-32.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime User's Guide. Pittsburgh: Psychology Software Tools Inc.

Spieler, D. H., & Griffin, Z. M. (2001). Aging and the scope of word preparation during language production. Presented at the Cognitive Aging Conference, Atlanta, GA.

Stallings, L. M., MacDonald, M., & O'Seaghdha, P. (1998). Phrasal ordering constraints in sentence production: Phrase length and verb disposition in heavy-NP shift. *Journal of Memory and Language, 39*, 392-417.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117-152). New York: Academic Press.

Turner, A., & Greene, E. (1977). *The construction and use of a propositional text base*. Boulder, CO: University of Colorado Psychology Department.

Ventry, I. M., & Weinstein, B. E. (1982). The Hearing Handicap Inventory for the Elderly: A new tool. *Ear and Hearing, 3*, 128-134.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence*. Baltimore: Williams & Wilkins.

Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language, 37*, 356-381.

Author Note

Table 1.

Characteristics of the Participants in Experiments 1 and 2. Means and Standard Deviations (in parentheses) are given.

Characteristic	Experiment 1				Experiment 2			
	Young Adults	Older Adults	$F(1,58)$	$p =$	Young Adults	Older Adults	$F(1,58)$	$p =$
Education	14.2 (1.4)	15.3 (2.5)	4.280	.043	14.9 (1.5)	14.8 (2.1)	0.018	.857
Vocabulary	31.6 (4.1)	34.4 (4.1)	7.350	.009	32.6 (3.7)	35.9 (2.9)	3.837	.000
Digits Forward	9.0 (2.1)	8.4 (2.6)	0.785	.387	9.5 (2.2)	7.2 (1.4)	4.892	.000
Digits Backwards	7.5 (2.5)	7.6 (2.2)	0.012	.913	7.4 (1.4)	6.3 (2.0)	2.006	.050
Reading Span	3.6 (1.0)	3.2 (1.0)	4.652	.035	3.4 (0.6)	3.0 (0.3)	3.131	.003
Digit Symbol	35.0 (5.1)	23.2 (4.9)	83.000	.000	35.3 (7.7)	22.6 (4.5)	77.838	.000
Stroop Xs	90.6 (8.3)	70.8 (12.8)	51.059	.000	85.5 (9.8)	73.0 (10.7)	4.713	.000
Stroop words	68.2 (13.8)	38.1 (10.6)	89.976	.000	56.2 (13.1)	39.9 (9.5)	5.501	.000

Table 2

Responses coded for sentence length in words, propositions, DLevel, and DSS. Presented words are underlined.

Response	words	DSS	DLevel	Propo-sitions
Experiment I				
The *preacher* condemned the *waiter.*	5	4	1	1
The *nurse carried* a crop as she rode the horse in the *meadow*.	13	20	6	4
The *waiter* laughed at the *photograph*.	6	3	1	2
The *witch* watched the *prisoner* when he fell off the *cliff*.	11	20	6	4
The blond dog left the *fireplace* to follow the *coach* to his *jeep*.	13	8	1	6
The angry *passenger* holding the *rope* claimed the ship was sinking at the *dock*.	14	16	7	6
The *pianist* threw the *luggage* out on the *lawn* and the reporter helped.	13	8	2	4
Experiment II				
The *nurse smiled* because the *reporter* was ill on the *dock*.	11	15	6	4
The *sheriff called* the *singer* from the *closet.*	8	4	1	2
The *reporter asked* the *butcher* about the hay growing in the *canyon.*	12	10	6	4

Table 3

Response errors by young and older adults. The percentage of trials on which each type of error occurred is given. Multiple errors could occur on a single trial.
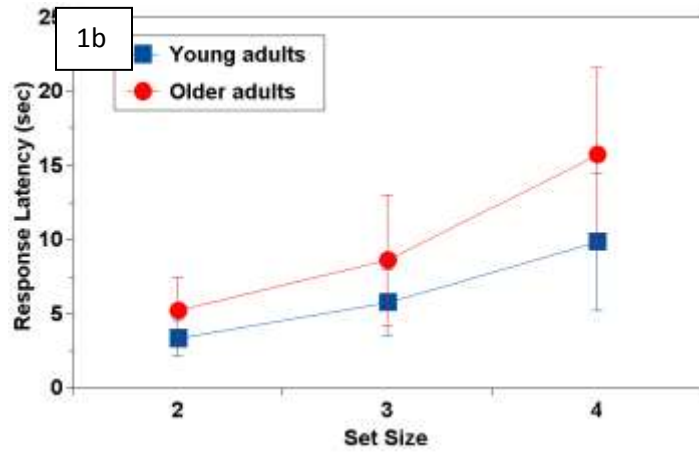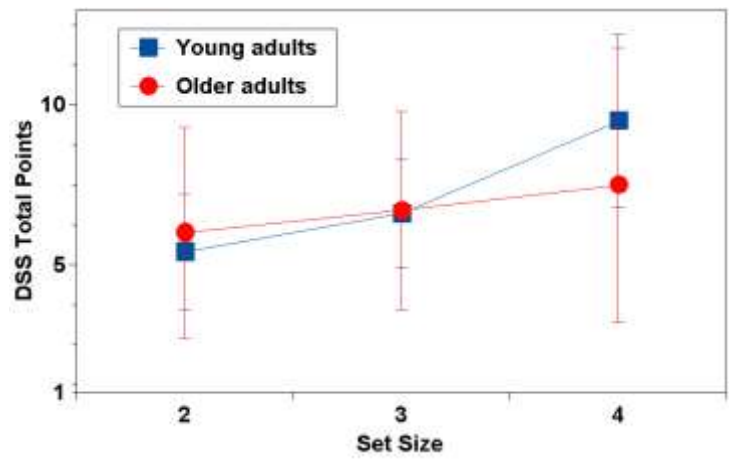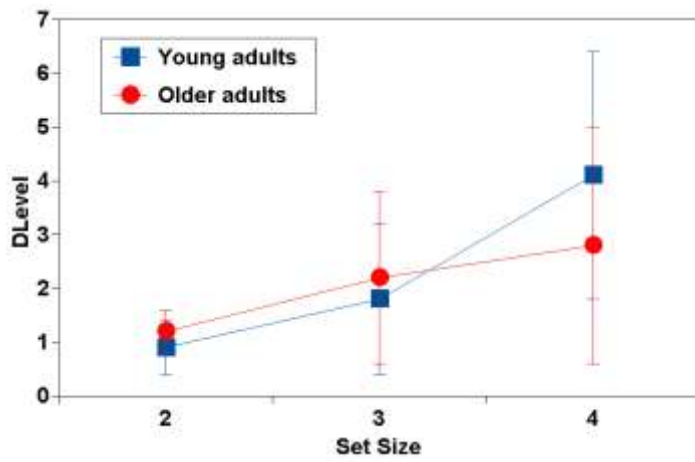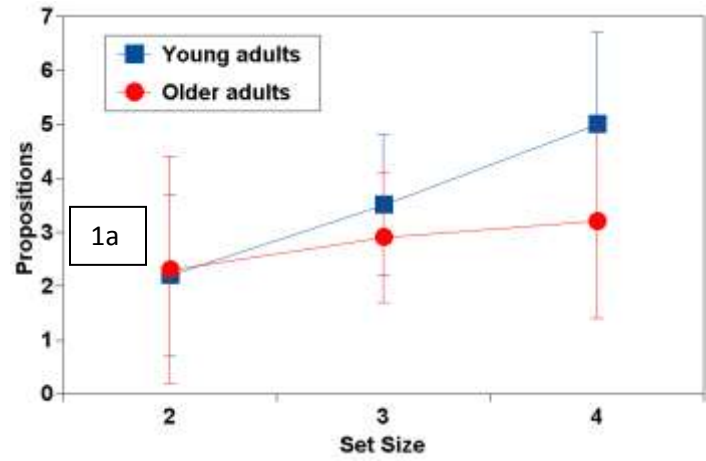
Error	Age group	Set size = 2	Set size = 3	Set size = 4
Non-Fluent responses	Young adults	2%	4%	6%
	Older adults	2	5	11
Sentence Fragments	Young adults	1	1	2
	Older adults	1	3	6
Memory errors	Young adults	<1	2	3
	Older adults	1	4	9

Table 4

Response errors by young and older adults. The percentage of trials on which each type of error occurred is given. Multiple errors could occur on a single trial.

Error	Age group	Agent Only			Agent + Location		
		Intransitive Verb	Transitive Verb	Complement-taking Verb	Intransitive Verb	Transitive Verb	Complement-taking Verb
Non-Fluent Responses	Young adults	1	3%	6%	8%	9%	11%
	Older adults	3	3	3	10	13	17
Sentence Fragments	Young adults	2	2	2	5	8	8
	Older adults	2	2	2	9	10	12
Memory Errors	Young adults	<1	<1	1	4	6	7
	Older adults	2	3	3	12	13	16

Figure Captions

Figure 1: Fluent responses of young and older adults were scored for length in words (1a), propositions (1b), DLevel (1c), and DSS points (1d). Response latencies are also shown (1e). Error bars represent $\pm$ 1 SE.

Figure 2: Fluent responses of young and older adults were scored for length in words (2a), propositions (2b), DLevel (2c), and DSS points (2d). Response latencies are also shown (2e). Error bars represent $\pm$ 1 SE.