Kemper, S., Herman, R.E., Liu, C. J. (2004). Sentence production by younger and older adults in controlled contexts. *Journals of Gerontology: Psychological Sciences, 58B*, P220-P224. Publisher's official version: http://dx.doi.org/10.1093/geronb/59.5.P220.Open Access version: http://kuscholarworks.ku.edu/dspace/.

Paper citation:

Kemper, S., Herman, R.E., Liu, C. J. (2004). Sentence production by younger and older adults in controlled contexts. *Journals of Gerontology: Psychological Sciences, 58B*, P220-P224.

Abstract:

This experiment compared young and older adults' abilities to produce complex sentences under controlled conditions. Participants were asked to memorize sentence stems differing in syntactical complexity and then to produce a complete sentence using the stem. The length, complexity, and content of young adults' responses varied with the syntactic complexity of the stems whereas the older adults' responses did not. These results suggest that working memory processing limitations impose a "ceiling" on older adults' production of complex sentences, limiting their length, complexity, and content.

Text of paper:

Sentence Production by Young and Older Adults in Controlled Contexts

Susan Kemper, Ruth E. Herman, Chiung-Ju Liu

University of Kansas

Address correspondence to:

Susan Kemper
Gerontology Center
1000 Sunnyside Ave., Rm 3090
University of Kansas
Lawrence, KS 66045
Telephone: 785 864-4130
Fax: 785 864-2666
Email: Skemper@ku.edu

Kemper, S., Herman, R.E., Liu, C. J. (2004). Sentence production by younger and older adults in controlled contexts. *Journals of Gerontology: Psychological Sciences, 58B*, P220-P224. Publisher's official version: http://dx.doi.org/10.1093/geronb/59.5.P220.Open Access version: http://kuscholarworks.ku.edu/dspace/.

Sentence Production by Young and Older Adults in Controlled Contexts

Most studies of aging and language production have compared oral or written language samples elicited from young and older adults (Kemper, Thompson, & Marquis, 2001). In contrast, controlled tasks have been widely used to assess psycholinguistic constraints on production (Bock, 1996). These tasks can be used to examine how aging affects production. For example, Kemper, Herman, and Lian (2003) found that older adults took longer, made more errors, and produced less complex sentences than young adults, and older adults' production difficulties increased with the number of words they were given to incorporate in their sentences. One limitation of the previous study is that participants rarely produced the sorts of complex, multi-clause constructions that have been the focus of naturalistic studies of speech. The present task was designed to probe participants' abilities to produce sentences involving complex constructions under controlled conditions.

The task modified the sentence generation task used by Kemper et al. (2003) and combined it with a task used by Ferreira (1991) to induce participants to generate complex sentences. Ferreira required participants to memorize a sentence and then repeat it when cued. The latency to respond was a function of the syntactic complexity of the sentence. This procedure was revised by requiring participants memorize sentence fragments or stems, not complete sentences; it was also revised to directly examine sentence planning processes by requiring the participants complete the stems. Memorization times, response latencies, and response errors were analyzed; in addition, sentence completions were scored for length, propositional content, and grammatical complexity.

Method

Participants Thirty-four young adults, 18 to 28 years of age, and 39 older adults, 70 to 80 years of age, participated. The young adults were recruited by solicitations on campus. The older adults were recruited from a registry of previous research participants.. All participants were paid a modest honorarium. Participants were required to produce at least 4 of 6 fluent responses in each experimental condition. Data from 9 young adults and 9 older adults were excluded for excessive errors. The final group of participants consisted of 24 young adults (M = 19.7, SD = 1.58) and 24 older adults (M = 72.6, SD = 2.31). 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 Hz using pure tone audiometrics or (ii) self-report of 6 or more problems on the Hearing Handicap Inventory (Ventry & Weinstein, 1982). One young adult and 6 older adults were excluded for hearing loss. All older adults were screened for possible dementia with the Short Portable Mental Status Questionnaire (Pfeiffer, 1975). The Digits Forward, Digits Backward, and Digit Symbol tests (Wechsler, 1958), the Daneman and Carpenter (1980) reading span test, the Shipley (1940) vocabulary test, and two versions of a Stroop test were administered to all participants. The Stroop tests required 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 summarized these comparisons of young and older adults. An alpha level of .05 was set for these and all subsequent *t* and *F* tests.

Materials

The stimuli consisted of 2 versions of each of 36 stems, left- or right-branching complements. All content words in the stems common nouns and verbs (10 or more occurrences per million; Francis & Kucera, 1982). The stems began as main clauses consisting of a subject, specified by a proper name, and verb. Each stem was revised to create a pair consisting of a right-branching, sentence-final complement stem and a left-branching, sentence-initial complement stem. All stems were 3 words in length. In addition, a set of 27 nouns referring to human characters, e.g., dentist, butcher, waiter, and 27 nouns referring to locations, e.g., kitchen, office, store, were also selected for use in the set size manipulation. All were common nouns (10 or more occurrences per million words).

Procedure

EPRIME (Schneider, Eschman, & Zuccolotto, 2002) was used to present the stimuli and collect responses . The stems were randomly ordered for presentation such that each participant was tested on 18 sentence-initial and 18 sentence-final stems, but on only 1 item from each pair of stems. Stems were presented with 0, 1 (a character or a location), or 2 (a character and a location) additional words to be incorporated in the sentence. The additional words were randomly selected for presentation. Each trial consisted of a fixation point presented for 2 seconds followed by the presentation of a stem. The participant was instructed to memorize the stem and told that they would be required to produce a sentence beginning with the stem. When the participant had memorized the stem, the participant pressed a response key. The time to memorize the stem was automatically recorded. When the participant pressed the response key, 0, 1 or 2 additional words would appear along with a cue to respond. If no additional words were presented, the response cue was presented alone. The participants were instructed to "produce a sentence, as quickly as possible, using the stem and any additional words presented on the computer screen". They were further reminded that their sentence should begin with the stem and they were instructed to "use all of the words presented" and encouraged to "add other words to make a complete, grammatical sentence." The words remained on the screen until the participant spoke into a microphone. As soon as a response box connected to the microphone 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.

Coding

Each response was initially classified as a response error or a valid response. Response errors were subcategorized as (i) non-fluent responses with lexical or non-lexical fillers, or false starts,; (ii) responses which were anomalous or meaningless; (iii) sentence fragments; and (iv) memory errors including incorrect or partial recall of the stem or stimulus words. Multiple errors could occur on a single-trial.

Valid responses were coded for propositional density (PDensity) (computed as the number of propositions expressed in the sentence divided by the number of additional words used in the sentence times 10), grammatical complexity in terms of Developmental Level (DLevel) and Developmental Sentence Scoring (DSS), and length in words (the number of additional words used in the sentence) using procedures described in Kemper et al. (2003). Coded examples are given in Table 2. Inter-coder reliability was assessed for each level of coding. Reliability averaged better than 90% for all levels of coding.

Results

Valid responses, memorization times, response latencies, and errors were analyzed by a 2 age group by 3 set size by 2 locus of embedding mixed ANOVA. Memorization times and response latencies were transformed to log RTs to correct for highly skewed distributions.

Valid Responses

Sentence length. The overall main effect of age group was significant for sentence length, $F(1,28) = 128.34$, $p < .01$, $\eta^2 = .27$; sentence length differed as a function of set size, $F(2,27) = 7.43$, $p < .01$, $\eta^2 = .36$, and locus of embedding, $F(1,28) = 13.40$, $p = .01$, $\eta^2 = .32$. In addition, the age group by locus of embedding interaction was significant, $F(1,28) = 5.57$, $p = .05$, $\eta^2 = .17$. Young adults produced longer responses with right-branching stems ($M = 9.8$, $SD = .3$) than left-branching stems ($M = 7.5$, $SD = .6$) whereas older adults' responses did not vary in length with locus of embedding ($M = 6.6$, $SD = .4$).

PDensity. The overall main effect of age group was significant for PDensity, $F(1,28) = 157.57$, $p < .01$, $\eta^2 = .65$; PDensity differed as a function of set size, $F(2,27) = 8.55$, $p < .01$, $\eta^2 = .39$, and the set size by age group interaction was significant, $F(2,27) = 5.51$, $p = .01$, $\eta^2 = .21$. Young adults produced more propositions as set size increased, increasing from PDensity = 4.2 ($SD = .2$) with set size = 0 to PDensity = 6.4 ($SD = .3$) with set size = 2; older adults limited their sentences to an average of PDensity = 4.4 ($SD = .3$) regardless of set size. Further, PDensity varied with locus of embedding, $F(1,28) = 6.41$, $p = .02$, $\eta^2 = .19$, and the locus of embedding by age group interaction was significant, $F(1,28) = 4.91$, $p = .02$, $\eta^2 = .27$. Young adults produced more propositions for right--branching stems (PDensity = 6.5, $SD = .2$) than for left-branching stems (PDensity = 4.3, $SD = .2$) whereas older adults averaged PDensity = 4.4 ($SD = .3$) regardless of the locus of embedding.

DLevel. The overall main effect of age group was not significant for DLevel, $F(1,28) < 1.0$. DLevel did not differ overall as a function of set size, $F(1,28) < 1.0$; however, the age group by set size interaction was significant, $F(2,27) = 4.79$, $p < .02$, $\eta^2 = .26$. For young adults, DLevel scores increased from an average of 4.7 ($SD = .2$) for set size = 0 to an average of 5.2 ($SD = .2$) for set size = 2. DLevel scores for older adults did not vary with set size (M = 3.7, SD = .3). DLevel scores also varied with locus of embedding, $F(1,27) = 758.336$, $p < .01$, $\eta^2 = .96$. As expected, DLevel scores for left-branching stems were systematically higher ($M = 5.0$, $SD = .1$) than those for right-branching stems ($M = 2.1$, $SD = .28$), reflecting the scoring system.

DSS. The overall main effect of age group was significant for DSS, $F(1,28) = 21.54$, $p < .01$, $\eta^2 = .44$. DSS varied as a function of set size, $F(2,27) = 9.64$, $p < .01$, $\eta^2 = .42$, and the set size by age group interaction was significant, $F(2,27) = 10.70$, $p < .01$, $\eta^2 = .44$. DSS scores for young adults increased monotonically with set size from 7.3 ($SD = .6$) points per sentence with set size = 0 to 13.5 ($SD = .8$) points per sentence with set size = 2. DSS scores for older adults' did not vary with set size, averaging 7.2 ($SD = .6$) points per sentence. In addition, DSS scores varied with the locus of embedding, $F(1,28) = 10.27$, $p < .01$, $\eta^2 = .27$, and the locus of embedding by age group interaction was significant, $F(1,28) = 19.89$, $p < .01$, $\eta^2 = .59$. DSS scores for young adults were higher for right-branching stems ($M = 11.4$, $SD = .5$) than for left-branching stems (*M* = 9.3, *SD* = .5) whereas DSS scores for older adults did not vary with the locus of embedding (*M* = 7.2, *SD* = .6).

Memorization Time. The latency to memorize the stem did not vary with age group, locus of embedding, or set size. Participants required an average of 3.2 s (SD = .2) to memorize the stems.

Response Latency. The latency to produce a fluent sentence using the stem and all additional words did vary with age group, $F(1,28) = 25.04$, $p < .01$, $\eta^2 = .48$. Response latency increased with set size, $F(2,27) = 12.67$, $p < .01$, $\eta^2 = .49$. Young and older adult's latencies increased with set size, set size 1 = 3.2 s (*SD* = .3), set size 2 = 3.6 (*SD* = .2), set size 3 = 4.4 (*SD* = .3). More interestingly, the age group by locus of embedding interaction was significant, $F(1,28) = 7.73$, $p < .01$, $\eta^2 = .48$. Older adults (*M* = 11.38, *SD* = .49) responded more slowly than young adults (*M* = 3.8 s, *SD* = .2) and older adults responded more slowly to left-branching stems (*M* = 4.4, *SD* = .3) than to right-branching stems(*M* = 13.2, *SD* = .3). Young adults' response latencies did not vary with locus of embedding.

Response errors

Young adults made 1 or more errors on 10% of the trials whereas older adults made 1 or more errors on 26% of the trials.

Non-fluent responses. Both main effects for age group and locus of embedding as well as the age group by locus of embedding interaction, $F(1,28) = 15.39$, $p < .01$, $\eta^2 = .36$, were significant. Older adults (*M* = 16.0%, *SD* = 4.5) produced more non-fluent responses than young adults (*M* = 2.0, *SD* = 1.5) and non-fluent responses by older adults were more common in response to left-branching stems (*M* = 20.3, *SD* = 5.5) than in response to right-branching stems (*M* = 11.4, *SD* = .5).

Sentence Fragments. The main effect of age group was significant, $F(1,28) = 4.44$, $p = .05$, $\eta^2 = .14$, as was the age group by locus of embedding interaction, $F(1,28) = 52.56$, $p < .01$, $\eta^2 = .65$. Older adults produced more sentence fragments (*M* =15.5 %, *SD* = 4.1) than young adults (*M* = 13.0, *SD* = 2.4). More sentence fragments occurred in response to left-branching stems (*M* = 20.1, *SD* = 7.9) than in response to right-branching stems (*M* = 8.5, *SD* = 3.0).

Anomalous sentences. The main effects for locus of embedding and age group were significant as was the age group by locus of embedding interaction, $F(1,28) = 10.72$, $p < .01$, $\eta^2 = .28$. More anomalous sentences were produced in response to left-branching stems (*M* = 10.4%, *SD* = 5.2) than in response to right-branching stems (*M* = 4.2, *SD* = 3.3) by older adults; young adults produced few anomalous sentences in response to either type of stem (*M* = 1.1, *SD* = 1.4).

Memory errors. The main effect of locus of embedding was significant, $F(1,28) = 18.59$ $p < .01$, $\eta^2 = .39$. More memory errors were committed in response to left-branching stems (*M* = 8.6%, *SD* = 2.8) than in response to right-branching stems (*M* = 3.4%, *SD* = 3.7).

Discussion

The length, grammatical complexity and propositional content of the young adults' valid responses varied with locus of embedding, hence syntactic complexity. They were longer, more informative, and more complexity when the young adults were given right-branching stems to complete than when they were given left-branching stems to complete, particularly when they were given 2 or 3 additional words to use in their response. This finding is consistent with theoretical arguments and experimental demonstrations of the asymmetries between left- and right-branching constructions (Gibson, 1998). Left-branching constructions impose a greater burden on working memory during production because they require the speaker to anticipant and plan for the main clause while producing the embedded clause.

The length, content, and complexity of older adults' responses did not vary with locus of embedding. Indeed, the older adults' responses appeared to be limited to 6.6 additional words, PDensity = 4.4, DLevel = 3.7, and 7.2 DSS points regardless of the type of stem or the number of additional words. Their response latencies were slower for left-branching than right-branching stems and increased when they were given 2 or 3 additional words to incorporate into their response. This pattern of results suggests that the increased memory load imposed by the left-branching stems and by the additional words impaired older adults' ability to produce valid responses.

The present study, like those of Kemper, Herman, and Lian (2003) suggests that there is a "ceiling" on older adult's speech at DLevel = 3.7 and PDensity = 4.4 since the complexity and content of older adults' sentence does not rise above these averages when they are asked to complete sentence stems, even left-branching stems. Processing limitations, arising from reduced working memory, inhibitory deficits, or slowed processing speed, may impose this "ceiling" by limiting older adults' abilities to construct complex, informative sentences, even during controlled production tasks.

References

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

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

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

Gibson, E. (1998). Syntactic complexity: Locality of syntactic dependencies. *Cognition, 68*, 1-76.

Kemper, S., Herman, R. E., & Lian, C. (2003). Age differences in sentence production. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, *58*, 260-269.

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.

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.

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

Shipley, W. C. (1940). A self-administered scale for measuring intellectual impairment and deterioration. *Journal of Psychology, 9,* 371-377.

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.

Kemper, S., Herman, R.E., Liu, C. J. (2004). Sentence production by younger and older adults in controlled contexts. *Journals of Gerontology: Psychological Sciences, 58B*, P220-P224. Publisher's official version: http://dx.doi.org/10.1093/geronb/59.5.P220.Open Access version: http://kuscholarworks.ku.edu/dspace/.

Author Note

This research was supported by grant RO1AG09952 from the National Institute on Aging. We thank Jennifer Nartowitz for her assistance. Requests for reprints may be addressed to: S. Kemper, 1000 Sunnyside Ave., Rm. 3090, Gerontology Center, University of Kansas, Lawrence, KS 66045 USA.

Table 1. Characteristics of the Participants

Characteristic	Young Adults	Older Adults	F(1, 28)	p
Education	15.2 (2.54)	16.1 (2.8)	<1.0	>.50
Vocabulary	33.0 (3.3)	36.9 (1.8)	16.67	<.001
Digits forward	9.5 (2.2)	8.8 (2.1)	<1.0	>.50
Digits backward	9.1 (2.4)	6.8 (1.6)	9.67	<.001
Reading span	3.7 (0.6)	3.1 (0.5)	8.83	=.001
Digit symbol	33.8 (4.4)	24.8 (3.8)	35.09	<.001
Stroop X blocks per 45 s	88.7 (11.3)	75.4 (16.6)	6.43	=.020
Stroop words per 45 s	66.6 (10.6)	42.6 (12.1)	33.37	<.001
Interference (Xs _ words/Xs) × 100	25.2 (8.7)	43.5 (6.4)	52.48	<.001

Note: Means are given in the table; standard deviations are shown parenthetically.

Table 2. Responses Coded for Sentence Length in Words, DSS, DLevel, Propositions, and PDensity

Response	Words	DSS	DLevel	Propositions
Right-branching stems				
Robert ordered that a pizza be delivered.	4	9	5	3
Henry cut what he was holding at the office.	5	11	5	3
Anne took what the nurse was wearing and put it in the closet.	8	19	7	5
Tom saw who was robbing the store.	4	9	3	2
George thought that he would remodel his kitchen.	5	5	5	2
Left-branching stems				
That Joan stole the jewelry was a surprise.	5	9	5	2
What Billy found was money at the store.	4	9	5	3
Whom Mary ordered to go to the library was me.	6	5	7	3
What Henry cut was the tent on the cliff.	5	9	5	3
That Alice said she saw the columnist in the park wasn't true.	8	18	7	4

Notes: The stem is italicized and any additional presented words are underlined. The measures of sentence length in words, DSS points, and PDensity do not include the stems. PDensity ¼ propositional density; DLevel ¼ developmental level; DSS ¼ developmental sentence scoring.

Kemper, S., Herman, R.E., Liu, C. J. (2004). Sentence production by younger and older adults in controlled contexts. *Journals of Gerontology: Psychological Sciences, 58B*, P220-P224. Publisher's official version: http://dx.doi.org/10.1093/geronb/59.5.P220.Open Access version: http://kuscholarworks.ku.edu/dspace/.

Table 3. Linguistic Characteristics of the Participants' Responses to the Left- and Right-Branching Stems as well as Stem Memorization Times and Response Latencies

	Left-Branching Stems: Additional Words			Right-Branching Stems: Additional Words		
Characteristic	0	1	2	0	1	2
Young Adults						
Sentence length (words)	6.9 (.4)	7.2 (.4)	8.5 (.4)	8.6 (.4)	9.3 (.5)	11.5 (.6)
PDensity	2.6 (.3)	5.0 (.4)	6.5 (.3)	5.8 (.3)	6.3 (.3)	7.5 (.4)
DLevel	6.0 (.2)	6.2 (.2)	6.5 (.2)	3.4 (.3)	3.8 (.2)	3.9 (.3)
DSS	6.8 (.6)	9.6 (.5)	11.8 (.8)	7.8 (.5)	11.3 (.6)	15.1 (.7)
Memorization time (s)	3.0 (.2)	3.2 (.2)	3.3 (.3)	3.2 (.3)	3.4 (.5)	3.2 (.4)
Response latency (s)	2.8 (.3)	2.9 (.3)	3.3 (.4)	2.4 (.3)	3.0 (.3)	3.2 (.4)
Older adults						
Sentence length (words)	6.1 (.3)	6.5 (.4)	7.3 (.4)	6.2 (.4)	6.4 (.4)	7.6 (.5)
PDensity	4.2 (.3)	4.4 (.2)	4.6 (.3)	4.4 (.3)	4.2 (.3)	4.4 (.3)
DLevel	3.7 (.3)	3.7 (.4)	4.2 (.3)	3.1 (.3)	3.5 (.2)	3.6 (.5)
DSS	6.8 (.5)	6.8 (.6)	7.2 (.6)	7.2 (.6)	7.4 (.5)	7.5 (.6)
Memorization time (s)	3.0 (.2)	3.2 (.2)	3.2 (.2)	3.2 (.3)	3.0 (.3)	4.4 (.4)
Response latency (s)	3.6 (.3)	4.2 (.2)	5.5 (.3)	2.8 (.2)	3.0 (.2)	3.7 (.4)

Notes: PDensity ¼ propositional density; DLevel ¼ developmental level; DSS ¼ developmental sentence scoring. Means are given in the table; standard deviations are shown parenthetically.