## Tracking Sentence Planning and Production

Susan Kemper,1 Daniel Bontempo,2 Whitney McKedy,3 RaLynn Schmalzried,4 Bruno Tagliaferri,2 and Doug Kieweg2

1Department of Psychology and Gerontology Center, 2Center for Biobehavioral Neurosciences in Communication Disorders, 3Gerontology Center and 4Department of Psychology, University of Kansas, Lawrence.

### Objective:
To assess age differences in the costs of language planning and production.

### Methods:
A controlled sentence production task was combined with digital pursuit rotor tracking. Participants were asked to track a moving target while formulating a sentence using specified nouns and verbs and to continue to track the moving target while producing their response. The length of the critical noun phrase (NP) as well as the type of verb provided were manipulated.

### Results:
The analysis indicated that sentence planning was more costly than sentence production, and sentence planning costs increased when participants had to incorporate a long NP into their sentence. The long NPs also tended to be shifted to the end of the sentence, whereas short NPs tended to be positioned after the verb. Planning or producing responses with long NPs was especially difficult for older adults, although verb type and NP shift had similar costs for young and older adults.

### Discussion:
Pursuit rotor tracking during controlled sentence production reveals the effects of aging on sentence planning and production.

### Key Words:
Cognition—Dual task demands—Language—Psycholinguistics—Speech.
they were paired with complement verbs. Planning time and response time increased with NP length.

In the present experiment, a controlled production task similar to that used by Stalling and colleagues was combined with pursuit rotor tracking. Pursuit rotor tracking has been shown to be sensitive to age group differences in language processing (Kemper, Schmalzlried, Herman, & Mohankumar, in press-b; Kemper, Schmalzlried, Hoffman, & Herman, in press-c; Kemper, Schmalzlried, Herman, Leedahl, & Mohankumar, 2008). Pursuit rotor tracking provides a continuous measure synchronized with the sequence of experimental events to separate out effects of sentence planning from those of sentence production. We expected older adults to show increased planning and production costs, reflecting declines in working memory and other cognitive resources tapped by sentence planning and production. We expected older adults to be more sensitive to heavy-NP shift than young adults, producing more shifted structures overall, if heavy-NP shift reflects an accommodation to working memory constraints on either sentence planning or production. We were especially interested in comparing the effects of verb type and NP length for young versus older adults in order to determine if older adults are more sensitive to these psycholinguistic manipulations.

We also investigated how cognitive abilities affect sentence planning and production. Vocabulary knowledge (MacDonald & Christiansen, 2002; Stine-Morrow, Soederberg Miller, Gagne, & Hertzog, 2008), working memory (MacDonald, Just, & Carpenter, 1992), processing speed (Stine, Wingfield, & Poon, 1986; Stine-Morrow, Loveless, & Soederberg, 1996), and inhibitory control (Hasher & Zacks, 1988) have been shown to affect language processing. Although vocabulary knowledge increases over the life span (Verhaeghen, Steitz, Sliwinski, & Cerella, 2003), most models of cognitive aging assume that working memory, processing speed, and inhibitory control decline (Park et al., 2002). These cognitive abilities may affect different aspects of sentence planning and production: Increased vocabulary knowledge may imply greater familiarity with low-frequency structures such as sentential complements, longer NPs, or shifted NPs (Roland, Dick, & Elman, 2007); decreased processing speed may provide an overall disadvantage during planning and production (Rabaglia & Salthouse, in press); decreased working memory capacity may affect the planning of complex sentences (Swets, Desmet, Hambrick, & Ferreira, 2007), especially those with long or shifted NPs; and poor inhibitory control may fail to suppress alternative syntactic constructions (Engelhardt, Corley, Nigg, & Ferreira, 2010), hence the likelihood of shifting NPs. We examined whether these cognitive abilities were moderators of age group differences in planning or producing long NPs or shifted NPs.

Methods

Participants

Forty young and 40 older adults were tested. The young adults were recruited by signs posted on campus and class announcements, whereas the older adults were recruited from a database of prospective and previous research participants. The participants were paid for their participation. Participants were given a battery of tests of working memory, processing speed, inhibition, and vocabulary. Table 1 summarizes their performance; an alpha level of .05 was set for these and all subsequent tests.

Three vocabulary measures were collected. On the Shipley (1940) Vocabulary Test, participants choose the best synonym from four choices, and the number correct (maximum = 40) is the outcome. On the North American Reading Test (AmNART; Grober & Sliwinski, 1991), participants read aloud a series of irregularly spelled words, and the number correctly pronounced words (maximum = 50) is the outcome. Finally, educational attainment in years served as a third indicator of vocabulary.

Three working memory measures were collected. On the Digits Forward and Digits Backwards tests (Wechsler, 1958), participants repeated strings of numbers either in the same (forward) or in the reverse (backward) order as presented. String length increased from 2 to 12 digits. Two strings at each length were given to the participants, and the longest string length at which the participant could remember all the digits (maximum = 12) determined their Forward or Backward Span. On the Daneman and Carpenter (1980) Reading Span Test, participants were asked to remember the last word of each sentence in a set; the number of sentences per set, hence the number of words to be remembered, increased from two to seven. Two sets at each length were given to participants, and the longest set length at which a participant could recall all the final words for both sets (maximum = 7) determined their Reading Span.

Four processing speed measures were collected. In the Digit Symbol Test (Wechsler, 1958), participants pair symbols to digits using a key. The number correct pairing within 45 s is the outcome. On the baseline condition of the Stroop Test (Stroop, 1935), participants had 45 s to name the color of the ink of a series of X’s printed in 5 columns of 25 rows on a page, and number correct (maximum = 125) served as the outcome. On the Trails A portion of the Trail Making Test (Reitan, 1958), participants connected 18 labeled dots in numerical order, and the total time in seconds required to correctly connect the dots is the outcome. Finally, the asymptotic rotor speed attained by the participant following practice (see subsequently) was used as an additional measure of processing speed.
For working memory, processing speed, and vocabulary, a summary composite was formed using maximum likelihood estimation, and factor scores were generated for each participant. For each composite, the respective factor analysis found a single latent factor with moderately high to high loadings for each indicator measure. The eigenvalues from the vocabulary, working memory, and processing speed factor models were, respectively, 1.91, 1.32, and 2.69. Loadings obtained for the vocabulary composite were Shipley ($\lambda = .81$), AmNART ($\lambda = .90$), and educational attainment ($\lambda = .66$). Loadings obtained for the working memory composite were Digits Forward ($\lambda = .62$), Digits Backward ($\lambda = .73$), and Reading Span ($\lambda = .63$). Loadings obtained for the processing speed composite were Digit Symbol ($\lambda = .86$), Stroop baseline ($\lambda = .78$), Trails A total seconds ($\lambda = -.87$), and asymptotic rotor speed ($\lambda = .76$). Factor scores were generated using Stata (StataCorp, 2009).

Lastly, the Stroop and Trail Making Tests were also used to derive two measures of inhibition. First, in addition to the baseline Stroop condition, participants were given a second condition requiring them to note the ink color of color names (e.g., green ink used to print the color name RED), printed in 5 columns of 25 names on a page. A Stroop interference score was then calculated as Stroop Interference = (blocks of Xs - color names)/blocks of Xs. Second, in addition to the Trails A test, on the Trails B test, participants connected 18 labeled dots in sequential order, alternating between letters and numbers (1-A-2-B-3-C and so on). A Trail Making interference score was calculated as Trail Making Interference = (seconds Trails A - seconds Trails B)/seconds Trails A. Because only two measures of inhibition were available, the Stroop and Trail Making interference scores were averaged for each participant to create a summary measure.

The four composites were weakly correlated ($r[80]$ for vocabulary with working memory = .07, with inhibition = .19, with speed = -.22; speed with inhibition = .03, with working memory = .39; working memory with inhibition = .22). Reliability estimates are reported in Table 1.

Table 1. Comparison of Young and Older Participants

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Young adults</th>
<th>Older adults</th>
<th>F(1, 78)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Age (years)</td>
<td>21.0</td>
<td>2.4</td>
<td>76.0</td>
</tr>
<tr>
<td>Vocabulary composite (.83)</td>
<td>–0.40</td>
<td>0.56</td>
<td>0.40</td>
</tr>
<tr>
<td>Years of education</td>
<td>14.7</td>
<td>1.8</td>
<td>16.0</td>
</tr>
<tr>
<td>AmNART (.88)</td>
<td>30.9</td>
<td>4.4</td>
<td>35.7</td>
</tr>
<tr>
<td>Shipley Vocabulary (.72)</td>
<td>31.1</td>
<td>4.9</td>
<td>34.9</td>
</tr>
<tr>
<td>Processing Speed composite (.89)</td>
<td>0.68</td>
<td>0.47</td>
<td>–0.68</td>
</tr>
<tr>
<td>Stroop X (.76)</td>
<td>84.8</td>
<td>12.1</td>
<td>67.8</td>
</tr>
<tr>
<td>Digit Symbol (.78)</td>
<td>31.9</td>
<td>4.9</td>
<td>23.1</td>
</tr>
<tr>
<td>Trail Making A (.86)</td>
<td>48.7</td>
<td>14.0</td>
<td>83.4</td>
</tr>
<tr>
<td>Working Memory composite (.70)</td>
<td>0.27</td>
<td>0.82</td>
<td>–0.27</td>
</tr>
<tr>
<td>Digits Forward (.64)</td>
<td>9.6</td>
<td>2.4</td>
<td>8.9</td>
</tr>
<tr>
<td>Digits Backward (.68)</td>
<td>7.9</td>
<td>2.1</td>
<td>6.9</td>
</tr>
<tr>
<td>Daneman and Carpenter (.56)</td>
<td>3.8</td>
<td>1.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Inhibition composite (.54)</td>
<td>0.03</td>
<td>0.17</td>
<td>0.03</td>
</tr>
<tr>
<td>Stroop Interference (.64)</td>
<td>–25.1</td>
<td>8.9</td>
<td>–45.9</td>
</tr>
<tr>
<td>Trail Making Interference (.76)</td>
<td>–18.3</td>
<td>3.4</td>
<td>–38.9</td>
</tr>
</tbody>
</table>

Notes: Alternate-forms reliability was assessed for the Digits Forward, Digits Backward, Digit Symbol, Stroop, and Trail Making tests; split-half reliability was assessed for the AmNART and Shipley tests; Cronbach’s alphas are reported for the four composites. Reliability estimates for the composites and individual measures are reported in parentheses. AmNART = North American Reading Test.

* $p < .05$; ** $p < .01$.

Task and Design

Following Stalnaggs and colleagues, each sentence consisted of three elements: a subject phrase, an NP, and a PP. The subject phrase consisted of a subject such as “Janet” and a verb, either a transitive verb such as “transferred” or a complement verb such as “revealed.” Short NPs, for example, “some plans,” were contrasted with long NPs created by adding additional modifiers to the short NP, for example, “some more specific plans for the brand new defense plant.” The PPs consisted of a preposition and an object, for example, “to Lee.” Sets of sentences were created with four alternatives contrasting transitive versus intransitive verbs and long versus short NPs sharing a common subject and a common PP; 40 sets of sentences were created (see Table 2). Pairs of transitive and complement verbs were matched for word frequency using the norms of Francis and Kucera (1982); long NPs were 13 or more ($M = 16.4$) syllables in length, whereas short NPs were 2–4 syllables ($M = 3.8$) in length. On each trial, the three elements were presented at the top, middle, and bottom of the computer screen centered within the pursuit rotor track. The subject phrase was always presented in the middle and was underlined. The positions of the NP and PP above versus below the subject phrase were counterbalanced across participants and trials. Each participant saw only 1 alternative from each sentence set but 10 exemplars of each combination produced by crossing verb type and NP length.
Table 2. Example Sets of Sentence Elements Contrasting Verb Type and NP Length

<table>
<thead>
<tr>
<th>Verb</th>
<th>Length</th>
<th>Subject phrase</th>
<th>NP</th>
<th>PP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transitive</td>
<td>short</td>
<td>Janet transferred</td>
<td>some plans</td>
<td>to Lee</td>
</tr>
<tr>
<td>Complement</td>
<td>short</td>
<td>Janet revealed</td>
<td>some plans for the brand new defense plant</td>
<td>to Lee</td>
</tr>
<tr>
<td>Transitive</td>
<td>long</td>
<td>Janet transferred</td>
<td>some more specific plans</td>
<td>to Lee</td>
</tr>
<tr>
<td>Complement</td>
<td>long</td>
<td>Janet revealed</td>
<td>some more specific plans for the brand new defense plant</td>
<td>to Lee</td>
</tr>
<tr>
<td>Transitive</td>
<td>short</td>
<td>Jake released</td>
<td>the figures</td>
<td>the figures</td>
</tr>
<tr>
<td>Complement</td>
<td>short</td>
<td>Jake explained</td>
<td>the figures</td>
<td>on Friday</td>
</tr>
<tr>
<td>Transitive</td>
<td>long</td>
<td>Jake released</td>
<td>all the figures about the new import and export taxes</td>
<td>on Friday</td>
</tr>
<tr>
<td>Complement</td>
<td>long</td>
<td>Jake explained</td>
<td>all the figures about the new import and export taxes</td>
<td>on Friday</td>
</tr>
</tbody>
</table>

Note: NP = noun phrase; PP = prepositional phrase.

Rotor Training

Participants were initially trained on pursuit rotor tracking following the protocol in Kemper and colleagues (in press-b). The pursuit rotor features a target that rotates along a circular track. Participants use a trackball mouse to track the target, attempting to keep a pointer centered on the moving target. Rotor speed can be adjusted from approximately .2 to 2 revolutions per min; the program samples the location of the pointer approximately every 16 ms and determines its distance (in pixels) from the center of the target. A moving average of the pointer status (on/off target) is taken over three successive 100-ms intervals, and percentage time off target (TOT) is determined. In addition, tracking error (TE) or the distance in pixels from the center of the target to the pointer is used as a second measure of tracking performance; it is also averaged over three successive 100-ms intervals.

Participants were trained on the pursuit rotor task to an asymptotic performance level. Initial tracking speeds for young and older adults were 1.2 and 0.45 rpm, respectively. Participants practiced tracking for 30 s and received feedback on their performance. A “two up/one down staircase” training procedure was used to gradually increase tracking speed on successive 30-s trials: If average TOT was 20% or less for a trial, the speed was increased by 10% for the next trial; if TOT was greater than 20%, the speed was decreased by 5%. The staircase procedure converged on an asymptotic rotor speed when the rotor speed oscillated around the same value, moving “up” and “down” past this value three times. Young adults (M = 16.5, SD = 4.4) required fewer trials to reach an asymptotic rotor speed than older adults (M = 18.4, SD = 4.9), F(1, 78) = 4.86, p < .05. Asymptotic tracking speed was greater for young adults (M = 1.7 rpm, SD = 0.3) than for older adults (M = 1.0 rpm, SD = 0.3), F(1, 78) = 97.99, p < .01. However, asymptotic TOT (M = 18.4%, SD = 3.8) and TE (M = 7.6 pixels, SD = 0.9) were comparable for young and older adults, both p > .50.

Experimental Procedure

Following rotor training, two experimental tasks were administered; order was counterbalanced across participants. Both were administered using paradigm (Tagliaferri, 2005). In addition to the controlled sentence production task, participants were also tested on a related experiment requiring them to read paragraphs aloud while tracking (Kemper et al., in press-a).

The controlled sentence production task involved 6 practice trials and 40 experimental trials. A version of the pursuit rotor was embedded within paradigm, and tracking speed was set to the asymptotic speed achieved by the participant during training. The design counterbalanced verb type (transitive vs. complement), NP length (short vs. long), and position of the NP (above or below the subject phrase) across trials. None of outcome measures varied with NP position; consequently, NP position is not discussed further. Prompts and sentence elements were presented centered within the circular rotor track and did not obscure the track or the pointer. Each trial involved four phases:

1. Baseline tracking: The pursuit rotor was initiated after a 3-s delay, and participants tracked it continuously for 20 s while a central fixation cross was presented.
2. Planning: The triad of sentence elements was presented for 15 s, participants were instructed to “make up a sentence beginning with the underlined phrase and using all the words and phrases,” and they were reminded “keep tracking.”
3. Delay: The sentence elements were removed, and “WAIT” was displayed centered within the track while the participants continue to track during a 2-s delay.
4. Production: “Speak” was displayed to prompt the participants to respond while continuing to track the rotor for an additional 15 s. The rotor reset at the end of each trial, repositioning the target to the “6 o’clock” starting position.

The tracking record was segmented into four phases: the baseline interval, the sentence planning interval, the delay interval, and the response production interval. Participants responded orally; their responses were recorded, and the audio (WAV) files were synchronized with their tracking record.

Results

The behavioral outcomes are presented first, followed by the analysis of the dual-task costs to pursuit rotor performance.
during the baseline, planning, and production phases. The delay phase was too short for the analysis of meaningful effects and is not discussed further. Data from 10 trials, all for older adults (<1% of all trials), were lost due to mechanical failure or because the participant stopped tracking during the trial. These trials were excluded from further analysis.

Behavioral Outcomes

Each response was coded off-line as valid (a fluent grammatically correct response using the correct NPs, verbs, and PP) or invalid (disfluent responses, responses that omitted one or more elements, or responses that substituted a different lexical item). Valid responses were coded as to whether the NP was shifted to the end of the sentence or not. The behavioral outcomes are summarized in Table 3. Participants produced valid responses on 84% of the trials (93% for young adults, and 75% for older adults). NPs were shifted to the sentence final position on 12.9% of the trials (7.4% for young adults and 19.8% for older adults).

Mixed-effects logistic regression (Jaeger, 2008) was used to analyze these data; odds ratios (ORs) for valid versus invalid responses and for shifted versus nonshifted NPs were analyzed as a function of age group, NP length, verb type, and their interactions. In addition, the models examined whether the person-level predictors for vocabulary, working memory, processing speed, and inhibition moderated effects of the NP length or verb type on the likelihood of valid trials or shifted NPs. Unless reported below, all other effects and interactions were not significant.

Valid/invalid.—Trials with short NPs were two times more likely to be valid than those with long NPs, as indicated by a significant effect of NP length, OR = 2.07, 95% confidence interval (CI) = 1.15–3.75. This effect of NP length was 12 times greater for older adults than for young adults, as indicated by the significant age group by NP length interaction, OR = 12.41, 95% CI = 4.56–33.75. Participants with larger vocabularies were somewhat more likely to produce valid responses than those with more limited vocabularies, resulting in a significant person-level effect of vocabulary for valid responses, OR = 1.71, 95% CI = 1.20–2.70.

NP shift.—The likelihood of NP shift was examined only for valid trials. NPs following complement verbs were two times more likely to be shifted than those following transitive verbs, as indicated by the significant effect of Verb Type, OR = 2.29, 95% CI = 1.18–4.56. There was a tendency for long NPs to be shifted more often than short NPs, OR = 1.86, 95% CI = 0.94–3.70. Older adults were 10 times more likely to shift long NPs than young adults, as indicated by the significant Age Group × NP Length interaction, OR = 10.60, 95% CI = 3.14–35.85.

Dual-Task Costs

Only valid trials were used in the remaining analyses. Tracking performance was sampled at approximately 16-ms intervals, resulting in more than 3,100 samples per participant per trial. Unlike the training program, a running average was not computed; for this analysis, samples were averaged over the entire trial event (i.e., approximately 940 samples were averaged for each 15-s planning or production event). TE was computed as the average distance in pixels between the center of the target and the tip of the pointer; TOT was computed as the proportion of samples where the distance was greater than the target diameter of 16 pixels.

Figures 1 and 2 provide an overview of the dual-task costs to pursuit rotor performance over the time course of the experimental trials. Local polynomial smoothing (Fan & Gijbels, 1996) was used to generate the figures using Stata (StataCorp, 2009). This technique fits the tracking performance measures to a polynomial function using locally weighted least squares regression, and the smooth polynomial function is then sampled at regular intervals (approximately every 240 ms) for plotting. As the figure indicates, the tracking performance of young and older adults was similar during the baseline phase. However, older adults’ tracking diverged from young adults’ during the planning phase and remained divergent during the production phases. Older adults were more likely to be off target.

Table 3. Summary of Behavioral Outcomes for Young and Older Adults

<table>
<thead>
<tr>
<th>Behavioral outcome</th>
<th>Transitive verbs</th>
<th></th>
<th>Complement verbs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Short NPs</td>
<td>Long NPs</td>
<td>Short NPs</td>
<td>Long NPs</td>
</tr>
<tr>
<td>Valid responses (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young adults</td>
<td>95.5 20.8</td>
<td>91.3 28.3</td>
<td>94.3 23.3</td>
<td>90.3 29.7</td>
</tr>
<tr>
<td>Older adults</td>
<td>96.7 17.8</td>
<td>56.7 49.6</td>
<td>94.2 23.4</td>
<td>52.6 50.0</td>
</tr>
<tr>
<td>Shifted NPs (%)</td>
<td>3.9 19.4</td>
<td>6.8 25.3</td>
<td>8.2 27.5</td>
<td>10.8 31.1</td>
</tr>
<tr>
<td>Young adults</td>
<td>7.3 26.0</td>
<td>35.1 47.8</td>
<td>9.1 28.8</td>
<td>45.7 49.9</td>
</tr>
<tr>
<td>Older adults</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: NP = noun phrase.
and to be further off target than young adults during sentence planning and production. Furthermore, an effect of NP length on tracking performance was apparent during the planning and production phases, especially for older adults.

Mixed effects modeling (Blozis & Traxler, 2007; Quene & van den Bergh, 2004) with restricted maximum likelihood estimation was used to analyze these results. This approach easily handles missing data, both categorical and continuous measures, and repeated observations. Separate models of the baseline tracking, planning, and production phases examined the effect of age group, the experimental manipulations of NP length and verb type, and their interactions. The effect of whether the NP was shifted or not was also included in the planning and production models. In addition, vocabulary, working memory, processing speed, and inhibition were examined as potential moderators of the
effects of age group, NP length, verb type, and their interactions in the planning and production models.

Separate models were estimated for the two dependent measures, TE (pixels) and tracking time off task (percent). Positive estimates (est.) indicate an increase in TE or TOT, hence worsening performance, whereas negative estimates indicate a decrease in TE or TOT, hence an improvement in tracking performance. Unless reported subsequently, all other effects and interactions were not significant.

Baseline tracking.—Tracking performance during the baseline phase was similar for all participants and for all types of experimental trials, resulting in nonsignificant estimates for age group, NP length, verb type, and their interaction. Over all, TE averaged 4.4 pixels, SE = 0.16, and TOT averaged 13.21%, SE = 0.31.

Planning.—Planning responses with long NPs was more difficult, increasing TOT, est. = 3.86%, SE = 0.31, p < .01, and TE, est. = 0.81 pixels, SE = 0.39, p < .05. The effect of NP length on TOT was somewhat attenuated for participants with greater processing speed, est. = −4.25%, SE = 1.10, p < .01, participants with larger vocabularies, est. = −2.74%, SE = 0.89, p < .01, and participants with better working memory, est. = −9.97%, SE = 3.01, p < .01. Overall, planning a response with a shifted NP was also more challenging than planning one with an NP immediately following the verb, as indicated by the significant effect of NP shift for TOT, est. = 5.73%, SE = 2.51, p < .05.

Older adults were more likely to be off target, est. = 11.30%, SE = 2.48, p < .01, and to be further off target than young adults, est. = 4.19 pixels, SE = 1.48, p < .01. Planning responses with long NPs was more challenging for older adults than young adults, as indicated by significant Age Group × NP Length interactions for TE, est. = 4.89 pixels, SE = 0.64, p < .01, and TOT, est. = 7.69%, SE = 1.16, p < .01. This effect of long NPs on older adults’ TE was attenuated slightly for complement verbs, as indicated by the significant Age Group × NP Length × Verb Type interaction, est. = −1.90 pixels, SE = 0.93, p < .05. Older adults who processed information more rapidly had a slight advantage, as indicated by the significant interaction between age group and processing speed for TOT, est. = −7.89%, SE = 3.58, p < .05.

Production.—Overall, tracking performance while participants were producing their response was not affected by verb type, NP length, or NP shift as indicated by nonsignificant estimates for these linguistic manipulations. Producing a response was more costly for older adults than for young adults, affecting TE, est. = 4.94 pixels, SE = 0.97, p < .01, and TOT, est. = 10.75%, SE = 1.87, p < .01. Producing responses with long NPs was more challenging for older adults than for young adults, as indicated by the significant Age Group × NP Length interaction for TOT, est. = 2.35%, SE = 0.86, p < .01. The age group disadvantage was attenuated for older adults who processed information more rapidly, as indicated by the significant interaction between age group and processing speed for TE, est. = −2.94 pixels, SE = 1.30, p < .05, and TOT, est. = −5.76%, SE = 2.63, p < .01.

Discussion

As in the Stallings and colleagues (1998) experiment, participants in this experiment were more likely to produce sentences with shifted NPs when given complement verbs than when given transitive verbs, and there was a tendency for participants to shift long NPs more often than short NPs. Moreover, by examining pursuit rotor tracking, this experiment showed that these linguistic manipulations not only affected behavioral outcomes but also affected sentence planning and production costs. Planning a sentence with a long NP or one with a shifted NP was costly to tracking performance.

Individual differences in cognitive ability do, however, affect sentence planning and production. A larger vocabulary provided some advantage on this task, perhaps due to increased familiarity with the lexical items or with alternative syntactic structures. Individuals who process information more rapidly, those with better working memories, and those with larger vocabularies had reduced costs for long NPs during the sentence planning phase. Acheson and MacDonald (2009) suggest that cognitive ability may affect how individuals deal with the demands of planning the serial order of elements within a sentence. Hence, greater processing speed may enable individuals to more rapidly construct a sentence around a long NP with many modifiers, a larger vocabulary may increase familiarity with lengthy phrases, and greater working memory may enable individuals to manipulate the entire phrase as they plan its position within the sentence. However, once the participant had planned how to incorporate the long NP into their sentence, actually producing the sentence was not affected by individual differences in cognitive ability. Inhibitory control did not seem to play a critical role in either sentence planning or production, although other executive functions such as switching or updating (Miyake, Friedman, Emerson, Witzki, & Howerton, 2000) may be more critical.

Aging also affects sentence planning and production. When confronted with the challenge of incorporating a long NP into a sentence, older adults were less likely to produce a valid sentence than young adults and when they did so, they were more likely to shift the long NP to the end of the sentence. Furthermore, planning and producing sentences with long NPs were more difficult for older adults than for young adults, as indicated by greater costs to pursuit rotor tracking. This approach avoids many of the measurement problems associated with using reaction times to index age group differences (Faust, Balota, Spieler, & Ferraro, 1999). It is important to this argument that young and older adults
were initially trained to their individual asymptotic performance as indicated by their comparable performance during the baseline tracking phase. Thus, although older adults could match the pursuit rotor performance of young adults in terms of TOT and TE during the baseline, pursuit rotor tracking during the planning and production phases was more difficult for older adults than for young adults, as indicated by increased TE and increased TOT. Planning and producing sentences with long NPs were especially challenging for older adults.

Prior controlled production studies had suggested that age-related changes to working memory impose a functional ceiling on the construction of long complex sentences. The findings that older adults had difficulty producing valid responses with long NPs and were more likely to shift long NPs are consistent with this hypothesis as is the finding that planning and producing sentences with long NPs were more costly for older adults than for young adults. However, this experiment suggests that the “height” of this functional ceiling is not fixed by working memory because working memory did not moderate the effects of NP length or NP shifting on sentence planning and production. Furthermore, older adults who processed information more rapidly had an overall advantage regardless of the linguistic manipulations of NP length and verb type. It may be that older adults attempted alternate or switch between the two tasks, and greater processing speed may provide an advantage by enabling older individuals to rapidly alternate tracking with sentence planning and production. Hence, although tracking during the dual task conditions appeared to be continuous, older adults may have adopted a strategy of allowing their tracking to lag somewhat during sentence planning or production operations then rapidly catching up with the moving target, alternating between the two tasks. Long NPs may disrupt this pattern of task alternation by increasing the lag, forcing older adults to devote more time to sentence planning and less time to tracking, thereby affecting time on target and TE. Systematic investigations of the role of executive functions such as switching and updating and of age differences in dual task performance with a greater range of lexical items and syntactic structures varying in frequency, familiarity, and processing demands can help to resolve these possibilities.

FUNDING

This work was supported in part by grants from the National Institutes of Health (NIH) to the University of Kansas through the Mental Retardation and Developmental Disabilities Research Center (grant number P30 HD-002528) and the Center for Biobehavioral Neurosciences in Communication Disorders (grant number P30 DC-005803) as well as by a grant from the National Institute on Aging (grant number RO1 AG-025906) to Susan Kemper. Its contents are solely the responsibility of the author and do not necessarily represent the official views of the NIH.

ACKNOWLEDGMENTS

We thank Whitney McKedy for her assistance with data collection and analysis. Doug Kieweg of the Biobehavioral Neurosciences in Communication Disorders developed the digital pursuit rotor; Bruno Tagliaferri developed paradigm; they collaborated on the integration of the two programs. A suite of digital pursuit rotor applications is available upon request; paradigm is available from Perception Research Systems. Preliminary results were presented at the 2010 Cognitive Aging Conference in Atlanta.

CORRESPONDENCE

Correspondence should be addressed to Susan Kemper, PhD, Gerontology Center, 3000 DHDHC, 1000 Sunnyside Ave., University of Kansas, Lawrence, KS 66045. E-mail: skemper@ku.edu.

REFERENCES


