Aging and word finding:
Reverse vocabulary and Cloze tests

EUGENE A. LOVELACE and VICKY E. COON
Alfred University, Alfred, New York

Two tasks were used to evaluate age differences in "word-finding difficulty,"—lexical access and retrieval—for 31 young adults (college students) and 24 healthy, community-dwelling older adults (aged 58 to 86). Comparison of performances on a traditional (forward) and a reverse vocabulary test for the same set of 43 nouns indicated that the aged could define the words as well as or better than the young, but that they had greater difficulty thinking of the word when given the definition. A Cloze task, in which 32 nouns had been deleted from each of four prose passages, required that participants try to guess the deleted words. On this task, the performances of young and old adults were very similar. The only suggestion of greater word-finding difficulty for the aged was that they more often failed to provide *any* response, although in absolute terms these omission errors were quite rare for both age groups.

A common observation of the aged regarding subjective changes in memory function is that they have trouble bringing some things they know to mind when they want to—most notably, people's names and words they wish to use in conversation (Cohen & Faulkner, 1986; Maylor, in press; Sunderland, Watts, Baddeley, & Harris, 1986). This is true even in "advantaged" samples (e.g., those in Hellebrandt, 1980; Lovelace & Twohig, 1990). This problem with semantic memory seems at odds with the common empirical finding that vocabulary remains very much intact for healthy older adults (Botwinick, 1984). However, the usual vocabulary test is a test of the ability to get from words to their meanings (speech comprehension), whereas the complaint is made with regard to a difficulty in moving from meanings to specific words (speech production).

We hypothesized that everyone experiences some asymmetry with respect to the relative difficulty of moving from meanings to words and words to meanings. We rarely hear someone use a word that we know without being able to think of its meaning, yet most of us will occasionally want to use a word to convey a particular meaning and find ourselves unable to think of the word. Given the perception of many aged that this word-finding difficulty increases in frequency as they grow older (Lovelace & Twohig, 1990), we hypothesized that this asymmetry would be greater for old than for young adults.

Previous work of Bowles and Poon (1985) and Burke, Worthley, and Martin (1988) provided some support for these hypotheses. Burke et al. had young and old adults keep diaries to record occurrences of tip-of-the-tongue states, clear cases of word-finding difficulty. Older participants recorded significantly more of these events. Bowles and Poon found that the elderly were significantly less likely than young adults to be able to recall the correct word when read the word's definition. They took this to "reflect a specifically directional breakdown in the connection from the semantic network to the lexical network" (p. 76).

In the present study, we provided a more direct test by giving forward and backward vocabulary tests on the *same items* to the *same individuals*. In addition, the Cloze procedure (Taylor, 1953, 1957) was used to see if there would be an age-related decline in ability to generate the deleted words specified by the running context of a prose passage.

METHOD

Participants

The 55 participants, run individually, were 31 young adults ($M = 20$ years) and 24 older adults ($M = 70$ years; range, 58-87). The young adults were undergraduate students, and the older adults were community-dwelling elderly. On a 4-point self-rating scale for health (poor, fair, good, excellent), the distribution for this sample of older adults was 2 fair, 16 good, and 6 excellent.

Procedure

Two word-retrieval tasks were administered in each of two sessions, each session lasting less than 1 h. The two sessions were separated by 10 days on the average (range, 7-16 days). One task involved vocabulary tests, the second a Cloze task. In the first session, a forward vocabulary test was given, followed by a brief practice passage for the Cloze procedure. The Cloze task was then carried out for two passages. In the second session, two more Cloze passages were followed by a reverse vocabulary test.

On the forward vocabulary test, 43 words were tested in the traditional manner; each word was spoken aloud, and participants were instructed to give the definition of the word. For the reverse vocabulary

test, the definition of each of the same 43 words was provided, and participants attempted to provide the word. Each of these 43 words was a noun that did not have an obvious synonym. They ranged from 4 to 13 letters in length, and from 0 to 17.78 in U frequency index (occurrences per million words), in *The American Heritage Word Frequency Book* (mean $U = 1.51$).

For the Cloze tests, subjects were informed that they would be asked to read four passages, drawn from magazines intended for the general public, in which key words had been deleted. The passages, drawn from stories in *The Reader's Digest*, were between two and three double-spaced typewritten pages in length. Since the word-finding problems for the aged appear more related to naming *things*, 32 common nouns had been deleted from each passage. The objective of the Cloze task was explained as follows: "We want you to read through ... magazine passages and try to guess what word the author used for each of the blanks created when we deleted some words." They were instructed not to read beyond the sentence containing the blank before filling it in, and told that they could not return to change a word once it had been entered.

Finally, for the aged sample, some personal information was gathered via an interview/questionnaire. The questions concerned their date of birth, their amount of schooling, their general health evaluation, and types and amount of reading that they normally did.

RESULTS AND DISCUSSION

Vocabulary Tasks

On the traditional (forward) vocabulary test, the average numbers of correct definitions were 40.7 and 41.3 for young and old adults, respectively, out of a possible 43. On the reverse vocabulary test, the mean numbers of words correctly provided were 39.8 and 38.9, respectively. Although overall performance was quite high, and the absolute differences in performance levels were small, there was a significant main effect of test direction, with the reverse vocabulary test producing lower performance than the forward [$F(1,52) = 22.0, p < .001$]. The interaction of age with forward versus backward test was also significant [$F(1,52) = 4.64, p < .05$]. This interaction reflects a reversal of the direction of the effect of age between the forward and reverse vocabulary tests; the old did better than the young at defining words, whereas the young were superior at getting from meanings to words.

For performance on the forward and backward vocabulary tests considered jointly, the responses to the 43 items for each individual can be tallied in the 2 × 2 array of correct versus incorrect for forward versus backward tests. The numbers of items falling in each of these four cells were compared. For both age groups, the majority of the items fell in the cell reflecting correct responses on both tests ($Ms = 38.5$ and 38.3 for young and old, respectively). Three individuals in each age group made no errors on either test; they were excluded from further analyses of the vocabulary tests.

Analyses were next performed on items that were correct on one test but not on the other. In both age groups, more individuals showed the pattern of being correct more often on the forward than on the backward test; this is consistent with the hypothesized asymmetry of access. Furthermore, as predicted, this effect was greater for old than for young adults, 17 versus 9 individuals from the young sample and 14 versus 1 for the old. (The remain-

ing 2 young and 5 old participants were correct for such items equally often on forward and backward tests.) This asymmetry is significant by simple sign test for the aged (14 vs. 1), but not for the young (17 vs. 9). Differential asymmetry for the two age groups was confirmed by a chi-square analysis of the frequencies in the 2 × 3 array (young vs. old × forward > backward, backward > forward, forward = backward), which showed a significant relationship of age to the relative distribution of more correct responses on forward or backward tests ($x^2 = 6.83$, $p < .05$).

Although the overall performance levels for the two groups were higher than one might prefer for maximum sensitivity of the tasks, there is clear evidence that the older adults had greater difficulty moving from the definition to the word than did young adults. This finding is consistent with other evidence, both from subjective reports of older individuals and from the laboratory research of Burke and her associates (e.g., Burke & Laver, in press) and Bowles and associates (Bowles, 1989; Bowles, Obler, & Poon, 1989), indicating a greater word-finding difficulty for the names of things for normal, healthy older adults than for young adults.

Since direction of the test was within-subjects and the reverse vocabulary test was given second, the interaction might reflect differential carryover from the forward to reverse tests. Although this cannot be ruled out, it does seem highly unlikely, given the combination of very low frequency words, lengthy delay between sessions, and the interpolation of four Cloze tasks.

Cloze Task

Each response on the Cloze task was coded as being one of six types: (1) the actual word that had been deleted, (2) a synonym of that word, (3) one that fit fairly well and preserved the general gist, (4) one that was syntactically correct but semantically anomalous, (5) syntactically incorrect, or (6) an omission. The criteria for the "actual word" category included cases in which the responses differed from the deleted word only in number (singular vs. plural).

Three research assistants scored the responses to the Cloze task. In each case, the first author also scored the data of at least 4 subjects on all four passages as a reliability check on the scoring. The percentage agreement on assignment of responses to categories ranged from 85% to 89%. Most of the disagreements were between Categories 2 and 3, reflecting some difficulty in deciding just how similar the words must be to be called a synonym and not simply a word that preserved the gist.

There was very little difference in the distributions of these six types of response as a function of age. The mean percentages of responses where the actual deleted word was given were 40% and 42% for young and old adults, respectively. Mean proportions for synonyms were 17% for both groups, and those preserving gist were 25% for young and 21% for old. These three categories,

where meaning was at least roughly preserved, account for 82% and 79% of responses of young and old adults, respectively.

Responses that were semantically anomalous were about equally likely for the two age groups, accounting for a mean proportion of 17% and 18% for young and old. Very few responses fell in the remaining two categories of syntactically incorrect words or omissions. The only cell to contain more than a fraction of 1% of the responses was the omission category for the aged participants, which contained a little more than 2% of their responses.

More omissions by the aged, of course, might represent the expected pattern of greater failure to retrieve an intended word. Since these are group mean proportions, this could, of course, result from a single older adult who frequently failed to provide a response. To assess this, the data for individuals were examined. For the young, 7 of the 31 individuals showed one or more omission responses, whereas 12 of the 24 old individuals showed omissions. This difference in proportion of individuals committing omissions for the two age groups is significant as a one-tailed test ($\chi^2 = 3.37$, $p < .05$).

SUMMARY

The vocabulary tests did indicate an asymmetry in ease of moving between the word and the meaning. Participants had a greater difficulty getting from the idea to the word than vice versa, an effect that showed a significantly greater asymmetry for old than for young adults, although the effect was not large in absolute terms for either group. This increased difficulty for the aged in getting from semantic to phonological features parallels the findings of Maylor (in press) and Burke and Laver (in press).

For the Cloze task, where the individuals sought a word on the basis of the context of the prior text and the current sentence, age differences were minimal. This Cloze procedure is similar to that in sentence completion studies by Cohen and Faulkner (1983) and Nebes, Boller, and Holland (1986), and it leads to the same general conclusion: that the aged are able to use the running context of a sentence to generate missing words with about the same facility as young adults.

As one might predict from the results of the vocabulary tests, and the general notion of age-related differences in word-finding ability, the aged were somewhat more likely to fail to provide any word. This was, however, a very infrequent event. In the 128 blanks to be filled, only half the aged subjects ever failed to produce some word (compared to 22% of the young).

This higher level of omission responses by the aged adults may, of course, reflect criterion differences (greater cautiousness in responding—see Botwinick, 1984) rather than, or in addition to, age differences in word finding. There is considerable doubt, however, about the generality of any age-related differences in cautiousness (Kausler, 1982; Salthouse, 1982).

Anyone who would interpret the small size of the effects in the present study must keep in mind that this was an attempt to induce an event for which the natural absolute frequency of occurrence is very low (Burke et al., 1988). Also, to the extent that the self-paced nature of the present tasks did not duplicate the normal time constraints placed on speakers operating in real time, the present tasks may be viewed as biased against finding age effects of a magnitude that would occur in actual daily speech. Bowles and Poon (1985) have shown that, even for words that both young and old successfully retrieved from the definition, latencies were significantly longer for aged than for young adults. Thus it is clear that the time constraints of conversation are likely to enhance differential age effects in word retrieval, and quite possibly greater age differences in performance on the Cloze task would have occurred had the task been paced.

REFERENCES

BOTWINICK, J. (1984). *Aging and behavior* (3rd ed.). New York: Springer-Verlag.

BOWLES, N. L. (1989). Age and semantic inhibition in word retrieval. *The Journals of Gerontology: Psychological Sciences*, **44**, P88-90.

BOWLES, N. L., OBLER, L. K., & POON, L. W. (1989). Aging and word retrieval: Naturalistic, clinical, and laboratory data. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 244-264). New York: Cambridge University Press.

BOWLES, N. L., & POON, L. W. (1985). Aging and retrieval of words in semantic memory. *Journal of Gerontology*, **40**, 71-77.

BURKE, D., & LAVER, G. (in press). Aging and word retrieval: Selective age deficits in language. In E. A. Lovelace (Ed.), *Aging and cognition: Mental processes, self-awareness, and interventions*. Amsterdam: Elsevier.

BURKE, D., WORTHLEY, J., & MARTIN, J. (1988). I'll never forget what's-her-name: Aging and tip of the tongue experience in everyday life. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 2, pp. 113-118). Chichester, England: Wiley.

COHEN, G., & FAULKNER, D. (1983). Word recognition: Age differences in contextual facilitation effects. *British Journal of Psychology*, **74**, 239-251.

COHEN, G., & FAULKNER, D. (1986). Memory for proper names: Age differences in retrieval. *British Journal of Developmental Psychology*, **4**, 187-197.

HELLEBRANDT, F. A. (1980). Aging among the advantaged: A new look at the stereotype of the elderly. *The Gerontologist*, **20**, 404-417.

KAUSLER, D. H. (1982). *Experimental psychology and human aging*. New York: Wiley.

LOVELACE, E. A., & TWOHIG, P. T. (1990). Healthy older adults' perceptions of their memory functioning and use of mnemonics. *Bulletin of the Psychonomic Society*, **28**, 115-118.

MAYLOR, E. A. (in press). Recognizing and naming faces: Aging, memory retrieval and the tip of the tongue state. *Journal of Gerontology: Psychological Sciences*.

NEBES, R., BOLLER, F., & HOLLAND, A. (1986). Use of semantic context by patients with Alzheimer's disease. *Psychology & Aging*, **1**, 261-269.

SALTHOUSE, T. A. (1982). *Adult cognition: An experimental psychology of aging*. New York: Springer-Verlag.

SUNDERLAND, A., WATTS, K., BADDELEY, A. D., & HARRIS, J. E. (1986). Subjective memory assessment and test performance in the elderly. *Journal of Gerontology*, **41**, 376-384.

TAYLOR, W. L. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism Quarterly*, **30**, 415-433.

TAYLOR, W. L. (1957). "Cloze" readability scores as indices of individual differences in comprehension and aptitude. *Journal of Applied Psychology*, **41**, 19-26.