The Dynamics of Memory Retrieval in Older Adulthood

ARTHUR WINGFIELD and MICHAEL J. KAHANA
Brandeis University, Waltham, MA

Abstract One of the most robust findings in cognitive aging is that of a significant decline in self-initiated recall from episodic memory. In laboratory studies this deficit can be seen in significant age differences in word-list free recall. In this article, we focus on free recall of categorized word lists where one observes "response bursting" in the form of a rapid output of within-category items with longer delays between categories. Age differences appear primarily in between-category latencies, results that are consistent with a relative sparing of semantic memory combined with an age-deficit in episodic retrieval. When adjusted for differences in overall mnemonic ability, we demonstrate that the relationship between organization and learning remains invariant with normal aging. We argue that the locus of the age deficit in free recall lies at the level of temporal coding of items and the use of temporal associations to guide recall.

One of the most common complaints among older adults is that of increasing difficulty with memory for recently experienced events, to include recently learned names, list of things to do, or narratives relating to recent activities of family and friends (Cutler & Grams, 1988; Dobbs & Rule, 1987; Gilewski & Zelinski, 1986). These difficulties are highlighted in laboratory tasks that involve explicit episodic memory. These include tests of free recall, paired-associate learning, and recognition memory (Craik & Jennings, 1992; Hartley, 1989, 1993; Moscovitch & Winocur, 1992; Light, 1991). However, not all of these are affected equally. For example, it is well known that age differences are larger in tests of free recall than for tests involving recognition or cued recall (Burke & Light, 1981; Craik, 1977; Craik & McDowd, 1987; Perry & Wingfield, 1994). These data support the general view that age-related memory changes are particularly pronounced in explicit memory (Mitchell, 1989; Schacter, Chiu, & Ochsner, 1993), and tasks that involve controlled processing (Jacoby, 1991).

Unlike recognition memory and cued recall, free recall requires the individual to initiate the formation of the retrieval cues that may facilitate access to the desired information. It may be for this reason that deliberate recollection (i.e., free recall) is among the most age-sensitive of cognitive tasks (Craik & Jennings, 1992). An additional factor in self-initiated recall is that success requires the effective inhibition of related, but nontarget, memories (Hasher & Zacks, 1988; Zacks, Hasher, & Li, 2000). Indeed, failure at achieving this inhibition may result in the production of false memories (Roediger & McDermott, 1955), a problem that is exaggerated among older adults (Norman & Schacter, 1997; Tun, Wingfield, Rosen, & Blanchard, 1998).

Adult Aging and Retrieval Mechanisms in Free Recall

These and other findings in cognitive aging cannot be fully understood without a detailed theoretical understanding of the retrieval mechanisms in free recall in young and older adults. After some years of neglect, the general cognitive literature has shown a resurgence of interest in the retrieval dynamics of free recall from episodic memory (Howard & Kahana, 1999; Kahana, 1996; Kahana & Loftus, 1999; Rohrer, Wixted, Salmon, & Butters, 1995; Wixted & Rohrer, 1996).

Although classical ideas of association and interference were largely replaced by more cognitive constructs in the 1960s and 1970s, more recent neural and computational models of memory have brought the earlier focus on simple associations to the foreground of contemporary theory. This new approach, however, has been able to overcome the limitations of classical association theory by providing much more powerful models for describing the mechanisms of associative storage and retrieval.

We believe that a fruitful approach to understanding the problem of the age-related deficits in human learning and memory is to take models of normative memory function that have been successful in accounting for a wide range of data in young adults, and then determining how "lesioning" such a model would produce behaviour that is isomorphic with the known pattern of memory changes in older adulthood.

Experimental studies of aging memory that have looked at overall levels of performance for young versus older adults have verified an effect closely akin to the "complexity hypothesis" championed many years ago by Cerella and colleagues (e.g., Cerella, Poon, & Williams, 1980). When tested across a variety of tasks, the general rule of aging is that tasks that the young find difficult, or require a longer time to perform, are especially difficult or take especially longer for older adults (Cerella, 1994; Stine & Wingfield, 1988, 1990; Wingfield, Tun, & Rosen, 1995). Interestingly, the argument has been made that the very ubiquity of this finding may have tended to obscure differences in young and older adults in the microstructure of the tasks being performed (Fisk & Fisher, 1994; Myerson, Wagstaff, & Hale, 1994). The goal in cognitive aging research, we would thus argue, should be to get below the surface of the general complexity hypothesis, and to attempt to elucidate these potential microstructure differences.

Recall from Semantic Memory: Output Rate and "Response Bursting"

From the time of Cattell (1886), it has been recognized that the time it takes to retrieve information from memory can reveal information about accessibility differences between items even when recall is accurate. One example of the usefulness of measuring response times can be seen in the classic finding that the temporal output pattern observed in the free recall of verbal items is frequently characterized by a phenomenon known as "response bursting" (Bousfield, Cohen, & Whitmarsh, 1958; Bousfield & Puff, 1964; Bousfield & Sedgewick, 1944; Bousfield, Sedgewick, & Cohen, 1954). In the earliest of these studies, Bousfield and Sedgewick (1944) asked young adults to name category exemplars from large semantic categories such as animals, U.S. cities or items of furniture. Using a variety of different categories they found that the rate of output of category exemplars typically begins rapidly and then progressively decreases as more and more of the exemplars are given and the category becomes exhausted. That is, the number of responses given over time takes the form of an exponential rise to an asymptote (Wixted & Rohrer, 1996).

Within this general trend, Bousfield and Sedgewick (1944) also observed the appearance of associative clustering: the rapid production of highly associated exemplars (e.g., within the "animal" category, animals typically found on a farm, animals typically found in a zoo, etc.) preceded by a relatively long response time. That is, as the search through memory proceeds, recalling an exemplar prototypical of a particular subcategory appears to trigger the rapid production of the remaining members of that subcategory. This, in turn, is often followed by a relatively long latency until another group of responses is produced. In Figure 1, we illustrate these effects with data from our own laboratory. University undergraduates were asked to name aloud as many animals as they could within a 20-minute period. (Participants were told that for this purpose fish and birds were to be considered part of the animal category, but that insects and microorganisms were to be excluded.) The insert in the upper left-hand corner of Figure 1 shows a typical temporal output pattern from a single participant. Time in minutes is plotted along the abscissa and the cumulative number of responses is plotted on the ordinate. We can see in this insert one of the features observed by Bousfield: a fairly rapid output rate as the participant begins to give his or her responses, followed by a slowing of output over time as the participant requires more and more time to retrieve and produce new category members that have not already been given.

Within this generally declining rate of output, however, one also sees several examples of the associative clustering with temporal response bursting described by Bousfield. This can be seen in the larger curve in Figure 1 that focuses on the area indicated by the box in the insert. This curve shows a sample of the participants' responses made from 6 minutes to just over 12 minutes into the production period. Each vertical rise represents the production of one additional word. Listed below the curve are the actual responses produced by the participant, with the vertical order from top to bottom representing the order of responses over time. In this enlarged sample of the output curve, one sees numerous examples of "response bursting" of subcategories (the rapid output of responses from within a subcategory with longer delays between subcategories). An example of this pattern is seen with the production of *Cougar* being followed rapidly by three other animals one might see on safari (*Elephant*, *Hippopotamus* and *Rhinoceros*), followed in turn by a long pause until *Poodle*, an item from another category, was produced.

Asking individuals to generate exemplars from specified semantic categories, or to produce words beginning with particular sounds, is frequently used in neuropsychological testing as a measure of "verbal fluency." Because of pragmatic constraints in clinical testing, typically only one or two minutes is allowed for responding. Even within this constraint, however, the number of responses generated within this specified

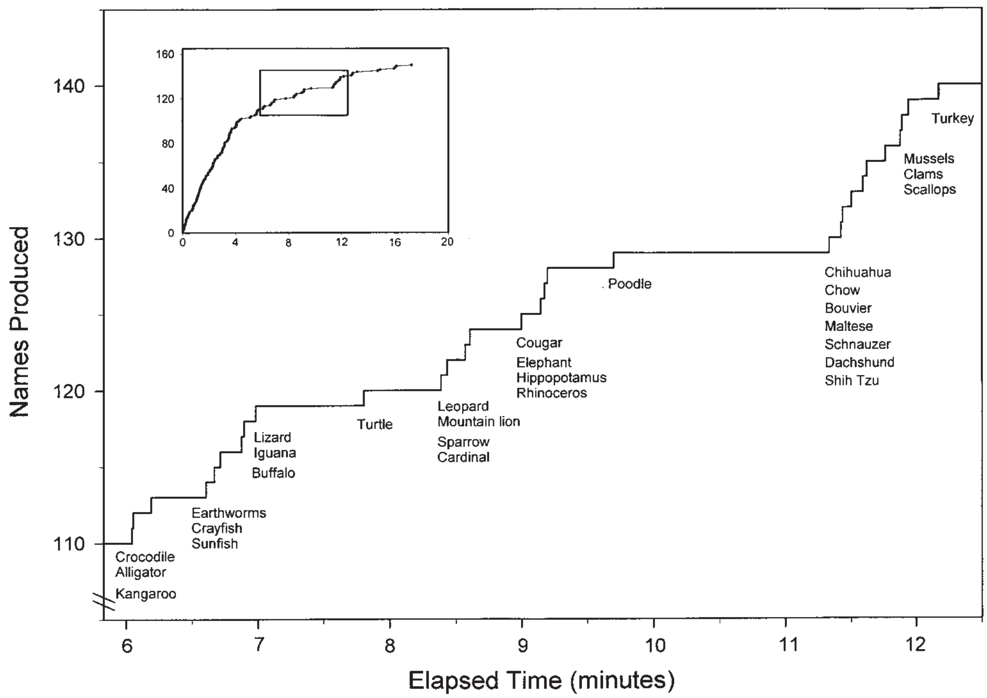

Figure 1. Cumulative number of animal names produced by a young adult over a 20-minute generation period. The full 20-minute temporal output pattern is shown in the insert in the upper left corner. The larger curve shows the output pattern and responses given by the participant from 6 minutes to just over 12 minutes into the production period. Where responses are listed vertically they are ordered from top to bottom in the order in which they were given.

period has been shown to be sensitive to a variety of neuropsychological disorders such as Alzheimer's disease, Huntington's disease, and traumatic brain injury (Kempler, Teng, Dick, Taussig, & Davis, 1998; Tombaugh, Kozak, & Rees, 1999). In addition, young adults typically produce more exemplars than do older adults, as do individuals with more years of formal education (Kempler et al., 1998; Tombaugh et al., 1999).

Unpublished work from our laboratory has shown that, even when given a full 20 minutes to respond, older adults on average still produce fewer category exemplars than their young adult counterparts. The age difference is thus not simply due to a general slowing of responses that would not allow the older adult time to give as many responses in a minute or two as would a young adult. The age difference in number of exemplars produced remained significant even when both young and older adults' curves were reaching asymptote with the extended time period. Other than this difference, however, the general shapes of the response curves and the appearance of response bursting in the older adults' responses were similar to that observed for the young adults.

Age Differences in the Temporal Output Pattern of Categorized List Recall

Following their early work on category generation, Bousfield and colleagues attempted a more focused approach to memory retrieval by shifting their efforts from retrieval from semantic memory (i.e., category exemplar production) to retrieval from episodic memory (Bousfield & Puff, 1964; Bousfield et al., 1954, 1958). A series of elegant studies by Pollio and his colleagues followed Bousfield by looking even more closely at the temporal pattern of recall from categorized lists (Pollio, 1974; Pollio & Gerow, 1968; Pollio, Richards, & Lucas, 1969). These were lists composed of words drawn from several different semantic categories, with the words scrambled in order of presentation.

As one considers the task of learning and recalling a categorized word list, one sees the questions that most researchers would consider among the central ones in the study of human memory: How is the stream of incoming sensory experience converted into a lasting memory trace? How are these traces represented and how are they bound to an internal representational of space and time that defines the context of the memory episode? During study, what are the control processes involved in effective encoding and organization of the studied material? During retrieval, what are the control processes involved in effective cue formation? How does that retrieval cue come in contact with the target information and redintegrate the target episode, and what information may be lost in recovery? For example, perhaps we will recover content information devoid of context (a failure of binding), or perhaps the retrieval cue will find a related (but incorrect) memory with the participant failing to inhibit the recovery of this extraneous information. How effective are individuals at focusing retrieval on the desired information and inhibiting the production of extraneous thoughts? And finally, how does the aging memory system differ from younger adults in these regards? These are questions that we suggest will engage memory researchers for many years into the future. At this point, we wish only to scratch the surface and to indicate some of the specific directions such future work might take.

In an attempt to explore potential age differences in the manner of recall, we conducted an experiment with young and older adults based directly on an experiment by Pollio et al. (1969) with young adults. The participants in our case were 15 young adults in their 20s and 15 older adults in their 60s and 70s, with the two groups matched for education and general verbal ability (Wingfield, Lindfield, & Kahana, 1998). The learning list was patterned after that used by Pollio et al., consisting of 25 words that consisted of five exemplars of each of five semantic categories (animals, gemstones, trees, transportation, vegetables). The 25 words were presented in random order, with the participant asked to recall as many of the words as possible in any order he or she wished. This procedure was repeated until the participant was able to recall all 25 words correctly. The order of presentation of the words was randomized on each presentation.

Our interest in this study was to compare the nature of the interresponse times (IRTs) produced by older adults and young adults learning the same list. One can immediately see the potential for a confound based on the traditional finding that older adults require more learning trials than the young to achieve equivalent levels of recall (Hultsch, 1975). Given this potential confounding between the occurrence of clustering in recall and the number of items recalled, we examined output latencies for young and older adults when both groups had reached a criterion of 100% recall. In this way, we were able to examine the age interaction with interresponse times (IRTs) without the confound of differences in degree of original learning.

Our question was not whether older adults would be expected to take longer in giving their responses than the young. One would expect to see such a main effect of slower responding for older participants (Cerella, 1994; Salthouse, 1991). Rather, our interest was in potential age differences in the pattern of growth rates in between-category and within-category IRTs.

Studies of young adults recalling categorized word lists have shown that both within- and between-category IRTs between successively recalled items tend to increase as recall proceeds. However, the rate of increase in between-category IRTs is significantly steeper than the rate of increase in within-category IRTs (Pollio et al., 1969). If it is the case that older adults suffer from a retrieval difficulty that is characterized by a differentially impaired episodic memory system relative to semantic memory, then one might expect to see a differential rate of growth in between-category IRTs as compared with that of within-category IRTs. This would be the case because within-category items in such lists share an already-learned semantic relationship, while the five categories chosen for the list have no previous connection.

The words were presented individually on a computer screen at a rate of one item per second. Items from the same semantic category were never presented in immediate succession and repetition of the items did not occur until all 25 words had been presented once. The end of the list presentation was signaled by a row of asterisks on the computer screen and an audible signal tone that was presented one second after the last word in the list appeared. When this signal occurred, participants were to attempt to recall as many words as possible from the studied list in any order they wished. Recall was given aloud into a desk-mounted microphone. Participants were told that their responses would be scored for recall accuracy, but no mention was made that we would also be examining the temporal output pattern of their responses, nor of the categorical nature of the list.

Although the rate of learning was faster for the young adults, the participants in both age groups predominantly recalled the items clustered by categories. That is, even though the words were presented in a random order, participants invariably recalled items grouped by category. In some instances, participants produced words from three or four different categories and then went back to produce a new item from an already recalled category. However, this was rare (less than 2% of all trials). These few out-of-category responses were excluded from the IRT analyses. Intrusion errors (production of responses that were not in the list) were also exceedingly rare.

The two participant groups' accuracy and recall latencies are summarized in the four panels in Figure 2. The top left panel shows the first learning and recall trial for the young adults, and the top right panel

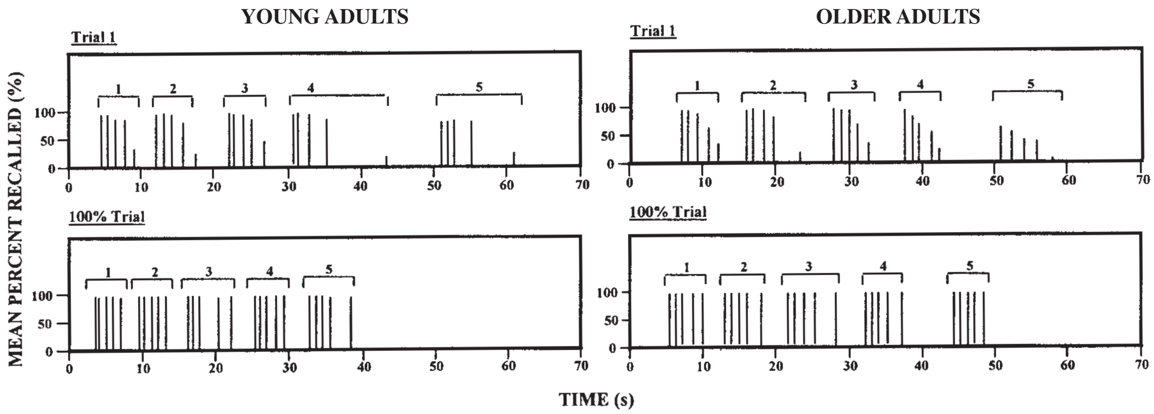

Figure 2. Recall accuracy and temporal output pattern for a categorized word list for young (left panels) and older (right panels) adults. Recall accuracy at each output position is indicated by the height of the bars. The position of each bar on the abscissa indicates the mean latency from the end of the list to the onset of each response. Category groups are indicated by the numbered brackets above the clusters of responses. The two upper panels show data for the first learning and recall trial. The two lower panels show data for the list when it was learned to a level of 100% recall accuracy. (Adapted from Wingfield, Lindfield, & Kahana, 1998, Fig. 1, pg. 259. Copyright 1998 by the American Psychological Association.)

shows these data for the older adults. The positioning of the vertical bars along the abscissa in all four panels represents the mean latencies (in seconds) from the signal to begin recall at the end of a list to the onset of each of the participants' responses. (These measures were obtained by digitizing the real-time recordings of participants' spoken recall responses and using a computer-generated visual display of the speech waveform to measure elapsed time from the recall signal at the end of the study list to the onset of each successive word recalled. Onset-to-onset latencies, rather than latencies from the endings of the words to the onsets of the next words, was chosen because of participants' tendency to drag out word endings as they attempted to think of the next word.)

Although participants predominantly recalled items within categories, the order of the items within a category and the order of the categories in their reports varied between participants. The numbered brackets above the responses indicate the participants' first, second, third, fourth, and fifth categories recalled. The heights of the bars within each category group show the mean percentage of participants recalling one, two, three, four, or all five items from within each category. It can be seen from the heights of the bars for the first learning and recall trial that participants' recall tended to fall off for the last category recalled, and this was especially so for the older participants. Indeed, one can see from the height of the bar that some older adults in Trial 1 were unable to recall even the first item of the fifth category.

This study was not the first one to examine age differences in category clustering in recall of categorized lists. As Kausler (1994) noted in his review of this literature, however, such studies have tended to have a natural confound between the occurrence of clustering in young and older adults' recall and the number of items recalled (Kausler, 1994, pp. 230-238). To overcome this confound, we show in the two lower panels the performance for the young (left panel) and older (right panel) adults when both groups had achieved 100% recall of all 25 list items.

In all four panels, we see the classic pattern of "response bursting," in which categorically related items are recalled in rapid succession with longer interresponse times separating the category groups. These findings are an exact replication of those observed by Pollio and colleagues for young adults (Pollio & Gerow, 1968; Pollio et al., 1969), findings that we can see also hold for older adults. For both age groups, we see that words presented in a random order are nevertheless recalled predominantly grouped by category, with short IRTs between items recalled from within a category, and longer IRTs at the transitions between categories. We see this result for both age groups on the first trial and again for both age groups on the 100% correct trial. That is, although the output rates are generally faster on the final trial, both within and between categories, the systematic "response bursting" by category evident in Trial 1 is still in evidence for the two groups on the 100% correct trial. Within this overall generality, however, important age differences appear. These differences take the form of an age dissociation between the within- and the between-category output latencies.

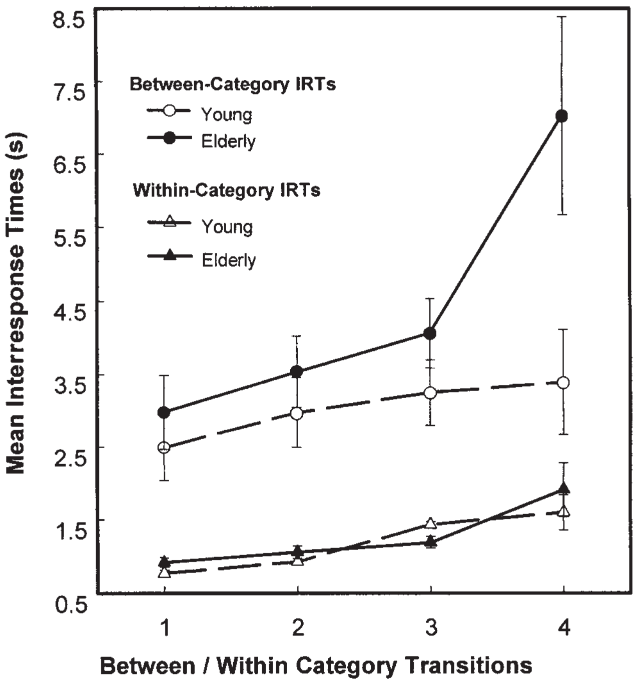

Figure 3. The top two curves show the average of the four between-category interresponse times (IRTs) for young and older adults and the bottom two curves show the mean IRTs for the four within-category transitions. These data are shown for the 100% trial on which all 25 items were produced. (From Wingfield, Lindfield, & Kahana, 1998, Fig. 2, pg. 260. Copyright 1998 by the American Psychological Association.)

Within-Category and Between-Category Recall: An Age-Dissociation

A clear way to see the pattern of the within-category and between-category IRTs on the 100% trial, where both groups had recalled all 25 items, is to plot the data from the lower two panels of Figure 2, as we have done in Figure 3. The upper two curves in Figure 3 show the young and older participants' mean IRTs for the four between-category transitions on the 100% trial. The two lower curves show the young and older participants' mean within-category IRTs as they proceeded through the five items from each of the five categories.

This figure shows clearly the finding that the older adults had significantly longer latencies than the young adults as they finished one category and attempted to retrieve the items for the next one. In addition, we see also that the rate of this increase from category to category was significantly steeper for the older participants than for the young participants, especially as participants executed their fourth transition to begin their recall of the items from the fifth category. These differences were confirmed by analysis of variance.

By contrast to these large age differences in the between-category IRTs for the young versus the older adults, the within-category IRTs were shorter and showed no significant age difference. Although there was a trend for IRTs to increase as participants moved from item to item within a category, the magnitude of this increase was very small compared to the latencies as participants moved from category to category.

It might be suggested that the rapid within-category IRTs resulted from participants simply generating highly prototypical category members as each category name was activated, rather than reflecting an actual recall response. Such a process, however, would be expected to produce not only responses that were correct but also intrusions of highly prototypical members of the studied categories that had not been in the list (cf. Deese, 1959; Roediger & McDermott, 1995). For both the young and older adults, however, such intrusions were extremely rare, suggesting that participants were not simply generating highly prototypical category items. The rarity of extralist intrusions, even among the older participants, is especially interesting in the context of the previously mentioned arguments for an inhibition deficit in normal aging (Hasher & Zacks, 1988; Zacks et al., 2000). Were such an inhibition deficit operating under these conditions one would have expected to see a higher incidence of intrusions, especially by the older adults, than we observed. This failure to find intrusions represents an important avenue for further study.

The meaning of between- and within-category IRTs has been characterized in a number of ways over the years since Bousfield et al. (1954) conducted their early studies of recall from categorized word lists. Patterson, Meltzer and Mandler (1971) took a similar position to Bousfield's original interpretation, postulating that between-category IRTs reflect the combined effects of the time needed to determine that no more items from a category can be recalled (category exit time) plus the time needed to gain access to the next category (category access time). Consistent with Tulving and Pearlstone (1996) and Tulving and Madigan's (1970) distinction between cue-dependent and trace-dependent retrieval, between-category IRTs could be seen as reflecting the accessibility of the higher-order memory units in which the list was organized by the learner (i.e., category names), while the within-category IRTs could be seen as reflecting the accessibility of the elementary memory units themselves.

To the extent that the long between-category IRTs reflect the time required to "find" the next category in memory, giving participants the names of the categories used in the study list either at the start of the study trials or at the time of recall should eliminate the

IRT increase as one moves from category to category. This finding was reported for young adults by Pollio and Gerow (1968). Giving participants the category names during learning and recall also eliminates the rise in between-category IRTs for older adults and brings their overall latencies more in line with those of younger adults (Wingfield et al., 1998, Experiment 2). Consistent with the notion that item-retrieval and category-retrieval represent different processes, presentation of category names had no effect on within-category IRTs.

In learning categorical word-lists, within-category items are part of a pre-existing semantic network. The learning task in this case involves primarily temporally based activation (i.e., activation by recency) of pre-existing associations. For example, if the words DIAMOND and RUBY are encountered on a study list, they are already bound together as part of the "gemstone" concept. One need only register the recency of their activation, or the time of their occurrence, as distinct from other gemstone names that were not in the list.

By contrast, the different categories represented in the learning list have little or no pre-existing connections between them. In this case, the participant must learn "animals," "gemstones," "trees," 'transportation," and "vegetables" as items that occurred in a common temporal context (i.e., the list). This kind of temporal association presents a special challenge to older adults (Naveh-Benjamin, 2000; Wingfield et al., 1998). It may be for this reason that the within-category responses are produced as rapidly by older adults as by the young adults whereas the between-category responses are not.

These findings from categorized list recall are consistent with Craik's (1986, 1994) view that older adults are particularly vulnerable in memory tasks he described as "self-initiated." These would include free-recall tasks unaided by environmental support from the presence of semantic context or other externally provided retrieval cues. By contrast, age differences are typically reduced or eliminated when environmental support is available from organizing cues present at the time of learning and recall (Craik & Jennings, 1992; Perry & Wingfield, 1994; Rankin & Firnhaber, 1986; West & Boatwright, 1983).

Organization and Learning: An Age Invariance

As we saw in our study of recall from categorized word-lists, young and older adults showed a marked difference in between-category recall that was not evidenced in within-category recall. Nevertheless, strong similarities also occurred: Both groups organized recall by categories and both groups showed longer latencies as they moved from category to category than they did when they moved from item to item within a category.

It has been suggested that older adults' slowness in their general rates of learning is due to an organizational deficit that retards learning by depriving the older learner of fully developed and available retrieval cues. It is well known that older adults generally exhibit less organization in learning than young adults (Hultsch, 1974; Rankin, Karol, & Tuten, 1984; Smith, 1980; Witte, Freund, & Brown-Whistler, 1993; Witte, Freund, & Sebby, 1990). Less clear, however, is whether the functional relationship between organization and learning is the same or different for young and older learners. This is so because their differences in mnemonic ability put them on a different scale. In our study of categorized word-list learning, we brought both young and older adults to the same recall level by giving older adults more learning trials than the young. When we did so, both groups showed an equivalent organization of recall by category membership.

Kahana and Wingfield (2000) examined this functional relationship between organization and learning in more detail. In this study, young and older participants were required to learn lists of 20 words, with lists consisting of five category exemplars drawn from each of four categories. As before, each list was presented in a series of learning and recall trials, with the list reordered in a different random sequence between each presentation. The learning-recall trials were continued until the list was mastered by the participant to a level of 100% recall. Participants received lists that consisted either of items that were highly prototypical of their category, or were less prototypical of their category based on judgment-ratings given by a separate group of participants.

As would have been expected, the older adults required significantly more trials to learn the lists than the young adults and, for both groups, high-prototypicality lists were learned faster than low-prototypicality lists. In addition to requiring additional trials to learn their lists, on a trial-by-trial basis the older adults exhibited less organization than their younger counterparts. Our experimental question, however, was whether older adults would show less semantic organization than the young adults when equated for their degree of learning. That is, we wished to plot the degree of organization of a list as a function of the number of items correct even though the older adults may have taken longer to achieve this level of performance. Our measure of organization was a modified version of Sternberg and Tulving's (1977) pair-frequency measure, which is based on the number of times that word pairs appear in adjacent output orders across trials. Detail of this method and calculation can be found in Kahana and Wingfield (2000).

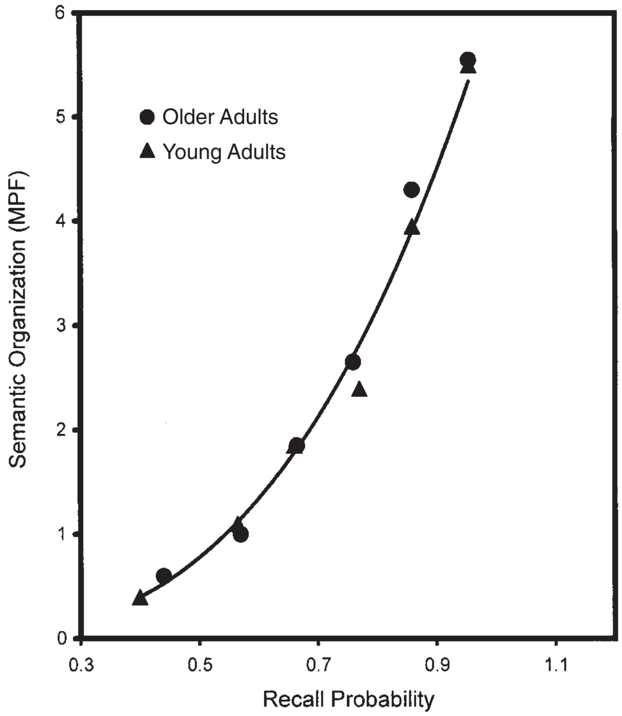

Figure 4. The relation between semantic organization, as measured by modified pair-frequency (MPF), shown on the y-axis, and recall probability, shown on the x-axis, for young and older adults. (Data from Kahana & Wingfield, 2000.)

In Figure 4, we show the buildup of organization over learning, not as a function of learning trials, but as a function of recall performance (i.e., percent correct recall). It can be seen that the young and the older adults' curves are virtually identical. This was true both for high-prototypicality and low-prototypicality lists, which are collapsed in this presentation of the data. That is, once we equated the young and the older adults for degree of learning, we demonstrated the same functional relation between learning and organization in free recall for the older adults as for the younger adults.

It may be interesting to note that in this study, as in the Wingfield et al. (1989) experiment, few errors of intrusion were observed. This was true either for intrusions from prior lists or from extra-list items. Such data do not negate an argument that older adults are often less able than young adults to inhibit inapproriate responses. This is a common finding in the aging literature, from studies of language procesing to studies of visual attention (Lahar, Isaak, & McArthur, 2001). It does suggest, however, that the associative connections formed in the list learning were strong enough to allow the retrieval system to distinguish between present-list and other items and to prevent production of the latter.

In both of the experiments just described, we observed that older adults took many more trials to learn categorized lists than did young adults. As we have also seen, however, once the categories had been learned, young and older adults exhibited the same organization as manifest in their order of recall. Whether lists are categorically organized or not, what items were present in the learning list must be determined by when they were encountered (i.e., their temporal context). Even when lists are categorically organized, one needs such temporal tagging to guide retrieval of the list items from among the larger array of items in semantic memory. It may be that the use of this temporal context underlies the age deficit in speed of list learning. In the next section, we describe some real-world implications of such a failure; in the final section, we consider the possible nature of this tagging deficit.

The Locus of the Age Impairment in Episodic Memory: Rich Encoding Versus Temporal Tagging

Many everyday memory failures can be due to mistaking a memory of an intended action for one that has actually been completed (Norman, 1981). This failure of "reality monitoring" (Johnson & Raye, 1981) can be as benign as "remembering" that one has put sugar in one's coffee when one has not. The cost of this failure would be to discover that one's coffee still needs sugar, or conversely, to add double the usual amount of sugar if the memory failure went the other way. A more serious concern is that one may think that one has taken one's medicine when one has not. This would result in a missed dose. The converse memory failure in this case would be the failure to remember that one has already taken one's medicine and, as a consequence, potentially taking additional pills beyond what should have been taken. This problem takes on extra importance in older adulthood where the amount of medication that needs to be taken per day can be quite large (Park & Jones, 1997).

Such failures to accurately distinguish between intended actions and completed actions may be due to a failure in effective temporal tagging for tasks that are routinely and frequently performed. In cases where an action is routinely performed (locking one's door, putting sugar in one's coffee, taking one's pills), the source of the failure would be in deciding not *whether* one has done these actions but *when*. That is, we may have a vivid and correct memory of our performance of the action in question. Our difficulty, however, may be in remembering whether we had completed the action a moment ago (putting sugar in this cup of coffee, locking the door when leaving the house this morning, or taking one's pills with this particular meal) or whether our vivid memory is actually a memory of

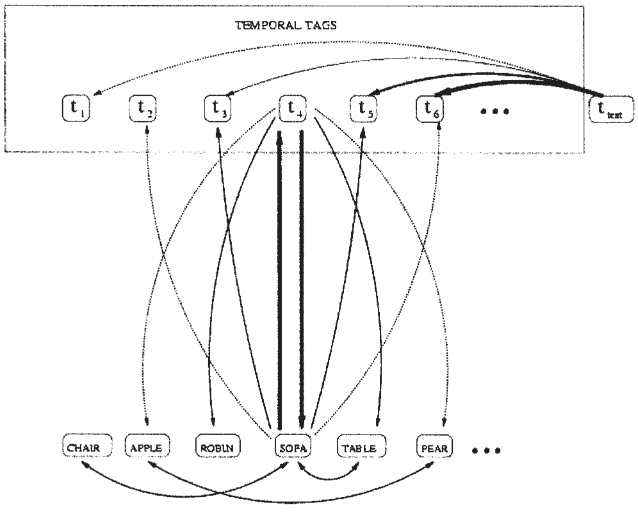

Figure 5. Schematic representation of temporal and semantic associations in word-list recall.

having done so on one of many previous occasions in which we have performed this identical act.

In the verbal learning laboratory when one "forgets" an item from a list (e.g., the word "cloud"), one would not say the participant has forgotten the word. The word cloud is still in the participant's lexicon. The difficulty, rather, is in determining whether the word had been heard a moment before (i.e., in the learning set), or whether the word had been heard some time before, perhaps spoken in conversation prior to the experiment. The most accurate recall, whether memory for an intended action or memory of a word in a word-list, is when that stimulus can be encoded using associations that would obviate the need for temporal dating as the only source of discrimination available. That is, a semantically rich memory code will be better than having to rely solely on temporal recency.

To the extent that older adults are less effective in detecting and rapidly encoding contextual information, as has been frequently argued (e.g., Burke & Light, 1981), so older adults would be forced to rely on temporal encoding alone to support their recall judgment. From this one might predict that older adults would be especially susceptible to memory failures traceable to mis-remebering an intended action for a completed one (Cohen & Faulkner, 1989). By contrast, if semantic connections can be easily made, or if the otherwise routine event is encoded in a way unique to other occasions when that task has been performed, older adults should show recall more in line with that of younger adults. In the following section, we propose a theoretical framework for understanding this age-related impairment in episodic memory.

Temporal Tagging and the Age Deficit in Episodic Memory

Figure 5 presents, in schematic form, the processes involved in encoding and retrieval of a list containing words drawn from several semantic categories. We illustrate this for a list of written or spoken words such as "CHAIR, APPLE, ROBIN, SOFA, TABLE, PEAR ..., " which are shown at the bottom of the figure in their order of presentation, going from left to right. At the top of the figure, we show a series of time tags, $\mathbf{t_1}$ through $\mathbf{t_6}$, which are associated with each of the six presented words. These time tags represent the relative positions of the items within a participant's autobiographical/episodic memory. Rather than thinking of these time tags as coding "clock" time, we think of them as bundles of elements whose values change slowly over the course of item presentation (e.g., Estes, 1950; Howard & Kahana, 2002; Mensink & Raaijmakers, 1988). As a consequence, nearby tags would share more elements than remote tags, endowing them with a relative representation of time.

As we outline in Figure 5, participants studying the word list form bidirectional associations between the items and their time tags (e.g., SOFA ⇨ $\mathbf{t_4}$; $\mathbf{t_4}$ ⇨ SOFA). At the time of test, the features that represent the current temporal tag, $\mathbf{t_{test}}$, can be used to cue memory. Due to its temporal proximity to recent tags, $\mathbf{t_{test}}$ is more likely to activate $\mathbf{t_6}$ than $\mathbf{t_1}$. This is represented in this example by the arrows of varying thickness going from $\mathbf{t_{test}}$ to less and less proximate tags. In this model, cueing with $\mathbf{t_{test}}$ will tend to activate items that have been recently experienced. This would account for the ubiquitous recency effect seen in self-initiated recall, and the fact that this recency effect depends solely on the relative (as opposed to absolute) spacing of list items. (Bjork & Whitten, 1974; Glenberg, Bradley, Stevenson, Kraus, Tkachuk, & Gretz, 1980).

After participants make their first response, this recalled item serves as part of the retrieval cue for subsequent recalls. As illustrated in Figure 5, recalling the word SOFA retrieves its time tag, $\mathbf{t_4}$, which will tend to activate temporally proximate items. We have illustrated this with the arrows running from $\mathbf{t_4}$ to ROBIN and to TABLE. This would account for the *lag-recency effect* in free recall, in which items temporally adjacent to a recalled item have the highest probability of being the next items recalled (Howard & Kahana, 1999; Kahana, 1996).

Semantic similarity also exerts an important influence on retrieval. This is represented by the arrows at the bottom of the figure linking items that belong to the same semantic category (e.g., SOFA - CHAIR, and SOFA - TABLE, and PEAR - APPLE). Because of the joint influence of semantic and temporal proximity on retrieval, participants would be more likely to retrieve TABLE following SOFA than ROBIN following SOFA.

We would suggest that the general form of this scheme is identical for both young and older adults. Our hypothesis is that older adults are impaired in their ability to utilize temporal proximity in guiding retrieval (e.g., going from SOFA to ROBIN), in contrast to the relative sparing of their ability to utilize pre-experimental semantic associations (e.g., going from SOFA to TABLE). Our recent finding that lag-recency effects become significantly diminished in older adulthood lends further support to this view (Kahana, Howard, Zaromb, & Wingfield, 2002).

We would argue that it is their deficiency in the use of temporal proximity that underlies older adults' general impairment in episodic memory tasks. Within our framework, this deficiency could take place at the level of the time tags themselves, or at the level of the associations between items and tags. We believe that further empirical work will be capable of distinguishing between these two possibilities.

Conclusions

It is generally true that those tasks that young adults find difficult, or that require a greater amount of time to perform, are especially so for older adults (Cerella, 1994; Cerella et al., 1980). This generality also holds for adult age differences in probability of recall of specific memory items in verbal memory studies (Stine & Wingfield, 1990). As we have tried to show, one of the most notable dissociations running counter to this generality is a sparing of semantic memory relative to significant age declines in episodic retrieval, especially when these tasks are self-initiated and must be performed in the absence of environmental support (Craik & Jennings, 1992).

In an attempt to understand the locus of older adults' difficulty with free recall, we have focused our theoretical interest on the temporal relations among items and their influence on retrieval. This focus suggests that the source of the age difference lies in a deficit in the formation and/or utilization of temporal coding of list items. Such a deficit would put older adults at a special disadvantage under conditions where temporal coding is the only or the primary retrieval cue available to the participant.

As we have indicated, one of the common cases where this condition occurs is in the attempt to distinguish between intended versus completed activities when the tasks are ones that are frequently and routinely performed. This is a case where claims have been made that older adults have special difficulty relative to young adults (Cohen & Faulkner, 1989), a finding that could be consistent with a temporal coding deficit in older adults along the lines we propose.

We believe that the resurgence of research interest in the dynamics of free recall (e.g., Kahana, 1996; Kahana & Loftus, 1999; Rohrer et al., 1995; Wixted & Rohrer, 1996), and its application to memory change in older adulthood, may offer valuable tools for our long-overdue understanding of these and related phenomena.

The authors acknowledge support from NIH Grants AG15852 and MH55687. We also gratefully acknowledge support from the W.M. Keck Foundation. We thank Jill Garland, Marc Howard, and Franklin Zaromb for valuable help in various stages of this research. Correspondence concerning this article and requests for reprints should be addressed to Arthur Wingfield, Volen National Center for Complex Systems, MS 013, Brandeis University, Waltham, MA, 02454-9110, U.S.A. (E-mail: wingfield@brandeis.edu).

References

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology, 6,* 173-189.

Bousfield, W. A., Cohen, B. H., & Whitmarsh, G. A. (1958). Associative clustering in the recall of words of different taxonomic frequencies of occurrence. *Psychological Reports, 4,* 39-44.

Bousfield, W. A., & Puff, C. R. (1964). Clustering as a function of response dominance. *Journal of Experimental Psychology, 67*, 76-79.

Bousfield, W. A., & Sedgewick, C. H. (1944). An analysis of sequences of restricted associated responses. *Journal of General Psychology, 30*, 149-165.

Bousfield, W. A., Sedgewick, C. H., & Cohen, B. H. (1954). Certain temporal characteristics of the recall of verbal associates. *American Journal of Psychology, 57*, 111-118.

Burke, D. M., & Light, L. L. (1981). Memory and aging: The role of retrieval processes. *Psychological Bulletin, 90*, 513-546.

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind, 11,* 63-65.

Cerella, J. (1994). Generalized slowing and Brinley plots. *Journal of Gerontology: Psychological Sciences, 49*, P65-P71.

Cerella, J., Poon, L. W., & Williams, D. M. (1980). Age and the complexity hypothesis. In L. W. Poon (Ed.), *Aging in the 1980s* (pp. 332-340). Washington, DC: American Psychological Association.

Cohen, G., & Faulkner, D. (1989). The effects of aging on perceived and generated memories. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 222-243). Cambridge, UK: Cambridge University Press.

Craik, F. I. M. (1977). Age differences in human memory. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the* *psychology of aging* (pp. 384-420). New York: Van Nostrand Reinhold.

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities, mechanisms, and performances* (pp. 409-422). North Holland: Elsevier.

Craik, F. I. M. (1994). Memory changes in normal aging. *Current Directions in Psychological Science, 3*, 155-158.

Craik, F. I. M., & Jennings, J. M. (1992). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 51-110). Hillsdale, NJ: Erlbaum.

Craik, F. I. M., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 474-479.

Cutler, S. J., & Grams, A. E. (1988). Correlates of self-reported everyday memory problems. *Journal of Gerontology: Social Science, 43*, S82-S90.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17-22.

Dobbs, A. R., & Rule, B. G. (1987). Prospective memory and self-reports of memory abilities in older adults. *Canadian Journal of Psychology, 41*, 209-222.

Estes, W. K. (1950). Toward a statistical theory of learning. *Psychological Review*, *57*, 94-107.

Fisk, A. D., & Fisher, D. L. (1994). Brinley plots and theories of aging: The explicit, muddled, and implicit debates. *Journal of Gerontology: Psychological Sciences, 49*, P81-P89.

Gilewski, M. J., & Zelinski, E. M. (1986). Questionnaire assessment of memory complaints. In L. W. Poon (Ed.), *Handbook for clinical memory assessment of older adults* (pp. 93-107). Washington, DC: American Psychological Association.

Glenberg, A. M., Bradley, M. M., Stevenson, J. A., Kraus, T. A., Tkachuk, M. J., & Gretz, A. L. (1980). A two-process account of long-term serial position effects. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 355-369.

Hartley, J. T. (1989). Memory for prose: Perspectives on the reader. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adulthood and later life* (pp. 135-156). New York: Cambridge University Press.

Hartley, J. (1993). Aging and prose memory: Tests of the resource-deficit hypothesis. *Psychology and Aging, 8*, 538-551.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193-225). San Diego, CA: Academic Press.

Howard, M., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25*, 923-941.

Howard, M., & Kahna, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology, 46,* 269-299.

Hultsch, D. F. (1974). Learning to learn in adulthood. *Journal of Gerontology, 29*, 302-308.

Hultsch, D. F. (1975). Adult age differences in retrieval: Trace-dependent and cue-dependent forgetting. *Developmental Psychology, 11*, 197-201.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 11*, 197-201.

Johnson, M. K., & Raye, C.L. (1981). Reality monitoring. *Psychological Review, 88*, 67-85.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory and Cognition, 24*, 103-109.

Kahana, M. J., Howard, M. W., Zaromb, F., & Wingfield, A. (2002). Age dissociates recency and lag recency effects in free recall. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28,* 530-540.

Kahana, M. J., & Loftus, G. (1999). Accuracy and latency in human cognition. In R. Sternberg (Ed.), *The concept of cognition*. Cambridge, MA: MIT Press.

Kahana, M. J., & Wingfield, A. (2000). A functional relation between learning and organization in free recall. *Psychonomic Bulletin & Review, 7*, 516-521.

Kausler, D. M. (1994). *Learning and memory in normal aging*. San Diego, CA: Academic Press.

Kempler, D., Teng, E. L., Dick, M., Taussig, I. M., & Davis, D. S. (1998). The effects of age, education, and ethnicity on verbal fluency. *Journal of the International Neuropsychological Society, 4*, 531-538.

Lahar, C. J., Isaak, M. I., & McArthur, A. D. (2001). Age differences in the magnitude of the attentional blink. *Aging, Neuropsychology and Cognition*, *8*, 149-159.

Laming, D. (1999). Testing the idea of distinct storage mechanisms in memory. *International Journal of Psychology, 34*, 419-426.

Light, L. L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Review of Psychology, 42*, 333-376.

Mensink, G. J. M., & Raaijmakers, J. G. W. (1988). A model for interference and forgetting. *Psychological Review*, *95*, 434-455.

Mitchell, D. B. (1989). How many memory systems? Evidence from aging. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 31-49.

Moscovitch, M., & Winocur, G. (1992). The neuropsychology of memory and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 315-372). Hillsdale, NJ: Erlbaum.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64*, 482-488.

Myerson, J., Wagstaff, D., & Hale, S. (1994). Brinley plots, explained variance, and the analysis of age differences in response latencies. *Journal of Gerontology: Psychological Sciences, 49*, P72-P80.

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*, 1170-1187.

Norman, D. A. (1981). Categorization of action slips. *Psychological Review, 38*, 1-15.

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory and Cognition, 25*, 555-583.

Park, D. C., & Jones, T. R. (1997). Medication adherence and aging. In A. D. Fisk & W. A. Rogers (Eds.), *Handbook of human factors and the older adult* (pp. 257-287). San Diego, CA: Academic Press.

Patterson, K. E., Meltzer, R., & Mandler, G. (1971). Inter-response times in categorized free recall. *Journal of Verbal Learning and Verbal Behavior, 10*, 417-426.

Perry, A. R., & Wingfield, A. (1994). Contextual encoding by young and elderly adults as revealed by cued and free recall. *Aging and Cognition, 1*, 120-139.

Pollio, H. R. (1974). *The psychology of symbolic activity*. Reading, MA: Addison-Wesley.

Pollio, H. R., & Gerow, J. R. (1968). The role of rules in recall. *American Journal of Psychology, 81*, 303-313.

Pollio, H. R., Richards, S., & Lucas, R. (1969). Temporal properties of category recall. *Journal of Verbal Learning and Verbal Behavior, 8,* 529-536.

Rankin, J. L., & Firnhaber, S. (1986). Adult age differences in memory: Effects of distinct and common encodings. *Experimental Aging Research, 12*, 141-146.

Rankin, J. L., Karol, R., & Tuten, C. (1984). Strategy use, recall, and recall organization in young, middle-aged, and elderly adults. *Experimental Aging Research, 10,* 193-196.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 803-814.

Rohrer, D., Wixted, J. T., Salmon, D. P., & Butters, N. (1995). Retrieval from semantic memory and its implications for Alzheimer's disease. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 1127-1139.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum.

Schacter, D. L., Chiu, C., & Ochsner, K. N. (1993). Implicit memory: A selective review. *Annual Review of Neuroscience, 16,* 159-182.

Smith, A. D. (1980). Age differences in encoding, storage, and retrieval. In L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg, & L. W. Thompson (Eds.), *New directions in memory and aging* (pp. 23-46). Hillsdale, NJ: Erlbaum.

Sternberg, R. J., & Tulving, E. (1977). The measurement of subjective organization in free recall. *Psychological Bulletin, 84*, 539-556.

Stine, E. A. L., & Wingfield, A. (1988). Memorability functions as an indicator of qualitative age differences in text recall. *Psychology and Aging, 3*, 179-183.

Stine, E. A. L., & Wingfield, A. (1990). The assessment of qualitative age differences in discourse processing. In T. M. Hess (Ed.), *Aging and cognition: Knowledge organization and utilization*, (pp. 33-92). New York: Elsevier.

Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology, 14*, 167-177.

Tulving, E., & Madigan, S. A. (1970). Memory and verbal learning. *Annual Review of Psychology, 21*, 437-484.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5*, 381-391.

Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging, 13*, 230-

West, R. L., & Boatwright, L. K. (1983). Age differences in cued recall and recognition under varying encoding and retrieval conditions. *Experimental Aging Research, 9*, 185-189.

Wingfield, A., Lindfield, K. C., & Kahana, M. J. (1998). Adult age differences in the temporal characteristics of category free recall. *Psychology and Aging, 13*, 256-266.

Wingfield, A., Tun, P. A., & Rosen, M. J. (1995). Age differences in veridical and reconstructive recall of syntactically and randomly segmented speech. *Journal of Gerontology: Psychological Sciences, 50B*, P257-P266.

Witte, K. L., Freund, J. S., & Brown-Whistler, S. (1993). Adult age differences in free recall and category clustering. *Experimental Aging Research, 19*, 15-28.

Witte, K. L., Freund, J. S., & Sebby, R. A. (1990). Age differences in free recall and subjective organization. *Psychology and Aging, 5*, 307-309.

Wixted, J. T., & Rohrer, D. (1996). Analyzing the dynamics of free recall: An integrative review of the literature. *Psychonomic Bulletin and Review, 1,* 89-106.

Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of aging and cognition* (2nd ed., pp. 293-357). Mahwah, NJ: Erlbaum.

Sommaire

Les adultes plus âgés se plaignent communément d'avoir de plus en plus de difficultés à se souvenir d'événements qui sont survenus dans un passé récent. Dans le cadre de tâches en laboratoire, ce déficit causé par l'âge est plus prononcé dans des tests de rappel libre que ceux de mémoire de reconnaissance ou de rappel avec indices (Kausler, 1994). Le besoin d'engager la formation des indices de récupération qui facilitent l'accès à l'information voulue constitue un facteur dans cette vulnérabilité. Pour réussir, il faut aussi freiner les souvenirs connexes, mais non ciblés (Hasher et Zacks, 1988; Zacks, Hasher et Li, 2000).

Nous montrons comment les facteurs dans le processus de récupération peuvent se révéler par les formes de sortie temporelles du rappel à partir de grandes catégories (p. ex. des noms d'animaux et de temps réponse du rappel de listes de mots catégorisés). Nous citons des données de Wingfield, Lindfield et Kahana (1998) montrant que dans le rappel de mots catégorisés, tant les jeunes adultes que les plus âgés manifestent une sortie rapide de mots reliés par catégorie, suivi de délais relativement plus long entre les catégories (« une salve de réponses »). Nous montrons que les adultes plus âgés produisent les articles au sein d'une catégorie aussi rapidement que les jeunes adultes, mais que les adultes plus âgés montrent des délais beaucoup plus longs entre les catégories comparativement aux jeunes adultes. Nous voyons cela comme une conséquence d'une réserve relative de la mémoire sémantique chez les adultes plus âgés, combiné à une déficience de la récupération épisodique des catégories qui étaient présentes dans une liste.

Il est bien connu que les adultes plus âgés ont besoin de plus d'essais pour apprendre une liste de mots que les jeunes adultes et pour découvrir les relations sémantiques qui pourraient exister entre les articles d'une liste (Kausler, 1994). Nous citons les données de Kahana et Wingfield (2000) qui ont mesuré l'organisation en utilisant le nombre de fois que des paires de mots apparaissaient dans des ordres de sortie voisines entre les essais (se reporter à Sternberg et Tulving, 1977). Ces données montrent que lorsqu'on contrôle la capacité mnémonique d'ensemble en fonction des différences d'âge, la relation entre l'organisation et l'apprentissage demeure invariable entre les âges.

Nous suggérons que les adultes plus âgés sont particulièrement désavantagés lorsque le codage temporel est le seul indice ou l'indice primaire de récupération dont dispose le participant pour guider le rappel. Dans le quotidien cela peut se produire lorsqu'il faut distinguer entre les activités envisagées et celles terminées lorsque les tâches sont accomplies fréquemment (p. ex. prendre ses médicaments). Nous offrons un cadre pour comprendre une déficience proposée en codage temporel chez les adultes plus âgés en notant que cette déficience pourrait avoir lieu au niveau de l'étiquette temporelle proprement dite ou au niveau des associations entre les articles et les étiquettes. Nous suggérons qu'il faut entreprendre d'autres études empiriques pour être en mesure d'établir une distinction entre ces deux possibilités.