Modeling age-related differences in immediate memory using SIMPLE☆

Aimée M. Surprenanta,*, Ian Neatha, and Gordon D.A. Brownb

a Psychology Department, Memorial University of Newfoundland, St. John’s, NL, Canada A1B 3X9
b University of Warwick, UK

Abstract

In the SIMPLE model (Scale Invariant Memory and Perceptual Learning), performance on memory tasks is determined by the locations of items in multidimensional space, and better performance is associated with having fewer close neighbors. Unlike most previous simulations with SIMPLE, the ones reported here used measured, rather than assumed, dimensional values. The data to be modeled come from an experiment in which younger and older adults recalled lists of acoustically confusable and non-confusable items. A multidimensional scaling solution based on the memory confusions was obtained. SIMPLE accounted for the overall difference in performance both between the two age groups and, within each age group, the overall difference between acoustically confusable and non-confusable items in terms of the MDS coordinates. Moreover, the model accounted for the serial position functions and error gradients. Finally, the generality of the model’s account was examined by fitting data from an already published study. The data and the modeling support the hypothesis that older adults’ memory may be worse, in part, because of altered representations due to age-related auditory perceptual deficits.

Keywords

Memory; Memory models; Aging and memory

Brown, Neath, and Chater (2002; see also Neath and Brown 2006) proposed a local distinctiveness model of memory called SIMPLE (Scale Invariant Memory and Perceptual Learning) in which memory retrieval, regardless of the paradigm, is viewed as discrimination of items in terms of their location along one or more dimensions. Items with fewer close neighbors on the relevant dimensions will be better remembered than items with more close neighbors. In a typical simulation, the researcher makes an assumption about what the relevant underlying dimensions should be. For many simulations, the choice of dimensions is relatively uncontroversial; for example, for free recall, the primary dimension is typically assumed to be time (Brown et al., 2002) and for absolute identification of tones, it is assumed to be frequency (Neath, Brown, McCormack, Chater, & Freeman, 2006). For other tasks, the exact nature of the underlying dimensions can be less clear: Both time and position have been suggested as the primary dimension for immediate serial recall tasks (see the discussion in Lewandowsky, Brown, Wright, & Nimmo, 2006), and both an item and a class dimension have been used when modeling the effects of word length (Hulme, Surprenant, Bireta, Stuart, & Neath, 2004).

☆ This research was supported by National Institute on Aging Grant AG021071 awarded to A.M.S. and I.N. and ESRC Grant RES 000 231038 to G.D.A.B. A preliminary account of part of this work was presented at the 46th Annual Meeting of the Psychonomic Society, Toronto, November 2005, and at the 11th Cognitive Aging Conference, Atlanta, April 2006.

* Corresponding author. Fax: +1 709 737 2430., E-mail address: asurprenant@mun.ca (A.M. Surprenant).
Although these assumptions may be reasonable, Neath and Brown (2006, p. 220) noted that “a more comprehensive approach could make use of an independently derived scaling solution to derive items’ locations in psychological space” (see also Nosofsky, 1992). One goal of the current paper is to do just that: We use a multidimensional scaling (MDS) procedure to determine the underlying dimensions used when people recall lists of acoustically confusable and non-confusable stimuli. If the basic idea behind SIMPLE is sound, the difference in recall of confusable and non-confusable items should be explicable in the model solely through differences in the representation of items.

A second goal is to assess the generality of SIMPLE by examining whether it can explain age-related differences in memory. Because SIMPLE was designed to account for memory performance in younger adults, it is not obvious that it can also explain age-related differences. However, given that memory representations are related to perceptual quality (Conrad, 1964; Surprenant & Neath, 1996) and older adults invariably have higher auditory thresholds than younger adults, we predict that they will have less distinct representations (as measured by the MDS procedure) than younger subjects. According to SIMPLE, this should result in worse overall performance. In addition to explaining the difference in recall of confusable and non-confusable items, SIMPLE might also be able to explain the difference in recall levels between younger and older adults through differences in item representations.

The third goal is to use the data to compare two extant versions of SIMPLE. Brown et al. (2002; Neath and Brown 2006) assumed that one of the primary dimensions used in immediate serial recall is temporal, and developed SIMPLE with relative time as the primary dimension that maintains order information. In contrast, Lewandowsky and colleagues (Lewandowsky et al., 2006; Lewandowsky, Duncan, & Brown, 2004) developed an otherwise identical version of SIMPLE that used a positional dimension in addition to a temporal dimension. The experiment is designed to provide data that would be useful in evaluating these two different assumptions.

SIMPLE

SIMPLE is an exemplar model of memory and has been described in detail elsewhere (see, for example, Brown et al., 2002; Neath & Brown, 2006; for its relation to other distinctiveness models of memory, see Neath & Brown, in press). Here, we present a brief overview. According to the model, memory is fundamentally a discrimination task. One factor that critically affects whether an item will be accurately recalled is the item’s distance from close neighbors in psychological space. At the heart of the model, then, is the similarity between representations. The similarity, \( \eta_{i,j} \), between two memory representations with values \( M_i \) and \( M_j \) on a psychological dimension is given by Eq. (1):

\[
\eta_{i,j} = e^{-c|M_i - M_j|^\alpha}
\]  

(1)

As in many other models, it is assumed that similarity falls off as a decreasing function of the separation between any two representations on the internal scale (Shepard, 1987). The main free parameter in SIMPLE is \( c \): Higher values of \( c \) correspond to greater distinctiveness of memory traces, i.e., less influence of more distant items. The additional parameter \( \alpha \) specifies the form of the similarity-distance function (e.g., exponential when \( \alpha = 1 \); Gaussian when \( \alpha = 2 \)). Eq. (1) can easily be extended to multiple dimensions (see Nosofsky, 1992). For example, consider a task that uses both a position dimension and a list dimension. Each dimension would be weighted, but the weights used must sum to 1.0, i.e., \( W_P + W_L = 1.0 \). In this case, the similarity between items \( i \) and \( j \) will be given by
\[ \eta_{i,j} = e^{-c(W_P \mid P_i - P_j \mid + W_L \mid L_i - L_j \mid)^{\alpha}} \]  

(2)

where \( P_i \) and \( L_i \) are the values of item \( i \) in memory on the position and list dimensions, respectively.

The probability of producing the response associated with item \( i, R_i \), when given the cue for stimulus \( j, C_j \), is given by Eq. (3), in which \( n \) is the number of items in the set:

\[ P(R_i \mid C_j) = \frac{\eta_{i,j}}{\sum_{k=1}^{n} \eta_{j,k}} \]  

(3)

Although additional assumptions and equations are needed for a complete instantiation of SIMPLE (see below for more details), these equations describe the basic version of SIMPLE and are sufficient to describe the main data of current interest. As the relative distinctiveness of a given item decreases (i.e., as similarity to near items increases), the probability of correctly recalling that item will decrease. Because of this, SIMPLE offers a natural explanation for the finding that lists of items that sound similar (e.g., B D G P T V) will be recalled less accurately than lists of items that sound different (e.g., F K L M Q R), the so-called acoustic confusion effect.

Although SIMPLE was not designed to explain age-related differences in memory, essentially the same explanation can be invoked to explain the lower levels of recall of older adults: The deficits in perceptual processing that occur during normal aging result in less distinct memory representations of older adults and thus worse memory compared to younger adults. We briefly review support for this idea.

### Perceptual deficits and reduced cognitive processing

It is well known that there is a decline in many types of memory performance as people grow older (see Zacks, Hasher, & Li, 2000, for a recent review). It is quite clear that this decline is not due to any single cause. On the contrary, multiple factors must be considered in order to generate a comprehensive account of the effects of aging on memory. Some of the major variables that have been identified as factors contributing to the difficulties in cognitive processing experienced by older adults in general, and memory in particular, are reductions in attentional or working memory capacity (Baddeley, 1986; Craik, 1986), slowed speed of processing (Salthouse, 1996) and lack of inhibitory control (Hasher & Zacks, 1988; Hasher, Zacks, & May, 1999).

In addition to those factors, recent research has demonstrated a strong relation between basic sensory/perceptual capabilities and cognitive functioning. For example, Baltes and Lindenberger (1997; Lindenberger and Baltes 1994) reported that up to 70% of the variance in measures of intellectual ability for subjects ranging in age from 25 to 101 years could be accounted for by a composite score that included age, vision and hearing abilities. In a separate study, speed of processing effects on intelligence were entirely mediated by vision and hearing scores (Lindenberger & Baltes, 1994). More recent experiments and reviews have further documented the substantial relations between sensory and cognitive functioning (e.g., Schneider, Daneman, & Pichora-Fuller, 2002; Scialfa, 2002). This effect of perceptual difficulties on cognitive processing has been called the information-degradation hypothesis (Schneider & Pichora-Fuller, 2000).

Although it is tempting to conclude that perceptual deficits can cause higher level cognitive deficits (the perceptual degradation hypothesis), the correlational nature of the previous studies
prevents a strong claim about the direction of causality. It is possible that the reverse is true, i.e., that cognitive declines cause a depletion of resources at a higher level which takes away resources that would normally be devoted to the perceptual system (the cognitive load on perception hypothesis). A third possibility was suggested by Lindenberger and Baltes (1994): Widespread neural degeneration might cause both perceptual and cognitive deterioration (the common cause hypothesis). A fourth possibility is that the relation between perception and cognition is highly complex because the two systems are very highly integrated and interdependent. Schneider and Pichora-Fuller (2000), for example, proposed an integrated system model of shared resources in which the flexible allocation of resources is a key ingredient. This model essentially combines all of the above hypotheses and suggests that each one could play a role at any point in time.

**Perceptual deficits and reduced memory performance**

There have been a number of studies demonstrating an effect of perceptual degradation (due to noise or hearing loss) on memory performance. Rabbitt (1991) found that even mild peripheral sensory hearing loss resulted in a substantial reduction in the number of words recalled, compared to an age-matched control group with no hearing loss, even though identification performance (measured by shadowing the presented stimuli out loud) was essentially perfect for both groups. Similarly, Pichora-Fuller, Schneider, and Daneman (1995), using a variant of the speech-perception-in-noise test (Bilger, Nuetzel, Rabinowitz, & Rzeczkowski, 1984), showed that older adults with mild hearing impairment recalled fewer items than younger subjects even when levels of identification were equated. Both Pichora-Fuller et al. (1995) and Rabbitt (1991) suggested that the added difficulty experienced by listeners with mild hearing loss depleted resources that might otherwise have been used to elaborate encode and rehearse the materials. Thus, although the distortion of the signal had no overt effect on identification (Rabbitt, 1991), and identification performance was equated (Pichora-Fuller et al., 1995), there were still effects of hearing loss and aging on memory and comprehension (see also Rabbitt, 1990; Surprenant, 1999, in press). Similar results have been reported when the input modality is visual (Salthouse, Hancock, Meinz, & Hambrick, 1996; Schneider & Pichora-Fuller, 2000).

Although the reduced resources idea is a compelling one, it is also possible that the noise or distortion at encoding leads to the creation of a more impoverished, less discriminable memory trace. One set of studies argues in favor of such an interpretation. Surprenant and Neath (1996) measured identification and recall of lists of consonants, vowels, and silent-center syllables. They reported a strong relationship between the probability of recalling an item and the perceptual discriminability of that item. However, it was possible to manipulate the items so that the opposite result sometimes obtained: some items that were poorly identified were nonetheless well recalled. One way of explaining this is to assume, in agreement with many kinds of dual coding theories (e.g., Fujisaki & Kawashima, 1970; Nairne, 1990; Pisoni, 1973, 1975), that the discriminability of an item includes at least two types of representation: one physical (or sensory, or modality dependent) and one abstract (or conceptual, or modality independent). Surprenant and Neath concluded that the relative discriminability of the memory representation, including both types of coding, affects recall of the item.

This type of analysis is clearly compatible with the underlying assumptions of SIMPLE: Recall in the model is dependent on the relative discriminability of the to-be-remembered items along multiple dimensions. If one of those dimensions is altered, whether it be through noise, age-related hearing loss, or acoustic discriminability, it will affect recall in predictable ways. Because older adults invariably experience some hearing loss, comparing memory in older and younger adults for various types of auditory stimuli is an ideal situation in which to test this prediction.
Experiment

Despite the large number of studies that demonstrate overall deterioration in measures of accuracy or latency in certain types of memory tasks as people age, there have been very few which report the measures necessary in order to assess the hypothesis that perceptual deficits are related to declines in memory performance and to fit SIMPLE to the data. In a recent review of 288 published studies on cognitive aging, Schneider and Pichora-Fuller (2000) found that less than 25% of the investigators measured visual or auditory acuity. The majority of studies either ignored the issue or obtained self-reports of auditory or visual abilities. In particular, we need (1) assessments of hearing ability, (2) measurements of output time, and (3) detailed error data. The experiment reported here was quite similar to that reported by Maylor, Vousden, and Brown (1999) except that all subjects were given an audiometric assessment and output time to recall each item was measured. In this study, younger and older adults recalled 6-item lists of acoustically confusable or non-confusable CV syllables presented auditorily over headphones. In addition to the hearing screening, participants were given a standard vocabulary test (Salthouse, 1993) in order to ensure that the level of intellectual ability was similar between the groups.

Method

Subjects—Twenty Purdue University undergraduates (M age 19.1, range 18–24) participated in exchange for credit in introductory psychology courses. Ten were female, ten were male. An additional twenty older participants (M age 68.8, range 62–81) from the community participated in exchange for $10. Four of the older participants had high school diplomas, four described their education as ‘some college’, six had college degrees, and six had post-graduate experience. Twelve were female, eight were male. All described their health as ‘good’ to ‘excellent’ with the exception of one participant who was only in ‘fair’ health. None were taking any medication that affected cognitive functioning. All subjects (younger and older) were native speakers of American English.

Stimuli and design—The stimuli were digitized consonants B, D, G, P, T, and V (confusable) and F, K, L, M, Q, and R (non-confusable) spoken by a male talker and digitized onto a computer. Each token was approximately 450 ms in duration. The stimuli were presented over headphones that were calibrated with a 440 Hz pure tone to a level of 70 dB SPL. Each participant received 15 lists in each of the conditions in random order.

Procedure—Subjects were informed that we were examining how well they could remember the order in which a series of letters was presented. They heard lists of six letters over headphones, with each letter presented one at a time with a 1 s onset-to-onset interval. After the last letter in the list had been presented, buttons labeled with the six confusable and six non-confusable letters appeared on the computer screen in alphabetical order. The subjects were asked to reconstruct the original order of presentation by clicking the letters (buttons) in the order that they had originally been presented using the mouse (i.e., strict serial reconstruction of order). When they had recalled as many letters as they could, even if this was fewer than six letters, they clicked on a button labeled ‘Next Trial’ to begin the next trial. Subjects were tested one at a time, and the experimenter remained in the room to make certain the subjects were following the instructions. No subject reported any difficulty using the mouse.

Screening procedures

Hearing test—A pure-tone hearing screening was performed and thresholds were recorded for 250, 500, 1000, 2000, 4000 and 8000 Hz for both the left and right ear. Mean audiometric thresholds for the two groups are shown in Table 1. The thresholds were analyzed using a 2
group (younger, older) × 2 ear (left, right) × 6 frequency mixed ANOVA. An alpha level of .05 was adopted for this and all other statistical tests.

As expected, the older and younger groups differed in hearing performance, with the younger group having better auditory acuity overall ($F(1, 38) = 56.75, MSE = 439$). There was a main effect of frequency ($F(1, 5) = 22.8, MSE = 95.35$) and a frequency by group interaction ($F(5, 190) = 26.65, MSE = 95.35$). This pattern is the usual one found for age-related hearing loss with greater loss at the higher frequencies. Note that the older subjects had thresholds that are within the range considered to be clinically normal up to 2 kHz. None of the other main effects or interactions were significant.

Vocabulary test—Items on the vocabulary test were adapted from those used by Salthouse (1993). Subjects were given a word and were asked to identify which of five possible alternative words was either a synonym or an antonym of the target word. There were ten synonyms and ten antonyms on the test. The vocabulary test was included in order to rule out the possibility that any differences between the groups was due to a difference in the cognitive capabilities of our older adults. The older subjects performed significantly better than the younger subjects on both the synonym (7.1 out of 10 correct versus 4.1) and antonym (6.2 versus 4.1) vocabulary tests (synonyms, $t(19) = 2.42$; antonyms, $t(19) = 2.73$).

Results and discussion

Correct recall—Each response was counted as correct only if it was the correct letter recalled in the correct position. Mean performance as a function of serial position, list type and age group is shown in the left panel of Fig. 1. The two main results of interest can be readily seen in the figure: (1) recall of confusable items was worse than non-confusable items, replicating the acoustic confusion effect (e.g., Baddeley, 1966; Conrad, 1964), and (2) performance was better for the younger subjects and worse for the older subjects, replicating age-related differences in short-term memory (e.g., Belleville, Peretz, & Malenfant, 1996; Maylor et al., 1999).

The data were entered into a $2 \times 2 \times 6$ mixed ANOVA with age group (younger, older), list type (confusable, non-confusable) and serial position (1–6) as factors. There was a main effect of group ($F(1, 38) = 18.72, MSE = 0.20$), with the younger subjects recalling more items correctly than the older subjects, 0.674 versus 0.501, respectively. There was also a main effect of list type ($F(1, 38) = 268.00, MSE = 0.05$), with better performance for non-confusable than confusable items, 0.743 versus 0.431. The main effect of position was significant ($F(5, 190) = 80.03, MSE = 0.02$), as was the interaction of list type by position ($F(5, 190) = 11.20, MSE = 0.02$). The three-way interaction of group, list type, and position was significant ($F(5, 190) = 3.20, MSE = 0.02$). None of the other interactions were statistically significant.

One result apparent in Fig. 1, and partly the reason for the significant three-way interaction, is the lack of any recency in the confusable items for the older subjects. One interpretation is that performance by the older adults is close to a floor effect. Although the older subjects showed a slightly larger difference than younger subjects between recall of confusable and non-confusable items at early positions, the difference was approximately the same at later positions.

Output time—The time to make each response was measured for both groups of subjects, and the data are shown in the right panel of Fig. 1. Overall, the younger subjects responded more quickly than the older subjects, and responses were faster when recalling lists of non-confusable items compared to lists of confusable items.
The response times for correct responses were submitted to a $2 \times 2 \times 6$ mixed ANOVA with age group (younger, older), list type (confusable, non-confusable) and serial position (1–6) as factors. There was a main effect of group, such that overall, older subjects took more time to respond than younger subjects, 1598 versus 1171 ms, respectively ($F(1, 28) = 11.94, \text{MSE} = 671178$). There was also a main effect of list type, with slower responding to confusable than to non-confusable lists, 1496 versus 1273 ms ($F(1, 28) = 20.00, \text{MSE} = 310492$). There was a main effect of position, ($F(5, 140) = 9.58, \text{MSE} = 633757$) and an interaction between position and age ($F(5, 140) = 6.66, \text{MSE} = 633757$). This interaction was due to older subjects taking less time to initiate recall than younger subjects, but taking more time to respond to later list items. Neither the list type by position interaction ($F(5, 140) = 1.46, \text{MSE} = 307431, p > .20$) nor the three-way interaction ($F(5, 140) = 1.59, \text{MSE} = 30.7431, p > .15$) were reliable.

Older adults took, on average, more time to make their responses than younger adults, but surprisingly, they initiated recall more quickly than younger subjects. This latter finding is an interesting result that, as far as we are able to determine, has not been reported before. It remains to be seen whether this is a stable finding or in some way was due to our experimental setup. There was also a particularly pronounced increase in response time approximately half way through recall of the confusable items for the older group. One possibility is that this is due to spontaneous grouping on the part of the older subjects (e.g., Maybery, Parmentier, & Jones, 2002).\footnote{We thank a reviewer for raising this possibility.} An alternate interpretation is that it reflects the fact that the subjects are now trying to recall items that are essentially no longer available.

**Audiometric threshold and recall**—There was insufficient variability in the audiometric thresholds of the younger subjects to perform any meaningful analyses on those data. However, we could calculate the correlation between hearing and recall for the older group. The mean threshold for the tones with frequencies similar to those most important in speech (i.e., 1000i.e., 2000, and 4000 Hz) was calculated for each older subject. The correlation between the mean threshold and overall performance with the non-confusable items was $-0.152$ (ns) compared to $-0.564$ ($p < .01$) for the confusable items. These two correlations differed from one another, $t(17) = 2.46$.

Older subjects with worse hearing recalled fewer items in confusable lists than those with better hearing. This is consistent with the idea that older adults may already be performing close to their capability, and when the task is made more difficult (by using confusable items), performance abruptly drops (Surprenant, 1999, in press).

**Patterns of recall errors**—The position error gradients are shown in Fig. 2. These show the proportion of times each item was recalled in each of the six possible positions. Errors in recall were classified as movement errors, where an item that had been presented in the list was recalled in the wrong position; omissions, where subjects did not respond with six letters; and extra-set intrusions, where an item recalled was from the set of stimuli for the other condition (confusable or non-confusable set). These error rates along with proportion of items correct are shown in Table 2. In general, the older subjects made more errors than the younger subjects, and more errors were made for the confusable items than for the non-confusable items.

Each of these error types was entered into a separate $2 \times 2$ age group (younger, older) × 2 list type (confusable, non-confusable) mixed ANOVA. Older adults had more movement errors ($F(1, 38) = 10.34, \text{MSE} = 0.02$) and more omissions ($F(1, 38) = 6.98, \text{MSE} = 0.01$) than younger adults, but the two groups did not differ in the number of extra-set intrusions ($F(1, 38) = 2.29, \text{MSE} = .01, p > .10$). There were more errors of all types in the confusable than in the non-confusable lists: movements ($F(1, 38) = 79.63, \text{MSE} = 0.01$), omissions ($F(1, 38) = 8.29$, $p < .01$).
\(MSE = 0.02\), and extra-set intrusions \((F(1, 38) = 58.49, MSE = 0.01)\). Finally, there was a significant interaction between group and list type in the number of omissions \((F(1, 38) = 5.15, MSE = 0.01)\), with older adults omitting relatively more responses in the confusable compared to the non-confusable condition. None of the other interactions were significant.

The pattern of errors was similar to those also previously reported. In particular, the shape of the positional uncertainty functions is different for confusable and non-confusable lists (Henson, Norris, Page, & Baddeley, 1996; Maylor et al., 1999). There is a particular flattening of the function when the items in the list are very similar to one another (confusable lists). That is, when errors are made in the confusable lists, they often come from farther serial positions than just the adjacent positions. One interpretation of this result is that there are two sources of confusability in the similar lists: positional similarity and item similarity. In contrast, errors in the dissimilar lists come mainly from transposing adjacent items.

**Multi-dimensional scaling**—Along with correct recall in position, we can analyze confusion errors, that is, the number of times an individual, when trying to recall a particular syllable, actually recalls a different one. Conrad (1965) showed that recall confusions were not random but, instead, mimicked the errors made when identifying auditorily-presented letters in noise. We generated confusion matrices for each stimulus for each group of subjects. These were then used as input to the ALSCAL multidimensional scaling procedure in SPSS (Takane, Young, & DeLeeuw, 1977). We used the individual differences Euclidean distance (INDSCAL) model, as it ensures that the resulting dimensions are the same for the two groups. The data were analyzed using the unconditional option because the matrix conditional approach has the effect of normalizing the data to a common space size whereas we were interested in the size of each group’s perceptual space (Borg & Groenen, 2005). This procedure results in a single scaling solution for both groups, from which the individual group solutions can be recovered by multiplying each coordinate by the square root of that group’s weight for that dimension (Chan, Butters, Salmon, & McGuire, 1993).

The scree plot showed that the stress values leveled off with more than two dimensions and there was little improvement of the fit with three compared to two dimensions. The derived two-dimensional solutions for the younger and the older adults are shown in Fig. 3. For both the younger and older adults, the confusable items are in a different part of the space than the non-confusable items and are much more tightly packed along Dimension 1. In addition, the space for the confusable items is even more compact for the older adults than for the younger adults. Specifically, the mean cityblock distance separating confusable items was 0.126 for the older adults and 0.172 for the younger adults. The corresponding values for the non-confusable items are 1.665 for the older adults and 2.329 for the younger adults.

The finding that the MDS solution is more compact in the data from the older adults agrees with other data that suggest as adults age, there is what is known as “de-differentiation” of cognitive abilities (e.g., Li et al., 2004). It also is consistent with the idea that the memory representations of the older adults are more compressed than those of the younger adults.

In terms of SIMPLE, the MDS solution revealed that confusable items are generally closer together (and thus relatively less distinctive) than non-confusable items for both younger and older adults. Previous work with SIMPLE has simply assumed this to be case. The MDS solution also revealed that the space for older adults is more compact than that for younger adults. This is consistent with the assumption that at least some age-related differences in memory might be attributable to impoverished representations due to perceptual deficits.
Simulations

The three goals of the current work were (1) to see if the difference in recall of confusable and non-confusable items could be explained in SIMPLE solely through differences in the representation of the items, (2) to determine whether SIMPLE could account for age-related differences in a standard short-term memory task solely through differences in the representation of the items, and (3) to provide data to evaluate two different extant versions of SIMPLE, one of which assumes items are represented primarily on a temporal dimension and the other of which assumes items are represented primarily on a positional dimension.

Simulation 1

Simulation 1 explores whether SIMPLE can account for the age-related difference in recalling confusable and non-confusable items through the assumed difference in the representation of the items. The coordinates from the MDS solution shown in Fig. 3 were used as specifications of the to-be-remembered items’ locations in psychological space.

The first simulation used a version of SIMPLE that assumes that order information is represented solely in terms of position along a temporal distance dimension (Brown et al., 2002). Each item is represented in terms of time until output. The last item in the list is given a value that corresponds to the retention interval, the measured mean time to respond with the first item for that age group and condition. For example, when modeling recall of the first item, we first calculate the time that elapses between the item’s presentation and recall. For the young non-confusable condition, this value is 7.138 s. There is a 1 s onset to onset presentation time, so list presentation ends 5 s after the first item has been presented. However, the first response is not made until (on average) 2.138 s after presentation of the final item. (the retention interval; see the right panel of Fig. 1). The second item would thus have a value of 6.138 s, the third item would have a value of 5.138 s, and so on. These values then underwent a log transformation and were used to obtain predictions about recalling Item 1. When modeling recall of the second item, the times are adjusted by adding the additional output time, in this case 0.852 s. When modeling recall of the third item, the times are again adjusted to reflect output time, this time by 0.904 s, and so on.

There are thus three dimensions: the temporal dimension and the two MDS dimensions. Eq. (2) is rewritten to be:

\[ \eta_{i,j} = e^{-c(W_1 \cdot T_{i-j} + W_{D1} \cdot D1_{i-j} + W_{D2} \cdot D2_{i-j})^a} \]  

In order to fit the serial position data and error gradients, one change has to be made to enable SIMPLE to produce omissions. Brown et al. (2002) suggested using a sigmoid function to increase recall probabilities that are already high, and reduce recall probabilities for items whose recall probabilities are already low. Eq. (5) shows the implementation, which calculates output probability, \( P_o \), based on the estimated recall probability, \( P \), from Eq. (3):

\[ P_o = \frac{1}{1 + e^{-s(P - t)}} \]  

The parameter \( t \) is the threshold and parameter \( s \) is the slope of the transforming function (which can be interpreted as the noisiness of the threshold). For example, if \( t \) is set to 0.8 and \( s \) is very large, the transformation will approximate a system that recalls all items with relative strengths greater than 0.8, and omits all items with strengths less than 0.8. As \( s \) becomes smaller, the transition from low to high recall probabilities becomes more gradual.
In fitting the model, MDS coordinates were assigned to a particular serial position to reflect the fact that different list items appeared in different serial positions over trials. All 720 possible item orderings were used in the model simulation, and the average output obtained. All parameters except for $\alpha$ were held constant across age group; we allowed $\alpha$ to vary with age group on the basis of independent evidence that higher values of $\alpha$ (e.g., 2, corresponding to a Gaussian similarity-distance function) may best describe performance when the stimuli to be discriminated are sufficiently close in psychological space (Ennis, 1988; Nosofsky, 1988).

Then, the best fit was obtained by minimizing RMSD. The resulting parameter estimates for both groups were as follows: $c = 2.78$, $W_{D1} = 0.38$, $W_{D2} = 0$, $W_T = 0.62$, $s = 12.11$, $t = 0.36$. The $\alpha$ parameter was 0.66 for the younger group, and 0.81 for the older group. The resulting serial position curves are shown in the left panel of Fig. 4.

First, SIMPLE produces the appropriate acoustic confusion effect for both groups of subjects. Second, SIMPLE also produces the appropriate age-related difference in recall. Note that with the exception of $\alpha$ the same parameter values are used for all conditions; that is, the key difference in the model for each of the conditions is the representation of the items, which was determined independently by the MDS procedure and by measuring output times. As expected, $\alpha$ was higher for the older subjects. However the value was less than 1.0 for both groups, indicating that neither an exponential nor a Gaussian similarity-distance function was appropriate. We discuss this in more detail after the next simulation, where more typical values of $\alpha$ were obtained.

According to the model, memory retrieval is viewed as discrimination in terms of location along (in this case) three dimensions (the two identified by the MDS procedure and an additional temporal dimension for the representation of serial order). Items with fewer close neighbors on the relevant dimensions will be better retrieved than items with more close neighbors. The confusable syllables are located more closely together in space than the non-confusable items, and more items are located closer together for the older than for the younger subjects. The ability of SIMPLE to account for the major pattern of data solely through the inferred psychological space within which the items are located is consistent with the idea that the older adults have an impoverished perceptual representation. Given the measured difference in pure tone thresholds, this result is also consistent with the perceptual degradation hypothesis, i.e., the suggestion that the reduced hearing ability of the older subjects results in a less fine-grained memory representation.

This simulation showed that a more objectively determined solution to derive the items’ locations in psychological space enabled SIMPLE to account for both the differences in recall of confusable and non-confusable items and the difference in performance of younger and older subjects. This remedies the limitation identified with previous simulations of the acoustic confusion effect of not specifying precisely the dimensions of interest. It also suggests that the idea that age-related differences might be due to differences in the quality of the representation, due to more perceptual difficulties on the part of the older subjects, was not without merit. Within the SIMPLE framework, both of these differences are explainable in the same way.

Importantly, SIMPLE produces appropriate results even when the temporal values are measured rather than just estimated. This is important because if the model could not simulate the data, it would have called into question other simulations that relied on estimated (but plausible) output times.

The model also produced a reasonable fit to the observed serial position curves. However, rather less primacy was observed in the model than in the data. This may in part reflect the lack of proactive interference from previous lists in the simulation (Brown et al., 2002). However in the second simulation we consider the possibility that the reduced primacy reflects
the use of a temporal (rather than positional) dimension, coupled with a lack of output interference.

**Simulation 2**

Simulation 1 addressed the first two goals of the paper; Simulation 2 addresses the third goal, comparing two extant versions of SIMPLE. Originally, it was assumed that items in memory were primarily represented on a temporal dimension, regardless of whether the test was free or serial recall. With plausible assumptions about average output time, SIMPLE was able to account for a wide range of data from many different paradigms (see Brown et al., 2002; Neath & Brown, 2006).

A key issue concerns the extent to which temporal manipulations affect serial recall. There is no question that such effects are readily observable in free recall (e.g., Brown, Morin, & Lewandowsky, in press), but recent data have failed to demonstrate equivalent effects of temporal manipulations in serial recall (e.g., Lewandowsky et al., 2004). Because of this, a version of SIMPLE was proposed that augmented the temporal dimension with a positional dimension (e.g., Lewandowsky et al., 2006). The current data set provide an opportunity to compare the outputs of a purely positional version of SIMPLE (Lewandowsky et al., 2006) with those of a purely temporal version (Brown et al., 2002). The temporal and positional versions of SIMPLE both use the two dimensions indicated by the MDS solution and both use the same implementation to produce omissions. They differ only in whether the third dimension is based on relative time until retrieval or on serial position.

The positional version of SIMPLE (e.g., Lewandowsky et al., 2004) used serial position as the third dimension (e.g., the first item had a value of 1, the second item had a value of 2, and so on). Instead of implementing output interference as temporal delay, it was implemented by systematically reducing the value of $c$ (see Lewandowsky et al., 2004). Specifically, the value of $c$ used when calculating retrieval of item $n$ was multiplied by a new free parameter, $o$, raised to the power of $n-1$:

$$c_n = co^{n-1}$$

If $o = 1$, there is no output interference; as $o$ gets increasingly smaller than 1, there is increasing output interference.

As in the previous simulation, all parameters except for $\alpha$ were held constant across groups. The parameters were set as follows: $c = 5.06$, $o = 0.84$, $W_{D1} = 0.72$, $W_{D2} = 0.03$, $W_p = 0.25$, $s = 6.33$, $t = 0.53$. The $\alpha$ parameter was 1.03 for the younger group, and 1.63 for the older group. The results are shown in the right panel of Fig. 4. Consistent with recent data that suggests immediate serial recall tasks are not sensitive to variations in relative time (other factors being held constant), the simulation shows that a version of SIMPLE that uses position rather than temporal cues is able to account for the acoustic confusion effect and for age-related differences in performance. The major discrepancies in the fit are the slight under prediction of performance for younger subjects with non-confusable items, and slight over prediction of performance for older subjects with confusable items.

The values of $\alpha$ were consistent with the interpretation that the younger subjects’ performance was governed by an exponential similarity-distance function, while the older subjects were closer to Gaussian (or, due to averaging across subjects, a mixture of Gaussian and exponential).

How do the temporal and positional versions of SIMPLE compare? The positional version has an additional free parameter, $o$. However, the temporal version uses different temporal values.
for each position in each of the four groups. Although these values are not free parameters in the traditional sense, they do provide additional differences between the groups over and above the MDS coordinates. The positional version fit the data only slightly less well than the measured temporal version.\(^2\) One reason for the relative lack of difference between the various versions is that both use the MDS coordinates. The relative equality of the temporal and positional models suggests that the original simplifying assumption to use a temporal dimension in serial recall tasks was not without merit. However, other data suggesting that relative time is not a key factor in immediate serial recall suggests further work focus on the positional version.

It is important to remember that SIMPLE predicts not only proportion correct, but also the position error gradients. Fig. 5 shows the predicted position error gradients from the positional version (compare with Fig. 2). The model is producing an appropriate amount of omissions in the various conditions, and, for the most part, an appropriate amount of movement errors. The model does not produce extra-list intrusions (but these are rare in the data also). The only apparent discrepancy is that the model does not produce quite as many movement errors for the older confusable condition as observed in the data.

**Simulation 3**

Both the temporal and positional versions of SIMPLE can account for the major results of interest, including both accuracy and error data, based primarily on the differing representations of the to-be-remembered items according to the calculated MDS solution. The final simulation concerns whether the solution can be applied to new sets of data. In particular, we wished to test the generality of the hypothesis that the relatively tighter clustering of phonologically confusable items in the psychological space of older subjects could be responsible for their impaired short-term memory performance. Fortuitously for this purpose, Maylor et al. (1999) examined serial recall by younger and older adults of 6-item lists of auditorily presented confusable and non-confusable letters. We therefore used the positional version of SIMPLE described above, derived a new set of MDS co-ordinates based on the Maylor et al. data to use as input to the model. These are shown in Fig. 6.

The estimated parameters for both groups were as follows: \(c = 3.81, o = 0.90, W_{D1} = 0.33, W_{D2} = 0.32, W_p = 0.35, s = 5.84, t = 0.46\). The \(\alpha\) parameter was 1.22 for the younger group, and 1.69 for the older group. Thus, as with the other simulations reported here, the only differences in the model between the groups were the calculated MDS solutions and the value of \(\alpha\). The data are shown in the left panel of Fig. 7 and the results of the simulation are shown in the right panel of Fig. 7.

The model is again clearly capturing the important aspects of the data. Note that the shape of the various serial position functions are quite different from those observed in our experiment, yet the model is able to capture the general pattern. Although there are six free parameters, they take the same value for all four conditions and are based on an MDS solution derived from a different set of subjects. Moreover, the model predicts not just the serial position functions shown in Fig. 6, but also the position error gradients (not shown). The difference in recall between older and younger adults and between recall of confusable and non-confusable items is accounted for in terms of the underlying dimensions on which the items are represented.

---

\(^2\)A general model, in which both a temporal and a positional dimension were included, gave some attentional weight to both of these dimensions.
Older adults performed worse on a task involving ordered recall of confusable and non-confusable items compared to younger adults. Unlike other studies, the present work demonstrated a relationship between peripheral sensory ability and recall in our older participants. Interestingly, hearing ability was correlated with recall of confusable but not non-confusable sets of items. This was explained by appealing to the idea that hearing loss has more of an effect on stimuli that are already quite confusable: Compression of the ‘memory space’ for the non-confusable items has less of an effect because the items start out in a more sparsely populated area of the memory space.

To test this idea we constructed confusion matrices made up of recall confusions and submitted them to a multidimensional scaling program. The memory space for older adults was compressed in relation to that of younger adults, and the space for confusable items was also compressed relative to non-confusable items.

The data were fit with SIMPLE, a scale invariant model of memory, which was able to account for the two main effects (list type and age) through differences in the coordinates from the MDS solution. That is, all parameters except for $\alpha$ were identical for both older and younger adults and for confusable and non-confusable items. The measured output times allowed a comparison of two extant versions of SIMPLE, one which uses a temporal dimension and one which uses a positional dimension. Both versions produced appropriate simulations, but in accordance with recent findings that immediate serial recall tasks are generally immune to temporal manipulations, further simulations assessed only the positional version. Finally, the generality of the model’s account was assessed by successfully fitting some previously reported data.

Because the key difference in the settings of the model for each of the conditions in all three simulations was the particular set of MDS coordinates used, it is possible—in principle—to model performance at the individual subject level, given sufficient observations. Thus, the model suggests a way of predicting memory performance from an individual’s underlying ‘memory space’ and removing chronological age, per se, as a variable. The present data did not include enough observations to make such an individual analysis feasible. However, given the promising results of the present study, future work will collect more observations per subject, but will space the experimental sessions over time to avoid inducing fatigue. In addition, future work should investigate the extent to which the dimensions used by younger and older adults correspond.

Obviously, SIMPLE is not the only model that addresses the acoustic confusion effect or age-related differences in memory. How does it compare to other models? In general, most models of the acoustic confusion effect have not been applied to the effects of aging and most models of the affects of aging have not been applied to the acoustic confusion effect (see Neath & Surprenant, submitted).

One notable exception is OSCAR (Brown, Preece, & Hulme, 2000). In OSCAR, an oscillator-based temporal context vector changes gradually over time, and is associated with the to-be-remembered items. Reinstantated states of this temporal context vector are used as recall probes. Maylor et al. (1999) fit OSCAR to data similar to that presented here, and found that it was necessary to change two parameters in OSCAR: older adults (according to the model) have worse context quality than younger adults and also have greater forgetting during output. The
acoustic confusion effect occurs primarily at retrieval: when a retrieval attempt is made, an approximation to the target item vector is retrieved and then compared with the potentially recallable items. Because confusable items are, by definition, more similar to each other than non-confusable items, there is more competition to confusable than to non-confusable items. A direct comparison is difficult, as the two models are intended to address different levels of analysis: SIMPLE has a higher level of abstraction and is omits mention of the process aspects of retrieval. In contrast, OSCAR has a lower level of abstraction and precisely specifies what happens at each stage. Neither model was built to account for age-related differences in memory, but both were able to account for such differences.

Although many explanations of age-related differences in memory rely on descriptive models, there is an increasing number of models that address either cognitive aging in general (e.g., Braver et al., 2001; Li et al., 2004; Myerson, Hale, Wagsta., Poon, & Smith, 1990) or age-related differences in memory in particular (e.g., Allen, Kaufman, Smith, & Propper, 1998; Balota, Duchek, & Paullin, 1989; Byrne, 1998; Maylor et al., 1999; Ratcli., Thapar, & McKoon, 2004). Although descriptive models may be well-specified, it is not always clear that a proposed explanation actually produces the observed pattern of results. In addition, verbal models leave little room for weighting different factors depending on the stimuli and type of task.

There are many advantages in using a formal memory model to help guide interpretation of the data (cf. Neath, 1999). It has become increasingly obvious that the effects of aging on memory are extremely complex and are caused by multiple interactions among factors. A formal model allows for the exploration of higher-order interactions that simply could not be worked through with a verbal model. In addition, clear and testable predictions can be made from a formal model in which psychologically plausible parameters are mapped on to particular human processes. Thus formal predictions can be made such that by manipulating a process (or combination of processes) a particular outcome will be observed. The current work suggests that researchers should consider how younger and older adults represent items and that differences in performance could be due to older adults having less distinct representations due to perceptual deficits.

References

Brown GDA, Morin C, Lewandowsky S. Evidence for time-based models of free recall. Psychonomic Bulletin and Review. in press

Neath I, Surprenant AM. Accounting for age-related differences in working memory using the Feature Model., submitted for publication. 2006


Rabbit PMA. Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. Acta Otolaryngology Supplement 1991;476:167–176.


Fig. 1.
Mean proportion of confusable (C) and non-confusable (N) letters correctly recalled in order by younger (Y) and older (O) subjects as a function of serial position (left panel) and corresponding mean output time in milliseconds (right panel). Error bars show the standard error of the mean.
Fig. 2.
Proportion of times each item was recalled in each position as a function of type of item (confusable or non-confusable) for younger and older subjects.
Fig. 3.
MDS solutions for younger (left panel) and older (right panel) subjects.
Fig. 4.
Fit of the temporal version of SIMPLE (left panel) and the positional version of SIMPLE (right panel) for the data shown in the left panel of Fig. 1.
Fig. 5.
Fit of the positional version of SIMPLE for the position error gradients. Compare to Fig. 2.
Fig. 6.
MDS solutions for younger (left panel) and older (right panel) subjects from Maylor et al. (1999).
Fig. 7.
Fit of the positional version of SIMPLE (right panel) for data reported by Maylor et al. (1999) (left panel).
### Table 1
Average pure-tone thresholds (dB HL), and their standard errors (in parentheses), for the younger and older subjects

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Younger</th>
<th>Older</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Right</td>
<td>Left</td>
</tr>
<tr>
<td>0.25 kHz</td>
<td>9.75 (.87)</td>
<td>10.00 (1.12)</td>
</tr>
<tr>
<td>0.50 kHz</td>
<td>6.25 (.82)</td>
<td>6.25 (.98)</td>
</tr>
<tr>
<td>1.00 kHz</td>
<td>5.50 (1.11)</td>
<td>5.25 (1.31)</td>
</tr>
<tr>
<td>2.00 kHz</td>
<td>5.00 (.91)</td>
<td>5.50 (1.34)</td>
</tr>
<tr>
<td>4.00 kHz</td>
<td>5.00 (1.29)</td>
<td>5.50 (2.13)</td>
</tr>
<tr>
<td>8.00 kHz</td>
<td>6.25 (1.70)</td>
<td>5.75 (1.41)</td>
</tr>
</tbody>
</table>
Table 2
Proportion of items correct, and proportion of error types as a function of group and list type

<table>
<thead>
<tr>
<th>Correct</th>
<th>Movement errors</th>
<th>Omissions</th>
<th>Extra-set intrusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-confusable younger</td>
<td>0.834</td>
<td>0.151</td>
<td>0.009</td>
</tr>
<tr>
<td>Confusable younger</td>
<td>0.515</td>
<td>0.372</td>
<td>0.017</td>
</tr>
<tr>
<td>Non-confusable older</td>
<td>0.660</td>
<td>0.276</td>
<td>0.037</td>
</tr>
<tr>
<td>Confusable older</td>
<td>0.342</td>
<td>0.443</td>
<td>0.100</td>
</tr>
</tbody>
</table>

*J Mem Lang.* Author manuscript; available in PMC 2008 January 2.