Reversing the phonological similarity effect

JAMES S. NAIRNE and MATTHEW R. KELLEY
Purdue University, West Lafayette, Indiana

The phonological similarity effect—poor retention of order for lists of similar-sounding items—is a benchmark finding in the short-term memory literature. In our first two experiments, we show that the effect actually reverses following relatively brief periods of distraction, yielding better order retention for similar than for dissimilar lists, provided that different items are used on every trial. In Experiment 3, the same items were used on every trial and similar lists produced poorer performance across all three retention intervals. The results are interpreted from a general discrimination framework: Items are viewed as occupying positions in a multidimensional space defined by list and within-list dimensions.

In immediate serial recall, lists composed of similar-sounding items are typically remembered less well than lists of dissimilar items (Baddeley, 1966; Conrad, 1964; Conrad & Hull, 1964; Hintzman, 1967). Increasing the phonological similarity among the members of a to-be-recalled set results in more omission errors at test, as well as more item-to-item transpositions in the output order (Coltheart, 1993; Conrad, 1965; Hintzman, 1965). The disruptive effects of phonological similarity are robust and have been replicated with a wide variety of experimental designs and materials, including the use of letters (Conrad, 1964), words (Baddeley, 1966), pictures (Hitch & Halliday, 1983), visual presentation (Baddeley, 1968), auditory presentation (Longoni, Richardson, & Aiello, 1993; Posner & Konick, 1966), and RSVP (Coltheart & Langdon, 1998). It is fair to say that the phonological similarity effect has achieved the status of a "benchmark" finding in the immediate memory literature, and most theories of short-term memory include mechanisms that are specifically designed to account for the phenomenon (e.g., Baddeley, 1986, 1992; Burgess & Hitch, 1992; Drewnowski, 1980; Nairne, 1990a; Schweickert, Guentert, & Hersberger, 1990; Shiffrin & Cook, 1978).

Nairne's (1990a) feature model, for example, is a member of a class of models that places the locus of the effect primarily at the point of retrieval, where the memory trace is interpreted or "deblurred." Confusions are thought to arise when degraded primary memory traces are compared with candidate members of a secondary memory search set. If list items sound similar (i.e., they share common vowel and/or consonant sounds), it is assumed that their mnemonic representations contain many overlapping features. Feature overlap, in turn, means that the representation of a given item in the sequence is apt to be confused with other list items, leading to incorrect trace interpretations at retrieval. Models that place the locus of such sim-

ilarity effects during the trace interpretation stage have been shown to mimic the empirical data quite well (Nairne, 1990a; see also Lewandowsky & Murdock, 1989).

The phonological similarity effect is well established in immediate serial recall, but it is less clear how the effect changes over the course of a retention interval. A number of studies have examined memory for confusable lists after distractor-filled retention intervals. Estes (1973) reported that sound-based confusion errors declined to near-chance levels after approximately 14 sec of digit shadowing (see also Conrad, 1967; Healy, 1975). Likewise, Baddeley and Ecob (1970) found that 20 sec of distractor activity (adding digit pairs) was sufficient to eliminate performance differences between similar and dissimilar lists. On the other hand, Baddeley (1968) and Posner and Konick (1966) found that the disruptive effect of similarity remained after 8 sec of digit copying (Baddeley, 1968) and 20 sec of adding digit pairs (Posner & Konick, 1966). Richardson, Longoni, and Di Masi (1996) reported that the phonological similarity effect was still intact after retention intervals ranging from 5 to 20 sec (see also Longoni et al., 1993). Apparently, then, the effect is robust enough to survive extended periods of distraction under some circumstances.

Recent pilot data from our laboratory complicate the matter even further. We examined whether the phonological similarity effect would persist through retention intervals (filled with digit shadowing) of 2, 8, or 24 sec. The results showed the standard phonological similarity effect after 2 sec of distraction and no effect after 8 sec of digit shadowing. However the effect actually reversed following the longest retention interval (24 sec); that is, similar lists yielded *better* overall order performance than did dissimilar lists after 24 sec of digit shadowing. This means that when testing is delayed by a distracting interval, phonological similarity can disrupt, facilitate, or have no effect on order retention, depending on the circumstance. Of course, these contrasting results could simply be the product of variations in experimental materials and/or design. The present experiments were designed, in part, to help gain further control over the situation.

Correspondence should be addressed to J. S. Nairne, Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907-1364 (e-mail: nairne@psych.purdue.edu).

Our pilot experiment differed from previous work in two important ways: (1) Different words were used on every trial, and (2) reconstruction of order was employed as the retention measure. Different words were used on every trial in an effort to minimize the effects of proactive interference (Keppel & Underwood, 1962; Nairne, Neath, & Serra, 1997). Repeating the same items across trials, or at least drawing from an extremely restricted set, is the norm for those interested in examining the immediate retention of order information, and the phonological similarity effect in particular (e.g., Richardson et al., 1996). By repeating items across trials, it is usually argued, one can reduce the demands of item encoding and obtain a more accurate measure of order retention. However, by consistently testing the same materials on each trial, one increases the probability that cross-trial confusions will occur (i.e., proactive interference) and potentially affect the results. Although earlier work by Coltheart (1993) verified that the phonological similarity effect still remains in immediate serial recall when different items are used on every trial, retention interval was not manipulated in her set of experiments.

We used reconstruction as the retention measure, rather than the more traditional serial recall, because serial recall has the problem of confounding item and order memory. Poor performance in serial recall can result from either the loss of item information (one forgets what items occurred) or the loss of order information (one forgets the position the item occupied on the list); there is no clear-cut way to disentangle the relative contributions of each in final performance. In a reconstruction task, the specific items that occurred on the just-presented trial are re-presented at the point of test, and the participant is required simply to place the items back into their original order of presentation. Because item information is completely available at the time of testing, reconstruction is often thought to be a better test of order or position memory (Healy, 1974; Nairne, 1990b, 1991). Reconstruction thereby accomplishes the same end as does repetition of items on every trial (reducing the role of item memory in performance) but without inducing a large buildup of proactive interference.

The experiments reported in the present article were designed to provide further data exploring how the effect of phonological similarity changes when testing is delayed by a distractor-filled retention interval. Experiments 1 and 2 replicated our pilot experiment, using different items on every trial and reconstruction as the retention measure; they revealed an enhancing effect of phonological similarity when testing was delayed by 24 sec of digit shadowing. In Experiment 3, in contrast, we employed the same items on every trial and obtained a quite different result—a consistently disruptive effect of phonological similarity across delays of 2, 8, and 24 sec. We will argue that these results are important because they elucidate the role that general discrimination processes play in immediate memory. More specifically, we will propose that items in immediate memory experiments are best viewed as occupying positions in a multi-dimensional memory space defined by list and within-list dimensions.

EXPERIMENT 1

In Experiment 1, participants were presented with 30 lists containing five unique items; no item was presented more than once in the experiment. Items were displayed visually and participants were instructed to read each item aloud as it was shown. Half of the lists contained items that were phonologically similar; half contained items that were phonologically distinct. The final item of each list in both conditions was immediately followed by a retention interval of 2, 8, or 24 sec filled with digit shadowing. During this retention interval, digits were displayed rapidly on the CRT, and the task was to repeat each digit aloud as it appeared. Immediately after the final digit, the five list items were presented again, in a new random order, and participants were required to place the words back into their original serial positions. Of main interest was the question of how memory for the similar and dissimilar lists might change over the course of the retention intervals.

Method

Participants and Apparatus. Participants were 42 Purdue University undergraduates who participated for credit in an introductory psychology course. Everyone was tested individually in sessions lasting approximately $\frac{1}{2}$ h. Stimuli were presented and controlled by IBM-compatible computers.

Materials and Design. Thirty lists of five rhyming nouns, of medium to high frequency, were constructed from sets reported by Libkuman (1994). The lists were then divided into two sets serving as the 15 similar and 15 dissimilar lists across participants. The dissimilar lists were formed by simply recombining words from similar lists in the following manner: For a given set, the 15 similar lists were divided into three groups of five lists. Within each group, five words, one from each list, were selected and recombined, creating a list of dissimilar-sounding items. By forming dissimilar lists in this fashion, we ensured that everyone would receive the same set of 150 words in the experiment. The words were simply grouped differently, serving in similar lists for one group of participants and in dissimilar lists for another. The order of the similar and dissimilar lists in each condition was randomly determined, as was the ordering of the words within each list. Two practice lists preceded the 30 experimental trials.

Three retention intervals (2, 8, and 24 sec) were employed in a within-subject design. The order of the retention intervals was determined randomly, and counterbalancing ensured that each of the three intervals, across participants, was associated equally often with each list. Participants had no way of knowing which of the three retention intervals would appear on a given trial.

Procedure. Each trial began with the word READY accompanied by a tone, followed by the presentation of a five-item list. Each list item was presented for 750 msec with a 250-msec interval separating the offset of one item from the onset of the next. Participants were instructed to say each item aloud as it appeared on the screen. Immediately after the last item, participants engaged in a digit-shadowing task that lasted for 2, 8, or 24 sec. The task required participants to read a number of randomly generated digits (0–9) aloud as each appeared individually in the center of the computer screen, at a rate of 500 msec per digit.

Following the last digit, the participant received a reconstruction test. The five list items were re-presented on the screen in a new random order. A series of five empty boxes, representing all possible list positions, was shown below the reordered list. The task was to reconstruct the original order of presentation by placing the items into their appropriate serial positions (boxes). Items were reordered in a two-step process. Step 1 required participants to select one of the words from the reordered list. Participants accomplished this by moving the cursor (via the arrow keys) to a word and pressing the space bar to select it. In Step 2, participants placed the word into its proper position by moving the cursor to one of the five boxes and pressing the space bar, thereby depositing the word. Participants were allowed to fill in the boxes in any order they wished and were instructed not to repeat or omit any items. Everyone was given as much time as needed to complete the reconstruction task; pressing the enter key initiated the next trial.

Results and Discussion

The data from the reconstruction test are shown in Figure 1 as a function of list type (similar vs. dissimilar), serial position (Positions 1–5), and length of the retention interval (2, 8, or 24 sec). The results are displayed as mean proportion correct reconstruction performance. An overall analysis of variance (ANOVA) was performed on the data, and each of its effects will be discussed in turn. First, the ANOVA revealed a highly reliable main effect of serial position [$F(4,164) = 60.78$, $MS_e = .05$, $p < .01$]. The serial position functions were characteris-

tically bow-shaped, with marked primacy and recency effects, for both similar and dissimilar lists at each of the three retention intervals. The interaction between retention interval and serial position was also significant [$F(8,328) = 2.18$, $MS_e = .03$, $p < .03$], reflecting somewhat reduced primacy and recency as the length of the retention interval increased. The presence of bow-shaped serial position curves, which flatten with an increasing retention interval, replicates previous work in immediate memory using reconstruction as a retention measure (e.g., Healy, 1974; Nairne, Riegler, & Serra, 1991).

Of main interest is the question of how phonological similarity affected memory for order over the course of the retention intervals. Although the main effect of list type was not reliable overall [$F < 1$], the ANOVA did show a reliable effect of retention interval [$F(2,82) = 8.24$, $MS_e = .11$, $p < .01$] as well as a reliable interaction between list type and retention interval [$F(2,82) = 6.28$, $MS_e = .09$, $p < .01$]. As Figure 1 indicates, there were sharp changes in the relative effect of phonological similarity over the course of the three retention intervals. After the shortest retention interval, participants showed the standard phonological similarity effect—better performance for lists composed of dissimilar-sounding items [$F(1,41) = 7.53$, $MS_e = 0.11$, $p < .01$]. After 8 sec of digit shadowing, list type had no significant effect on perfor-
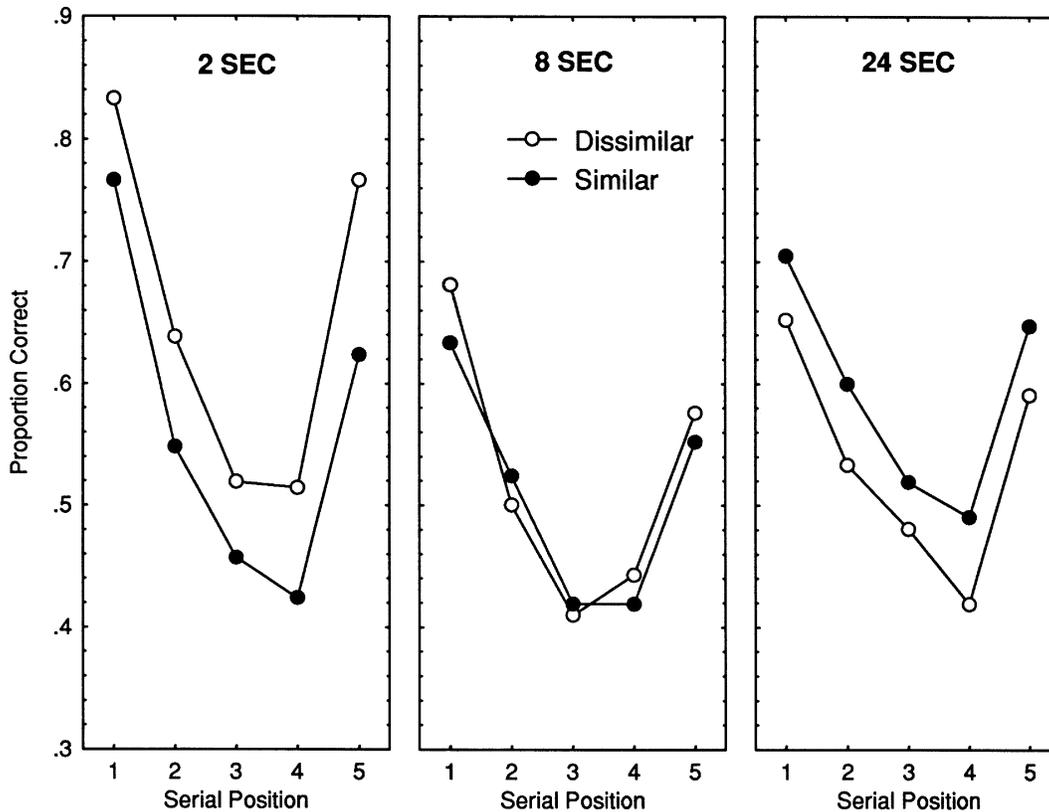

Figure 1. Serial position curves for the similar and dissimilar groups for each of the three retention intervals.

mance [$F(1,41) < 1$], and for the 24-sec retention interval, the phonological similarity effect was effectively reversed. The similar lists were reconstructed significantly better than the dissimilar lists [$F(1,41) = 5.11$, $MS_e = 0.07$, $p < .03$].

If we look only at the 2- and 8-sec retention intervals, our results replicate previous work by Estes (1973) and others in showing that the effects of phonological similarity disappear after a brief period of digit shadowing. The unique feature of the present results, of course, is the finding that after a relatively lengthy period of distraction (24 sec), similarity among list items actually improved order memory relative to a dissimilar condition. Interestingly, for the similar lists there appeared to be little, if any, decline in performance over the course of the retention intervals; in fact, average performance for the similar lists, but not for the dissimilar lists, actually improved significantly from the 8- to the 24-sec interval (as determined by a Newman–Keuls test).

Although surprising, this enhancing effect of phonological similarity after a relatively lengthy delay is not without precedent. A similar result was found in a long-term memory experiment by Nairne and Neumann (1993) when testing was delayed by a 10-min distracting interval. That experiment differed from the present one, however, in that testing was conducted after a number of lists had been presented rather than after each individual list. Nairne and Neumann argued that similarity can sometimes help participants discriminate current trial information from the residual effects of prior trials, thus improving relative reconstruction performance. We will offer a similar interpretation for the data of the present Experiment 1, but first we report the results of a second experiment that was designed to replicate Experiment 1 using a slightly different experimental design.

EXPERIMENT 2

Experiment 2 followed the procedures of Experiment 1 in all details except for the presentation order of the similar and dissimilar lists within the experimental session. In Experiment 1, the similar and dissimilar lists were randomly intermixed throughout the session. In the present experiment, the lists were presented in blocks corresponding to the type of list (15 similar then 15 dissimilar lists, or vice versa). Although there was no particular reason to expect blocking by list type to change the patterns found in Experiment 1, it seemed possible that grouping the similar lists together might enhance the salience of the similarity variable to the participants.

Method

Participants and Apparatus. Participants were 42 Purdue University undergraduates who participated for course credit. The stimuli were presented and controlled by IBM-compatible computers. The entire experiment lasted ½ h.

Materials and Design. The stimulus materials from Experiment 1 were used again in Experiment 2. In all respects, the present experiment was a direct replication of Experiment 1, except that presentation of the lists was blocked rather than intermixed. Participants received one of two sets of lists, each containing a block of 15 similar lists and a block of 15 dissimilar lists. In one set, the block of similar lists was presented first, followed by the dissimilar block; the opposite pattern was employed with the remaining set. The order of the lists within each block was randomly determined for each set. A new random ordering of the words within each list was created and was held constant for all participants.

Procedure. The procedure of Experiment 2 matched that of Experiment 1 in all details.

Results and Discussion

Figure 2 shows the data from the reconstruction test, broken down by list type, serial position, and retention interval. As in Experiment 1, an overall ANOVA revealed a reliable effect of serial position [$F(4,164) = 62.53$, $MS_e = .05$, $p < .01$] as well as a reliable interaction between serial position and retention interval [$F(8,328) = 2.20$, $MS_e = .03$, $p < .03$]. The individual serial position curves were generally bow-shaped and primacy and recency appeared to flatten somewhat as the retention interval increased. In addition, block order (similar followed by dissimilar vs. dissimilar followed by similar) produced no significant effects in the analysis.

For the effects of main interest, as in Experiment 1 the ANOVA revealed a reliable effect of retention interval [$F(2,82) = 14.73$, $MS_e = .10$, $p < .01$] and no main effect of list type [$F < 1$]; list type interacted significantly with retention interval [$F(2,82) = 5.73$, $MS_e = 10$, $p < .01$]. Following the 2-sec retention interval there was a significant phonological similarity effect [$F(1,41) = 7.21$, $MS_e = .07$, $p < .01$], but no differences were found between similar and dissimilar lists after the 8-sec delay ($F < 1$). Following the 24-sec interval, the similarity effect was reversed and a comparative advantage for the similar lists appeared [$F(1,41) = 7.13$, $MS_e = 0.07$, $p < .01$]. Once again, there was very little, if any, evidence of forgetting for the similar lists over the course of the retention intervals; unlike in Experiment 1, a Newman–Keuls test revealed no significant increase in performance for the similar lists from 8 to 24 sec.

As previously mentioned, the disappearance of the phonological similarity effect after a retention interval of about 8 sec is not surprising; a number of studies in the literature have reported comparable findings (Conrad, 1967; Estes, 1973; Healy, 1975). The reversal of the phonological similarity effect after the long retention interval (24 sec), however, requires explanation. As mentioned previously, Nairne and Neumann (1993) obtained a similar enhancing effect of phonological similarity in a long-term memory experiment when testing was delayed by 10 min, but testing in their case occurred after multiple lists had been presented. To interpret their results, Nairne and Neumann (1993; see also Nairne, 1991) suggested that one could view individual list items as occupying positions in a multidimensional memory space, defined by a within-list dimension and a list dimension. Phonological similarity, they argued, potentially affects discriminability along each of these dimensions in a dif-
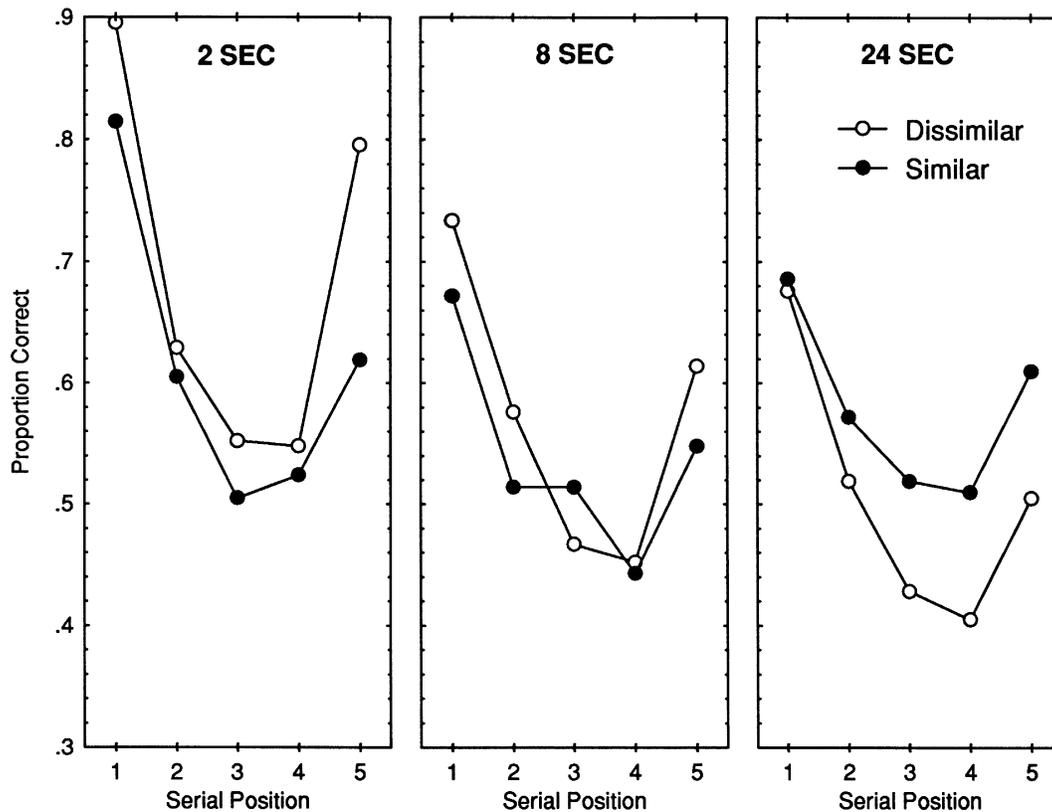
Figure 2. Serial position curves for the similar and dissimilar groups for each of the three retention intervals in Experiment 2.

ferent way. When list items sound similar, it is difficult to recover an item's correct position along the within-list dimension because the within-list traces contain a number of overlapping features (see the introduction; see also Nairne, 1990a). On the other hand, phonological similarity could well help one to discriminate an item's correct position along the list dimension—that is, to determine the correct *list* representation from other list representations in memory—because, again, all the list items share a common feature (such as a common vowel and/or consonant sound).

For example, suppose a participant is presented with a list such as *block stock clock lock rock*. In this case, he/she can presumably use the common phonological feature among the list items (in this case, the sound *-ock*) as a cue to discriminate the correct list representation from other representations in memory (e.g., from lists that ended in the sound *-are* or *-ere*). With dissimilar items, there is no distinctive list-based phonological cue that clearly distinguishes one list from another in the memory space. Nairne and Neumann (1993) reasoned that manipulations of within-list similarity will either enhance or impair performance depending on the respective roles that list discriminability and within-list discriminability

play in final task performance. Similarity will enhance performance whenever the disadvantage that similarity gives to within-list discrimination is more than offset by the advantage that similarity gives to discrimination along the list dimension.

These arguments can be extended to explain the results of our present experiments. In Experiments 1 and 2, the retention intervals were much shorter than in the Nairne and Neumann (1993) study, ranging only from 2 to 24 sec. Over these shorter intervals, it is possible to see how the effects of the discrimination processes on memory performance might change over time. It is reasonable to argue that when testing occurs immediately after presentation, performance will be driven primarily by locating the item's position along just the within-list dimension; the list representation itself is temporally distinct, having just occurred in time, and should therefore be relatively easy to access. However, as the length of the retention interval increases, the list representation will recede backward in time, making it more difficult to discriminate current trial information from that of prior trials— much like railroad ties seem to come together perceptually with increasing distance. After the 24-sec retention interval, the similarity-based list cue (e.g., the trial contain-

ing words that end in -ock) is therefore likely to play a relatively more important role and performance will be better than in the dissimilar condition.

If the reversal of the phonological similarity effect is a result of improved discriminability along the list dimension, as suggested by Nairne and Neumann (1993), then we can generate a relatively straightforward prediction. If the same set of items were to be repeated across trials, rather than different items on every trial, as in Experiments 1 and 2, we should expect to see the standard, disruptive, phonological similarity effect, regardless of the length of the retention interval. When items are repeated across trials, the list cue will be rendered ineffective, no longer aiding in list discrimination for similar-sounding lists. The cue will no longer uniquely specify a particular list in memory; it will specify a number of lists. For example, if the list *block lock stock clock rock* has been seen 10 times throughout the experimental session, the list cue (all items ending with the -ock sound) will not help in discriminating the current presentation from prior presentations of the same items. In such a case, list access would presumably be equivalent for both similar and dissimilar lists. One would expect performance then to be driven by the within-list discrimination process, which should result in a dissimilar-list advantage or a phonological similarity effect. Experiment 3 was designed to test this prediction.

EXPERIMENT 3

The presentation, distraction, and test procedures of Experiment 3 were identical to those of the previous experiments, except for a change in the number and type of stimuli used. In the previous experiments, each similar and dissimilar list consisted of novel items not previously seen in the experiment. In the present experiment, participants received only two lists, one with five phonologically similar items and one with five phonologically distinct items, and these same lists were repeated throughout the session.

Method

Participants and Apparatus. The participants were 40 undergraduates participating for course credit. Sessions were conducted individually and lasted for 1 h. Stimuli were presented and controlled by IBM-compatible computers.

Materials and Design. The stimulus materials consisted of 10 similar and 10 dissimilar lists. The 10 similar lists were selected, at random, from the materials used in Experiments 1 and 2. The 10 dissimilar lists were formed by recombining the words from the similar lists in the following manner. The 10 similar lists were divided into two groups of five lists. Within each group, five words, one from each list, were selected and recombined, creating a list of dissimilar-sounding items. As a consequence of forming the dissimilar lists in this fashion, each word served in both a similar and a dissimilar list, although never in the same experimental session.

Ten sets of lists were constructed, each set containing one similar and one dissimilar list. A set consisted of 48 lists (24 similar and

24 dissimilar) and was formed by repeating the two lists throughout the session. The order of the items within each list was randomly determined for each trial. A participant received 1 of these 10 sets of lists. As in Experiment 1, similar and dissimilar lists were randomly intermixed throughout the session. All remaining details were identical to those in the previous experiments.

Procedure. Except for the change in the number of trials, the procedures of Experiment 3 matched those of the previous experiments in all details.

Results and Discussion

The proportion of correct responses in the reconstruction task is shown in Figure 3 as a function of serial position, list type, and retention interval. The overall ANOVA revealed a reliable effect of serial position [$F(4,156) = 37.97$, $MS_e = .05$, $p < .01$] and a reliable retention interval $\times$ serial position interaction [$F(8,312) = 3.11$, $MS_e = .02$, $p < .01$]. As in the previous experiments, the serial position curves exhibited a bow-shaped form, with pronounced primacy and recency effects, and flattened as the length of the retention interval increased.

As Figure 3 shows, the pattern found in Experiment 3 differed significantly from the one found in Experiments 1 and 2. The effect of phonological similarity did not change over the course of the three retention intervals; the standard phonological similarity effect persisted through each of the three intervals. The main effect of list type was reliable in this case [$F(1,39) = 9.89$, $MS_e = .08$, $p < .01$], as was the effect of retention interval [$F(2,78) = 17.63$, $MS_e = .09$, $p < .01$]. All remaining interactions, including the critical list type $\times$ retention interval interaction, were not statistically significant. An analysis of the individual comparisons revealed a significant phonological similarity effect after 2 sec [$F(1,39) = 7.92$, $MS_e = .06$, $p < .01$]; after 8 sec, the phonological similarity effect did not quite reach conventional levels of significance [$F(1,39) = 2.60$, $MS_e = .05$, $p < .11$], and the effect was again reliable after the longest interval of 24 sec [$F(1,39) = 4.66$, $MS_e = .06$, $p < .04$].

We predicted that if the same items were repeated across trials, the enhancing effect of phonological similarity after the 24-sec retention interval would disappear and, instead, we would see the standard phonological similarity effect. This prediction followed from our proposal that the enhancing effect of similarity stems from an advantage that is sometimes seen during the list-discrimination process. When list items are phonologically similar, they share a common feature that can be used as a cue in discriminating the correct list representation from other list representations in memory. However, this list cue will be effective only when it uniquely specifies a particular list in memory. If the same items are repeated across trials, the list cue will no longer effectively specify one particular list. In such a case, performance should be guided primarily by the within-list discrimination process, and one would expect a disruptive effect of phonological similarity after each retention interval. The results
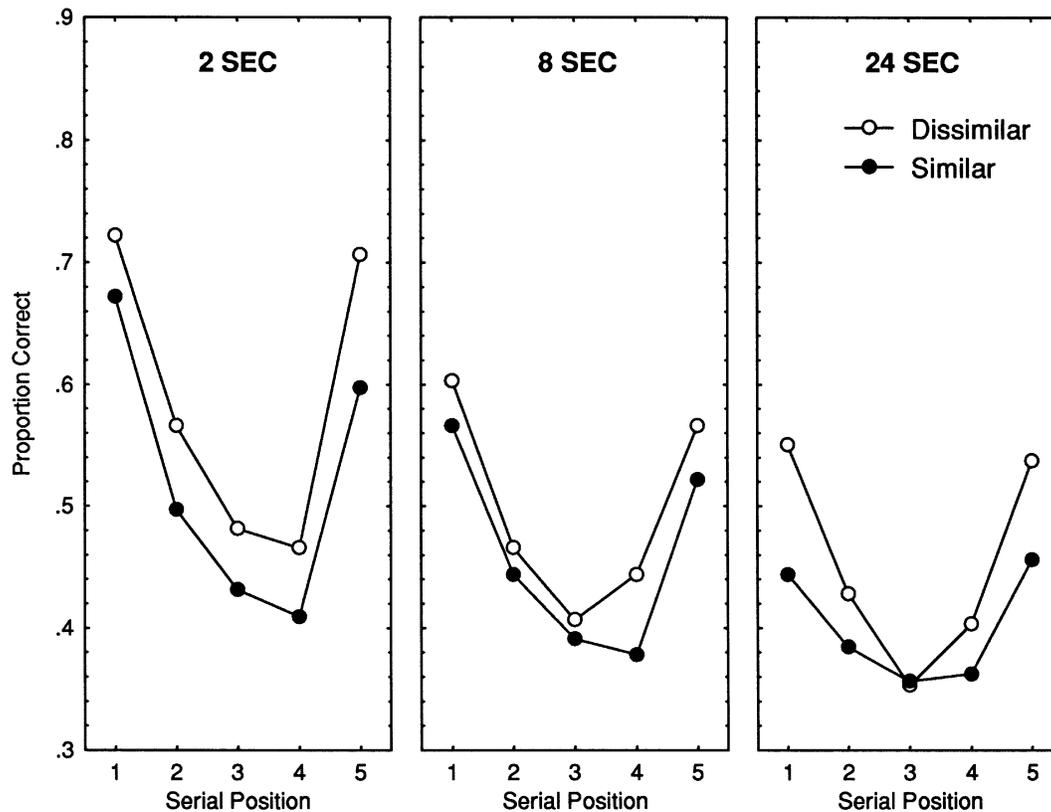
Figure 3. Serial position curves for the similar and dissimilar groups for each of the three retention intervals in Experiment 3.

of Experiment 3 generally support this prediction. An advantage for dissimilar lists over similar lists was seen at each of the three intervals.

GENERAL DISCUSSION

The present experiments were conducted in an effort to examine how the effect of phonological similarity changes over the course of a retention interval. Previous investigations have produced somewhat mixed results: Some studies have reported that the phonological similarity effect disappears when testing is delayed by a distractor-filled retention interval (Baddeley & Ecob, 1970; Conrad, 1967; Estes, 1973; Healy, 1975); others have reported that the effect persists through retention intervals of up to 20 sec (Posner & Konick, 1966; Richardson et al., 1996). The present experiments have demonstrated yet another result—namely, a reversal of the phonological similarity effect after a filled retention interval of 24 sec. In Experiments 1 and 2, we used different items on every trial and used reconstruction as the retention measure; we found the standard phonological similarity effect after the short retention interval of 2 sec, no effect of similarity after the 8-sec interval, and a reversal of the phonological similarity effect after the longest retention

interval of 24 sec. Experiment 3, in contrast, employed the same items across trials and obtained a disruptive effect of phonological similarity across delays of 2, 8, and 24 sec. Because these are cross-experiment comparisons, we conducted an additional ANOVA to look for the presence of an experiment (1 vs. 3) × similarity (similar vs. dissimilar) × delay (2, 8, and 24 sec) interaction; the interaction was significant [$F(2,160) = 3.51$, $MS_e = .068$, $p < .032$].

To explain these results, we proposed an account that appealed to differential discriminability along a within-list and a list dimension in a multidimensional memory space. We argued that phonological similarity affects discriminability along each of these two dimensions in a different way. Phonological similarity consistently disrupts one's ability to discriminate an item's correct *within-list* position, while at the same time it enhances one's ability to discriminate a correct *list* representation from other list representations in memory. Manipulations of within-list similarity, as a result, will either enhance or impair performance depending on the respective roles that list and within-list discriminability play in task performance. We argued that the relative contributions of these discrimination processes can change over the course of a retention interval.

When testing occurs immediately, or after a brief delay (2 sec), it is likely that performance is driven primarily by recovering each item's correct within-list position. This should lead to the standard phonological similarity effect because recovering an item's correct position within a list is more difficult when list items sound similar. As the length of the retention interval increases (24 sec), the list dimension is assumed to play a more important role in determining final performance. Because phonological similarity can, in principle, aid in list discrimination (by providing a list-based retrieval cue), it should be possible for similar lists to yield better overall order performance relative to dissimilar lists—a reversal of the phonological similarity effect—under some circumstances (e.g., Experiments 1 and 2).

At the same time, the list-based similarity cue should be effective only when it uniquely specifies a particular list in memory. When the same items are repeated across trials, item-based list information (e.g., that all items end in *-ock*) will no longer specify any particular list representation; instead, this cue information will be consistent with every list that occurred in the experiment. Under these conditions, we would expect correct performance to be guided primarily by the within-list dimension, and the standard phonological similarity effect should emerge regardless of the length of the retention interval. As we showed in Experiment 3, when the same items are repeated across trials, there is no longer any reversal of the phonological similarity effect after a long delay; instead, dissimilar lists show an ordering advantage across all retention intervals (although the effect was not statistically significant at the 8-sec interval).

A similar list-based discriminability hypothesis was used by Nairne (1990b) to explain how semantic similarity can enhance long-term order performance when testing was done on same-category materials. Using an incidental learning paradigm with a 10-min distracting interval, Nairne (1990b) showed that order performance was better when the items within a list were all drawn from the same semantic category than when list items were dissimilar. It was argued that category information could provide a powerful retrieval cue for accessing the appropriate list processing record. However, the similarity advantage occurred only when each list was drawn from a different semantic category. When all lists items were drawn from the same category, even though the individual items were different on every trial, the enhancing effect of similarity disappeared. Similarity enhanced order performance only when the list-based cue uniquely specified a particular list in memory, as in the present experiments.

There is one aspect of the present account that remains perplexing, however. It could be argued that since the list items were re-presented at test (reconstruction), list discrimination should not play a role in final performance; one need not remember the appropriate list when the list itself is given at test. How then can a list-based "cue" possibly lead to an enhancement in performance when the information provided by the cue is fully available to the remember? It should be noted, however, that an important part of any act of remembering is the recovery of the original processing record (Hunt & Einstein, 1981). Importantly, even when the list items are re-presented at test, participants must still access the appropriate representation in memory. The items themselves simply serve as cues, albeit "copy cues," and do not guarantee access to the proper mnemonic representation (Tulving, 1983). This means that reconstruction should not be considered as a "pure" test of order, although it is a nominal test of order, because item information clearly plays a role (Neath, 1997).

Overall, the fact that phonological similarity, depending on experimental condition, can either impair or enhance reconstruction performance presents a significant challenge to most current models of short-term memory, especially the working memory framework proposed by Baddeley (1986, 1990). All current versions of working memory assume that the locus of the phonological similarity effect lies in the phonological store, a storage facility that is a component of a more general phonological loop. It is assumed that people can hold a limited amount of verbal material, in an abstract phonological form, for a brief amount of time in this store. The phonological store's contents are thought to decay within 1–2 sec (Baddeley, 1986), although Richardson et al. (1996) have recently shown that phonological similarity, as in the present experiments, can exert a negative effect after filled intervals of up to 20 sec; thus, in order to salvage the working memory account, one must assume that the contents of the phonological store can potentially have a much longer duration (10–20 sec) than was previously thought.

In this case, however, one should still expect to see a *disruptive* effect of phonological similarity as long as testing occurs within the 10- to 20-sec decay window. As we have shown here, this prediction does not always hold true. In Experiments 1 and 2, the phonological similarity effect disappeared after an 8-sec retention interval and reversed following the 24-sec interval. One could, of course, argue that performance after 24 sec of distraction is no longer mediated by the working memory system. It is conceivable that after a lengthy filled delay, performance is governed by some form of long-term memory. The enhancing effect of similarity might then result from improved discriminability in this system rather than in working memory. Yet in Experiment 3, when the same items were used on every trial, phonological similarity was shown to impair performance throughout an identical 24-sec filled interval. What matters is not the passage of time per se, but rather whether similarity can act as an additional, and beneficial, list cue to offset the detrimental effect that it normally has on within-list discriminability.

We believe that our data are more adequately handled by cue-based theories of immediate memory (e.g., Nairne, 1990a). Thus, items are not proposed to reside in stores, nor are they recovered "intact" from consciousness provided that they escape the effects of decay. Instead, the short-term memory system is seen as a repository for

cues, and all immediate memory performance results from interpretation of these cues (Nairne, 1990a; Tehan & Humphreys, 1995). People in immediate memory experiments are trying to solve a discrimination problem: What occurred on the just-presented list? To solve this discrimination problem, they use whatever information is available at the point of test. Similarity, as we have seen, is an important source of cue information that can either help or hinder the discrimination process depending on the circumstance.

REFERENCES

BADDELEY, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology*, **18**, 302-309.

BADDELEY, A. D. (1968). How does acoustic similarity influence short-term memory? *Quarterly Journal of Experimental Psychology*, **20**, 249-264.

BADDELEY, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

BADDELEY, A. D. (1990). *Human memory: Theory and practice*. Oxford: Oxford University Press.

BADDELEY, A. D. (1992). Working memory. *Science*, **255**, 556-559.

BADDELEY, A. D., & ECOB, J. R. (1970). Simultaneous acoustic and semantic coding in short-term memory. *Nature*, **227**, 288-289.

BURGESS, N., & HITCH, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory & Language*, **31**, 429-460.

COLTHEART, V. (1993). Effects of phonological similarity and concurrent irrelevant articulation on short-term memory recall of repeated and novel word lists. *Memory & Cognition*, **21**, 539-545.

COLTHEART, V., & LANGDON, R. (1998). Recall of short word lists presented visually at fast rates: Effects of phonological similarity and word length. *Memory & Cognition*, **26**, 330-342.

CONRAD, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, **55**, 75-84.

CONRAD, R. (1965). Order errors in immediate recall of sequences. *Journal of Verbal Learning & Verbal Behavior*, **4**, 161-169.

CONRAD, R. (1967). Interference or decay over short retention intervals? *Journal of Verbal Learning & Verbal Behavior*, **6**, 49-54.

CONRAD, R., & HULL, A. J. (1964). Information, acoustic confusion, and memory span. *British Journal of Psychology*, **55**, 428-432.

DREWNOWSKI, A. (1980). Memory functions for vowels and consonants: A reinterpretation of acoustic similarity effects. *Journal of Verbal Learning & Verbal Behavior*, **19**, 176-193.

ESTES, W. K. (1973). Phonemic coding and rehearsal in short-term memory for letter strings. *Journal of Verbal Learning & Verbal Behavior*, **12**, 360-372.

HEALY, A. F. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **13**, 644-655.

HEALY, A. F. (1975). Coding of temporal-spatial patterns in short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **14**, 481-495.

HINTZMAN, D. L. (1965). Classification and aural coding in short-term memory. *Psychonomic Science*, **3**, 161-162.

HINTZMAN, D. L. (1967). Articulatory coding in short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **6**, 312-316.

HITCH, G. J., & HALLIDAY, M. S. (1983). Working memory in children. *Philosophical Transactions of the Royal Society of London: Series B*, **302**, 325-340.

HUNT, R. R., & EINSTEIN, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning & Verbal Behavior*, **20**, 497-514.

KEPPEL, G., & UNDERWOOD, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning & Verbal Behavior*, **1**, 153-161.

LEWANDOWSKY, S., & MURDOCK, B. B., JR. (1989). Memory for serial order. *Psychological Review*, **96**, 25-57.

LIBKUMAN, T. M. (1994). Norms for words that rhyme. *Behavior Research Methods, Instruments, & Computers*, **26**, 278-322.

LONGONI, A. M., RICHARDSON, J. T. E., & AIELLO, A. (1993). Articulatory rehearsal and phonological storage in working memory. *Memory & Cognition*, **21**, 11-22.

NAIRNE, J. S. (1990a). A feature model of immediate memory. *Memory & Cognition*, **18**, 251-269.

NAIRNE, J. S. (1990b). Similarity and long-term memory for order. *Journal of Memory & Language*, **29**, 733-746.

NAIRNE, J. S. (1991). Positional uncertainty in long-term memory. *Memory & Cognition*, **19**, 332-340.

NAIRNE, J. S., NEATH, I., & SERRA, M. (1997). Proactive interference plays a role in the word-length effect. *Psychonomic Bulletin & Review*, **4**, 541-545.

NAIRNE, J. S., & NEUMANN, C. (1993). Enhancing effects of similarity on long-term memory for order. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 329-337.

NAIRNE, J. S., RIEGLER, G. L., & SERRA, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 702-709.

NEATH, I. (1997). Modality, concreteness, and set-size effects in a free reconstruction of order task. *Memory & Cognition*, **25**, 256-263.

POSNER, M. I., & KONICK, A. F. (1966). On the role of interference in short-term retention. *Journal of Experimental Psychology*, **72**, 221-231.

RICHARDSON, J. T. E., LONGONI, A. M., & DI MASI, N. (1996). Persistence of the phonological trace in working memory. *Cahiers de Psychologie Cognitive*, **15**, 557-581.

SCHWEICKERT, R., GUENTERT, L., & HERSBERGER, L. (1990). Phonological similarity, pronunciation rate, and memory span. *Psychological Science*, **1**, 74-77.

SHIFFRIN, R. M., & COOK, J. R. (1978). Short-term forgetting of item and order information. *Journal of Verbal Learning & Verbal Behavior*, **17**, 189-218.

TEHAN, G., & HUMPHREYS, M. S. (1995). Transient phonemic codes and immunity to proactive interference. *Memory & Cognition*, **23**, 181-191.

TULVING, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.