Journal of Experimental Psychology:
Learning, Memory, and Cognition
2003, Vol. 29, No. 5, 838–849

0278-7393/03/$12.00 DOI: 10.1037/0278-7393.29.5.838

Dissimilar Items Benefit From Phonological Similarity in Serial Recall

Simon Farrell
Northwestern University

Stephan Lewandowsky
University of Western Australia

In short-term serial recall, similar sounding items are remembered less well than items that do not sound alike. This phonological similarity effect has been observed with lists composed only of similar items, and also with lists that mix together similar and dissimilar items. An additional consistent finding has been what the authors call *dissimilar immunity*, the finding that ordered recall of dissimilar items is the same whether these items occur in pure dissimilar or mixed lists. The authors present 3 experiments that disconfirm these previous findings by showing that dissimilar items on mixed lists are recalled better than their counterparts on pure lists if order errors are considered separately from intrusion errors (Experiment 1), or if intrusion errors are experimentally controlled (Experiments 2 and 3). The memory benefit for dissimilar items on mixed lists poses a challenge for current models of short-term serial recall.

The standard phonological similarity effect in serial recall refers to the well-replicated finding that lists composed of similar sounding items are recalled less accurately than lists in which items do not sound alike (e.g., Baddeley, 1966, 1968; Conrad, 1964; Henson, Norris, Page, & Baddeley, 1996; Wickelgren, 1965a, 1965b). This effect is of considerable generality, occurring with lists of letters (Baddeley, 1968; Conrad & Hull, 1964) and word lists (Baddeley, 1966; Coltheart, 1993; Henry, 1991), and it has had unparalleled theoretical impact and is considered to be a crucial feature of short-term serial recall (Farrell, 2001).

The phonological similarity effect also arises when phonologically similar (e.g., *B, P, T*) and dissimilar (e.g., *K, Q, R*) items are presented together in a mixed list. Mixed lists also give rise to an additional, highly diagnostic finding: The accuracy of serial recall of dissimilar items in mixed lists is not affected by the presence of similar items. That is, the recall of dissimilar items in mixed lists is the same as that of dissimilar items at the same serial positions on pure dissimilar lists (e.g., Baddeley, 1968; Bjork & Healy, 1974; Henson et al., 1996). This immunity of dissimilar items to the list environment has led numerous theorists to propose that serial recall involves two independent stages of processing, with order errors occurring between abstract positional representations in a primary stage, and with item-based similarity confusions occurring in a separate secondary stage (e.g., Burgess & Hitch, 1999; Henson, 1998; Page & Norris, 1998a, 1998b).

This article critically reexamines the mixed list effect and reports three experiments that show that, contrary to previous reports, phonological similarity enhances memory for the order of dissimilar items on the same list. Previous studies did not detect this effect because they did not examine or control the distribution of different types of recall errors, in particular a trade-off between transposition and intrusion errors. When intrusion errors are left uncontrolled, the enhanced memory for dissimilar items is expressed in the transposition rates (Experiment 1). When intrusions are eliminated by using a reconstruction task (Experiment 2) or by controlling guessing strategies (Experiment 3), serial recall of dissimilar items on mixed lists is boosted relative to pure-list controls. These results undermine the empirical motivation for two-stage models of serial recall (e.g., Burgess & Hitch, 1999; Henson, 1998; Page & Norris, 1998a, 1998b) and are more compatible with approaches that expect similarity to have an effect at encoding (e.g., Brown & Chater, 2001; Brown, Neath, & Chater, 2002; Farrell & Lewandowsky, 2002; Nairne, 1990).

Mixed-List Phonological Similarity in Serial Recall

Baddeley (1968, Experiment 5) presented participants with lists in which dissimilar items were interleaved with similar items. For those lists, similar items were similar to each other but not to the dissimilar items, which in turn were dissimilar to each other. For example, a mixed list might contain the letters *M, V, K, D, R,* and *T,* in which *V, D,* and *T* are phonologically similar, and *M, K,* and *R* are each phonologically distinct from all other list items. An abstract representation of this mixed list is *DSDSDS,* where *S* and *D* represent, respectively, a similar and dissimilar item. Previous research on mixed lists has used a variety of list types, typically including the cases *SSSSSS* (pure similar), *DDDDDD* (pure dissimilar), and the alternating cases *DSDSDS* and *SDSDSD* (e.g., Baddeley, 1968; Henson et al., 1996; but see Bjork & Healy, 1974).

An early theoretical impetus for mixed-list research was the chaining view of serial recall, which posits that lists are repre-

---

Simon Farrell, Department of Psychology, Northwestern University; Stephan Lewandowsky, Department of Psychology, University of Western Australia, Nedlands, Australia.

Preparation of this article was supported by Grant A00105535 from the Australian Research Council to Stephan Lewandowsky. Simon Farrell was supported in part by National Institute of Mental Health Grant HD MH44640 and National Institute on Aging Grant AG17083-01.

We thank Leo Roberts for his help at all stages of manuscript preparation. We also thank Gordon Brown, Mathew Duncan, and Tony Fallon for their comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Simon Farrell, Department of Psychology, Swift Hall 102, Northwestern University, 2029 Sheridan Road, Evanston, Illinois 60208. E-mail: s-farrell@northwestern.edu

sented as a chain of items, with each pair of items being joined by an association that is used at retrieval to cue recall of the next item (e.g., Lewandowsky & Murdock, 1989; Wickelgren, 1965b). Computational instantiations of chaining (e.g., theory of distributed associative memory [TODAM]; Lewandowsky & Murdock, 1989) furthermore specify that cues similar to the one initially associated with an item will also be effective retrieval prompts. By implication, cueing with any of the similar items on mixed lists should lead to interference on the following dissimilar item because, the cue being similar to others, several dissimilar candidates are retrieved and compete for report. In consequence, recall of dissimilar items on mixed lists should be worse than on pure lists. Baddeley, Papagno, and Norris (1991) presented simulations of TODAM that provide a quantitative illustration of these predictions.

This prediction by chaining models was first called into question by Baddeley's (1968) study, which found that the accuracy with which *D* items were recalled from mixed lists (e.g., *SDSDSD*) was identical to that of *D* items in the corresponding serial positions from pure dissimilar lists (*DDDDDD*). The effect was replicated by Henson et al. (1996), who used a more refined conditional measure that presented an even greater problem for chaining models.

The immunity of dissimilar items to the list environment is now considered a benchmark result in short-term serial recall, and has posed a considerable challenge to theorists. In order to handle the mixed-list dissimilar immunity effect, several models incorporate the assumption that phonological similarity affects a separate output stage that is not involved in representing serial order.

Theoretical Accounts of Dissimilar Immunity

The primacy model (Page & Norris, 1998a, 1998b) accounts for mixed-list dissimilar immunity through addition of an extra "confusion" stage that affects only phonologically similar items. Thus, the first stage of the model is thought to store the order of items without regard to their phonological similarity. At recall, the output of the first stage is passed to the second stage, where phonological confusions are assumed to take place. Critically, these confusions only occur between similar items, so that items dissimilar to other list items will pass through this second stage unaffected. Henson's (1998) start–end model (SEM) accounts for mixed-list effects in a similar manner, by assuming a separate stage of phonological confusions after normal retrieval operation of the model (see Henson, 1998, for details).

The Burgess and Hitch (1999) model also accounts for the mixed-list dissimilar immunity effect through multiple stages of item selection and output. In their model, item selection, which is driven by a representation of temporal context, is unaffected by phonological similarity. (This stage loosely corresponds to the first stage in the primary model and SEM.) Similarity effects instead arise because of phonemic feedback. The model contains two layers of phonemes, one at input and one at output, that are mutually interconnected. When a candidate item selected in the first stage cues the output phonemes, feedback is unavoidably sent to the input phonemic layer, which in turn "increases the likelihood of a (nearby) similar item being recalled in its place" (Burgess & Hitch, 1999, p. 569).

In summary, to accommodate the mixed-list dissimilar immunity effect, these three models postulate two (or more) stages of processing, with order errors occurring between positional tokens at an early stage, and with item-based similarity confusions occurring in a separate, later stage. Furthermore, order representations involved in the primary stage are assumed to be unaffected by phonological confusability. Notably, Page and Norris (1998b) reported that explorations of several single-stage models to handle the mixed-list results were unsuccessful, and concluded that "the data alone appear to force us to accept a two-stage model" (p. 243).[1] The dissimilar immunity effect has therefore played a major role in motivating the architecture of contemporary models of serial recall.

It follows that any empirical disconfirmation of the dissimilar immunity effect would have considerable theoretical significance. Specifically, a demonstration that dissimilar items are not always immune to the similarity of surrounding list items would call into question the basis for dual-stage models, in particular the assumption that representation of order is unaffected by phonological similarity. Although the studies to date give an apparently clear empirical picture, there are several reasons why a further exploration of the dissimilar immunity effect is warranted. First, data analysis in previous studies focused primarily on accuracy (e.g., Baddeley, 1968; Henson et al., 1996), perhaps at the expense of a more detailed analysis of errors and a possible trade-off between different types of errors. For example, although Henson et al. (1996) reported transposition gradients for one of their experiments, they did not break down intrusion and omission rates between different list types, although errors together accounted for up to 23% of all responses (Henson et al., 1996, Experiment 2).

Second, the immunity of dissimilar items may be tied to the relatively weak similarity manipulations used in research to date. For example, the mixed lists used by Henson et al. (1996) contained an equal number of similar and dissimilar items, which may have limited the potential distinctiveness of dissimilar items (Newman & Jennette, 1975). It is possible that the immunity of dissimilar items could be shattered by a stronger manipulation of similarity. For example, only a single dissimilar item might be embedded in a mixed list, as in research on the isolation effect.

The Isolation Effect

The isolation effect, often called the *von Restorff effect* after its initial investigator, refers to the ubiquitous finding that free recall of an item is facilitated if it is dissimilar from a homogeneous set of surrounding list items. For example, the word *tiger* (referred to as the *isolate*) will be remembered better if it is presented in a list of vegetables than if it is presented in a list of mutually unrelated items (for a review see Hunt, 1995; Wallace, 1965). Translated to the present context, the isolation effect implies that a single dissimilar item on a list of similar items should be better recalled than an item in the same serial position on a pure dissimilar list. Thus,

---

[1] The need for two stages is indirectly supported by the oscillator-based associative recall (OSCAR) model (Brown, Preece, & Hulme, 2000). In critical differentiation with the other three theories, OSCAR assumes that similarity plays a role not only during final response selection but also at an earlier stage that is devoted to maintaining order among list items. OSCAR has not been applied to the mixed-list effect, but Brown et al. (2000) acknowledged that the presence of similarity effects at an early stage may be problematic in light of the dissimilar immunity in mixed lists.

the isolation effect lies in empirical opposition to the dissimilar immunity effect.[2]

Research on the isolation effect has typically involved free recall (e.g., Dunlosky, Hunt, & Clark, 2000; Fabiani & Donchin, 1995; Hirshman & Jackson, 1997; Winters & Hoats, 1989), although a number of studies have addressed order memory (e.g., Bone & Goulet, 1968; Cimbalo, Nowak, & Soderstrom, 1981; Lippman, 1980). Lippman and Lippman (1978) found enhanced reordering of an item isolated by color with a reconstruction task. Similar effects were reported by Kelley and Nairne (2001), who also used a reconstruction task. In a reconstruction task, participants are provided with the identity of all list items at recall, and the task is to rearrange these items into the order in which they were presented at study. This task is considered a relatively pure measure of order memory because people are not required to remember the identity of items, only their positions.

These relevant precedents call for a further examination of possible phonological isolation effects in serial recall. Demonstrating the existence of a phonological isolation effect would contradict the existing empirical evidence motivating dual stages of processing, and would question the insensitivity of order encoding to phonological similarity. Instead, an isolation effect would support alternative views that acknowledge a role of similarity at encoding.

Similarity at Encoding

There are at least three memory models that ascribe a role to similarity at encoding. These models consequently reject the idea that similarity affects only a separate output stage.

Farrell and Lewandowsky (2002; see also Farrell, 2001) recently presented a model called serial-order-in-a-box (SOB), named after the brain-state-in-a-box (BSB) algorithm (e.g., Anderson, Silverstein, Ritz, & Jones, 1977) on which it is based. In SOB, each incoming list item is compared to the contents of memory (i.e., the composite matrix of connection weights that represents all previous items) and the strength of encoding of the item is based on the results of this comparison. If items are encoded with a strength that is inversely related to their similarity to the memory contents (Farrell, 2001), and an incoming item is similar to one previously presented, its encoding strength will be reduced compared to the case in which all preceding items are dissimilar. Because retrieval is based on the relative strengths of the items that remain to be recalled, the reduced encoding of similar items renders dissimilar items on mixed lists more retrievable than their counterparts on pure lists. The similarity-sensitive encoding in the SOB model (Farrell, 2001; Farrell & Lewandowsky, 2002) therefore qualitatively predicts an isolation effect.

Similar predictions can be derived from the scale-invariant memory, perception, and learning (SIMPLE) model (Brown & Chater, 2001; Brown et al., 2002), which is based on local distinctiveness principles. According to SIMPLE, the success of retrieval of items is based on their distance in psychological space from their nearest neighbors. Items that are isolated from their nearest neighbors are recalled better than items in close proximity to others. In the case of mixed lists, psychological space is considered to be two-dimensional, with one dimension representing temporal distance (between encoding and recall of an item) and the other separating items according to their phonological similarity. This representation gives rise to an isolation effect because on mixed lists, each *D* item is temporally adjacent to two *S* items, and hence maximally distant from its neighbors along the similarity axis; whereas on a pure list, the *D* items that happen to be temporally adjacent are not guaranteed to be maximally distant in two-dimensional space.

Finally, Kelley and Nairne (2001) discussed how the feature model of immediate memory (Nairne, 1990) could be adapted to explain the effects of isolation on memory. They suggested that participants encode characteristics of items at study and that the distinctive characteristics of isolates could be used as diagnostic cues at recall. In Nairne's (1990) feature model, items are represented as vectors of features, and recall proceeds by matching degraded items in primary memory to a set of intact item representations in secondary memory. Kelley and Nairne (2001) showed that making one of the items in the model distinct by giving it unique features enhances recall for that item because this causes the enhanced item to overlap less with other items in secondary memory, and thus decreases the probability of a confusion.

The experiments presented here were intended to distinguish between the broad predictions of dual-stage models (the primary model, SEM, and the Burgess & Hitch [1999] model) and models positing a role of similarity at encoding (SOB, SIMPLE, and the feature model). Unlike previous mixed-list research, we varied the number of dissimilar items and focused on analysis of different error types.

Experiment 1

Our major aim in Experiment 1 was to examine the dissimilar immunity effect when only a single dissimilar item was shown on a mixed list. In Experiment 1, we manipulated a single variable, list type, within subjects. Participants recalled visually presented lists in forward order immediately after presentation.

Method

Participants and apparatus. Twenty undergraduate and postgraduate students from the Department of Psychology at the University of Western Australia participated voluntarily in exchange for course credit or remuneration of $5/hr.

---

[2] It is important to differentiate between two techniques to measure the isolation effect. Both techniques use an isolate list that consists of a set of homogeneous items plus one other item that in some way differs from its surroundings. For example, one of the list items might be printed twice as big as the other items (e.g., Kelley & Nairne, 2001) or one of the items might be a digit embedded in a list of nonsense syllables (von Restorff, 1933, cited in Hunt, 1995). The techniques differ, however, with respect to the comparison list: One technique compares recall of the isolate to performance on the same item embedded in the same surroundings but without rendering the critical item distinct (i.e., all items are printed in the same font size; Kelley & Nairne, 2001). A more conservative technique compares recall of the isolate to recall of the same item embedded in a completely heterogeneous list. For example, recall of a single digit among nonsense syllables might be compared to recall of the same digit on a list composed of a word, a nonsense syllable, a picture, and a variety of other heterogeneous stimuli (von Restorff, 1933, cited in Hunt, 1995). It is this more conservative comparison that we focus on here.

The experiment was controlled by a PC that presented all stimuli and collected and scored all responses. The same apparatus was also used in Experiments 2 and 3.

Materials. Six phonologically similar (*B, D, G, P, T, V*) and six dissimilar (*H, K, M, Q, R, Y*) consonants, taken from Henson et al. (1996), were used to construct 6-item lists. Although lists were randomly generated, familiar sequences or acronyms (e.g., *TV, BHP*) were disallowed to discourage chunking.

Twenty-five lists were constructed for each of the six types. Three list types involved a single dissimilar isolate embedded in a list of similar items. The isolate could appear at Serial Position 2 (*SDSSSS*), Position 4 (*SSSDSS*), or Position 6 (*SSSSSD*). An example of the *SDSSSS* type might be the list *B, K, T, D, P, G*. The remaining three list types replicated previous work and involved pure lists of similar (*SSSSSS*) or dissimilar (*DDDDDD*) items plus an alternating list with dissimilar items at Serial Positions 2, 4, and 6 (*SDSDSD*).

One known determinant of serial-recall performance is bigram predictability, with predictable sequences (i.e., those containing bigrams that are common in written English) more readily recalled than nonpredictable ones (cf. Henson et al., 1996). Given that the present similarity manipulations necessarily involved different sequences, which may inadvertently engender differences in predictability (Henson et al., 1996, Experiment 1), it is important to control bigram frequency between list types. The average predictability of stimuli within list type was computed using the logarithms of bigram occurrence in the corpus of Solso and Juel (1980). For example, for the *SSSSSS* list, this measure consisted of the average bigram frequency across all 30 possible combinations of *B, D, G, P, T,* and *V* (excluding repetitions). The resulting predictability values are shown in Table 1. It is clear from the table that all list types were comparable in predictability. In particular, all lists containing dissimilar items—whether 1, 3, or 6—were indistinguishable.

Procedure. We informed participants that they would be presented with lists of six letters not containing any repetitions. For each list, the uppercase letters appeared one by one in the middle of the screen, each item being overwritten by its successor. Letters appeared for 400 ms, with a 100-ms interstimulus interval (ISI). We instructed participants to read the letters in silence.

Participants recalled the list immediately after the last item was presented by typing the letters on the keyboard. As each key was pressed, the corresponding letter appeared on the screen, replacing the previously recalled item. Participants were not allowed to correct any mistakes, but they could repeat a letter at a later position if they felt it had occurred there rather than at the earlier output position. We instructed participants to type the first letter that came to mind if they were unsure, although omissions could be recorded by pressing the space bar if necessary.

Lists were separated by brief self-paced breaks. Experimental sessions lasted about 45 min.

Results and Discussion

Analyses of all experiments focused on conventional unconditional serial-position curves, which have formed the benchmark for much modeling (e.g., Burgess & Hitch, 1999; Henson, 1998; Page & Norris, 1998b). Henson et al. (1996) additionally reported conditional serial-position curves that were particularly useful for refuting chaining, but there is little theoretical reason to report them here. Nonetheless, all analyses reported below were repeated for the conditionalized data, and the results were nearly identical in all cases.

Analysis of accuracy. The serial-position curves in Figure 1 show the average proportion of participant trials on which an item was recalled in the correct position. All accuracy analyses reported in this article use this correct-in-position score.

To avoid visual clutter, the single-isolate curve is a composite serial-position curve that was obtained by collapsing across the three isolate list types. Specifically, for Serial Positions 2, 4, and 6, the composite curve shows the isolate responses for lists *SDSSSS*, *SSSDSS*, and *SSSSSD*, respectively (i.e., for the *D* items only from these lists). For the remaining serial positions (1, 3, and 5), the composite curve shows the average performance across the three isolate lists (i.e., the average performance across *S* items at these positions).

Because we were interested primarily in responses to dissimilar items at the isolate positions, statistical analysis focused on a 3 (list type: pure dissimilar, single-isolate, or alternating) × 3 (serial position: 2, 4, or 6) repeated measures analysis of variance (ANOVA; the pooled error terms of which were used to construct the confidence intervals shown in all figures). For this analysis, the *SDSSSS*, *SSSDSS*, and *SSSSSD* lists were again combined into a composite single-isolate list incorporating Serial Positions 2, 4, and 6. Likewise, for the pure and the alternating list, only responses in Positions 2, 4, and 6 were considered. The ANOVA revealed a main effect of serial position, $F(2, 38) = 7.43$, $MSE = 0.018$, $p < .01$, but did not show a significant main effect of list type, $F(2, 38) < 1.0$, $MSE = 0.011$, or an interaction between those two variables, $F(4, 76) = 1.66$, $MSE = 0.005$, $p = .17$. This shows that the D items on mixed lists were recalled no more accurately than their counterparts on the pure list, thus replicating the dissimilar immunity effect observed in previous research.

Analysis of errors. Recall errors (i.e., the complement of the proportions correct shown in Figure 1) were separated into two classes: transpositions and intrusions. A transposition occurs when a list item is recalled at the wrong output position, whereas an intrusion error refers to the recall of an item that was not presented. Figure 2 shows the proportion of responses that were transpositions plotted by output position, the construction of the composite curve being the same as for Figure 1.

Figure 2 shows that there were fewer transpositions for dissimilar items on the mixed lists than for the corresponding items on the pure list. This effect was statistically supported by a 3 (list type: pure dissimilar, single-isolate, or alternating) × 3 (serial position: 2, 4, or 6) within-subjects ANOVA using the transposition probabilities as the dependent measure. The ANOVA revealed significant main effects of list type, $F(2, 38) = 10.75$, $MSE = 0.007$, $p < .01$, and serial position, $F(2, 38) = 10.454$, $MSE = 0.008$, $p < .01$, and a significant interaction of both variables, $F(4, 76) = 3.30$, $MSE = 0.003$, $p < .05$.

The main effect of list type can be further explored by examining the confidence intervals shown in Figure 2, which confirm that the transposition rates for the dissimilar items on mixed lists

Table 1
Average Stimulus Predictability For All List Types in Experiments 1 and 2

List type	Mean predictability
Pure similar (*SSSSSS*)	1.42
Pure dissimilar (*DDDDDD*)	1.87
Alternating (*SDSDSD*)	1.84
Single isolate (e.g., *SDSSSS*)	1.83

Note. S = similar; D = dissimilar.

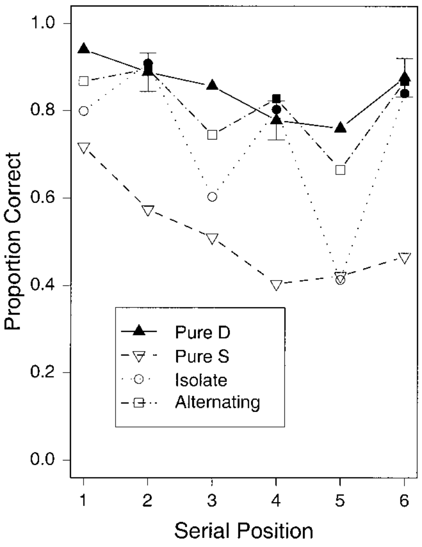

Figure 1. Serial position curves for correct-in-position recall in Experiment 1. The figure contains a single composite curve for all single-isolate lists that shows performance on the isolate in Positions 2, 4, and 6, and the average across all three isolate lists in the remaining serial positions. S = similar; D = dissimilar. The 95% confidence intervals around the dissimilar items at Serial Positions 2, 4, and 6 were calculated from the pooled error variance for the associated 3 × 3 ANOVA involving dissimilar items (see text for details). Confidence intervals in this and all other figures were computed using the recommended techniques for within-subjects designs (Loftus & Masson, 1994). Solid symbols represent dissimilar items, whereas open symbols correspond to similar items.

were significantly lower than for the corresponding items on the pure dissimilar lists (at Output Positions 2, 4, and 6). The interaction likely reflected the larger difference between mixed and pure lists at Position 4 than at Positions 2 and 6; the latter differences barely reached significance for both mixed-list conditions.

Because the proportions of all types of responses must sum to unity at each serial position, it follows that the effect on mixed-list transposition rates, which remained hidden in the accuracy analysis, must be accompanied by a compensatory effect on intrusions and omissions. Intrusion rates across output positions are shown for all conditions in Figure 3; omissions were not examined given their infrequency (just over 1% of all responses). It is clear that there were more intrusions for mixed lists than pure lists, with that difference growing larger across output positions. It is well established that intrusions tend to increase across output positions (e.g., Henson, 1998).

Implications. Before discussing the fate of mixed-list items, it should be noted that Experiment 1 replicated the common finding that phonological similarity harms order memory while leaving item memory intact when rhyming items are used in immediate serial-recall tasks (Bjork & Healy, 1974; Fallon, Groves, & Tehan, 1999; Poirier & Saint-Aubin, 1996; Lian, Karlsen, & Winsvold, 2001; Wickelgren, 1965b). In the present data, this was reflected in the large difference in transposition rates between the pure similar and pure dissimilar lists, accompanied by virtually identical intrusion rates (see Figures 2 and 3).

Turning to the dissimilar immunity effect, the accuracy serial-position curves were consistent with those of Baddeley (1968) and Henson et al. (1996) in showing no difference between dissimilar items on mixed and pure lists. However, further analysis of the underlying errors revealed that dissimilar items on mixed lists were more accurately ordered than the corresponding items on pure lists, as was reflected in the transposition probabilities. The

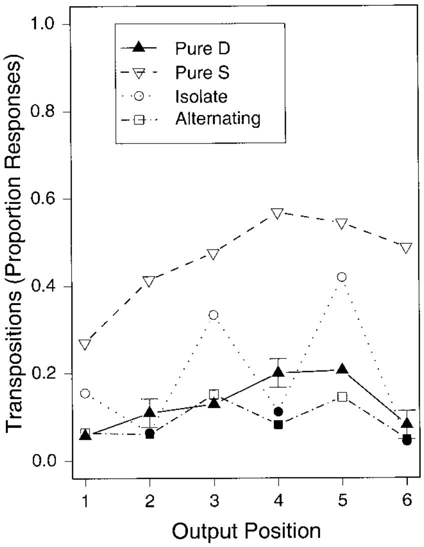

Figure 2. Transposition errors across output positions in Experiment 1. The composite curve for the isolate conditions and confidence intervals were formed as in Figure 1. Solid symbols represent dissimilar items, whereas open symbols correspond to similar items. D = dissimilar; S = similar.

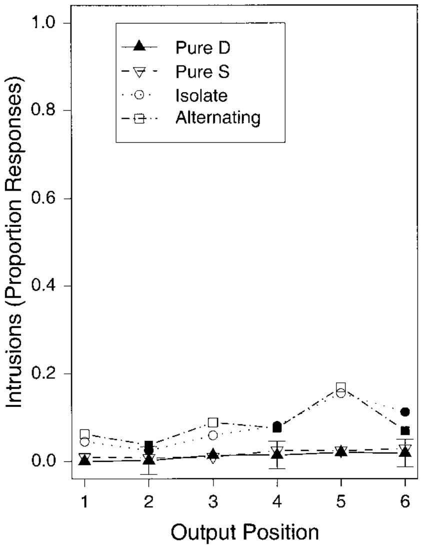

Figure 3. Intrusion errors across output positions in Experiment 1. The composite curve for the isolate conditions and confidence intervals were formed as in Figure 1. Solid symbols represent dissimilar items, whereas open symbols correspond to similar items. D = dissimilar; S = similar.

ordering advantage was offset by a greater rate of intrusions for mixed lists.

There is no a priori reason to expect these two opposing effects to be of the same magnitude, and we expect it is mere coincidence that they cancelled each other out in Experiment 1 and, at least within the constraints of statistical power, in previous reports of dissimilar immunity as well. That said, careful inspection of the data of Henson et al. (1996) reveals occasional deviations between mixed and pure lists that confirm that the trade-off between order memory and intrusion rates need not always be perfectly balanced (see further discussion after presentation of Experiment 2).

Although the accuracy data were consistent with the predictions of the primacy model, SEM, and the Burgess and Hitch (1999) model, the transposition data prove challenging to these models. Because order confusions are taken to be separate from confusions based on phonological similarity, in their present form the models cannot account for the advantage in ordering accuracy, reflected in lower transposition rates, which was observed for dissimilar items in mixed lists. By contrast, for the reasons outlined at the outset, the results are largely compatible with models that postulate an effect of similarity at encoding (e.g., SIMPLE [Brown et al., 2002], SOB [Farrell & Lewandowsky, 2002], and the feature model [Kelley & Nairne, 2001]).

We next examine whether the isolation effect observed for the transpositions is necessarily tied to an increased rate of intrusions for mixed lists, which might be explained by assuming a trade-off of encoding of item and order information (cf. Hockley & Cristi, 1996; Lewandowsky & Murdock, 1989). A connection between intrusions and the isolation effect would not be expected from models such as SOB, SIMPLE, and the feature model. We designed Experiment 2 to discourage intrusions during recall, in the expectation that this would shift the locus of the isolation effect from transpositions into the accuracy serial-position curves.

Experiment 2

In Experiment 2, we used a reconstruction task similar to that used in previous experiments on the isolation effect in serial recall (Kelley & Nairne, 2001). By providing participants with all list items at recall, the occurrence of intrusions and omissions should be reduced, if not eliminated, thus equalizing these errors between mixed and pure lists.

Method

Participants. A new sample of 20 undergraduate and postgraduate students from the Department of Psychology at the University of Western Australia participated voluntarily in exchange for course credit or remuneration of $5/hr.

Materials and procedure. Experiment 2 used the same stimuli as Experiment 1. To maximize comparability between the studies, each participant in this experiment received identical lists, in the same order, as a randomly matched participant in Experiment 1.

With the exception of the recall task, the procedure was identical to that of Experiment 1. Immediately after list presentation, all list items were shown in random order at the top of the screen. Participants were instructed to use this information to aid recall. Serial recall then proceeded as in Experiment 1, except that each recall resulted in that letter being darkened in the display of all list items.

This procedure was similar to the common variant of the reconstruction task (Crowder, 1979; Whiteman, Nairne, & Serra, 1994); but because recall still had to be in strict forward order it also retained comparability with Experiment 1 and other serial-recall studies. Although the presence of list items should have minimized intrusions, erroneous repetitions, and omissions, participants could still commit these errors by entering any item or by pressing the space bar. These responses were allowed to retain maximal compatibility with Experiment 1 and previous experiments (Henson et al., 1996).

Results and Discussion

Analysis of accuracy. The accuracy serial-position curves for Experiment 2 are shown in Figure 4. In contrast to Experiment 1, Figure 4 suggests that isolates on mixed lists were better recalled than items in the corresponding serial positions on the pure *D* list. To confirm this effect, the data in Figure 4 were subjected to a 3 (list type) × 3 (serial position) ANOVA using the same technique to aggregate data across isolate lists as in Experiment 1. The ANOVA revealed significant main effects of list type, $F(2, 38) = 4.23$, $MSE = 0.019$, $p < .05$, and serial position, $F(2, 38) = 9.04$, $MSE = 0.024$, $p < .01$, as well as a significant interaction between both variables, $F(4, 76) = 3.19$, $MSE = 0.010$, $p < .05$. The list-type effect and the interaction can be explored by considering the confidence intervals for the control items on the pure dissimilar lists shown in Figure 4. With the exception of the alternating list at the second serial position and the single isolate at Position 6, recall of dissimilar items on mixed lists was more accurate than that of the control items on the pure dissimilar lists. As for Experiment 1, the interaction between list type and serial position can be explained by noting the larger differences at Serial Position 4.

The serial-position curves from this experiment are clearly inconsistent with those in the previous studies of Henson et al. (1996) and Baddeley (1968), with the dissimilar immunity effect in those earlier studies being replaced by an isolation effect here.

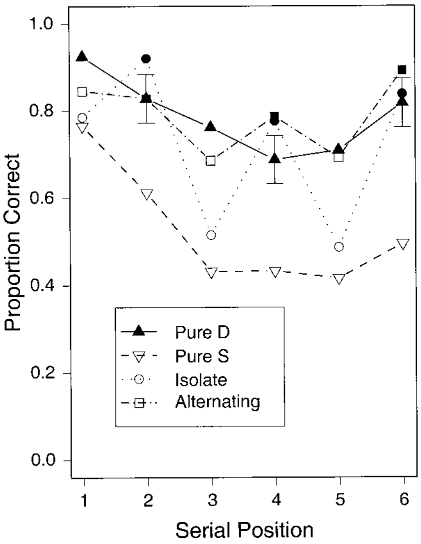

Figure 4. Serial-position curves for correct-in-position recall in Experiment 2. The composite curve for the isolate conditions and confidence intervals were formed as in Figure 1. Solid symbols represent dissimilar items, whereas open symbols correspond to similar items. D = dissimilar; S = similar.

Moreover, in contrast to Experiment 1, the isolation effect here was observed not only in the transposition rates but also at the level of overall accuracy. This was due to the low incidence of intrusion and omission errors, which formed 1.4% and 0.01% of responses, respectively.

The reconstruction results provide further evidence against current accounts of phonological similarity effects that imply dissimilar immunity (e.g., Burgess & Hitch, 1999; Henson, 1998; Page & Norris, 1998a). Although it might be argued that these models were not developed to handle the reconstruction task, this objection can be dismissed by considering the fact that our variant of the task was very similar to standard forward serial recall, the only difference being that participants were provided with the identity of items at recall (and indeed were free to ignore them).

Implications of Experiments 1 and 2. The results of Experiments 1 and 2 can be reconciled with previous results if one assumes that those earlier experiments also produced an undetected trade-off between intrusions and transpositions. What remains to be resolved is why the trade-off occurred in Experiment 1 (and perhaps in previous studies) in the first place. In particular, why were there more intrusions with mixed lists than pure lists in Experiment 1? Our preferred explanation invokes a guessing strategy. After some experience with the experimental lists, participants in Experiment 1 are likely to have become aware of the set of vocabulary items, and may have used this awareness to eliminate response candidates for certain types of lists. As pure lists contained all possible similar or all possible dissimilar items, the set of response candidates for those lists was limited to six items. However, in mixed lists, the set of potentially recallable items included all 12 letters, thus opening the door to extralist intrusions when participants guessed. This guessing account is compatible with the results of Baddeley's study (1968, Experiment 5), in which participants could see the 12 vocabulary items throughout the experiment, which undoubtedly facilitated analysis of the list structures. Thus, to avoid intrusions on pure lists, participants needed to remember only whether the list items were similar or dissimilar and to constrain the set of response candidates accordingly.

In this context, it is also noteworthy that when the intrusion rates in the Henson et al. studies were low (1996, Experiment 1; intrusions 3% and omissions 5%), the data visually suggested an isolation effect, as we would expect, although its magnitude was not statistically significant.[3] Conversely, when the rate of item errors (i.e., intrusions and omissions) was higher (Henson et al., 1996, Experiments 2 and 3; up to 23% item errors), no trace of an isolation effect was found.

If such a strategy were employed by participants, it follows that the trade-off should be prevented if ensemble size is experimentally equalized across all list types. We explored this possibility in Experiment 3.

Experiment 3

In Experiment 3, we used standard serial recall but equalized ensemble size across list types. We expected this to equalize intrusion rates between mixed and pure lists, which in turn would be expected to reveal an isolation effect in the accuracy serial-position curves even with standard serial recall.

Ensemble size was controlled by introducing a larger set of dissimilar items, all of which were used on pure lists but only some of which were used on mixed lists. This was expected to equalize intrusion rates between pure dissimilar and mixed lists. If this manipulation were to produce a phonological isolation effect with the recall task and serial position analysis used by Baddeley (1968) and Henson et al. (1996), it would provide a particularly strong challenge to models that relegate the effects of phonological similarity to a confusion stage at output.

Method

Participants and design. A new sample of 17 members of the campus community at the University of Western Australia participated voluntarily in exchange for remuneration of $5 per hr.

We manipulated the single experimental variable, list type, within subjects. There were four different types of 6-item lists: Two single-isolate lists, *SDSSSS* and *SSSDSS*, the alternating list (*SDSDSD*), and the pure dissimilar control list (*DDDDDD*). Pure similar lists were not present in this study. We randomly constructed 40 lists of each type for each participant.

Materials. We constructed lists by randomly sampling letters from three 5-letter ensembles. All similar stimuli were drawn from the set *B, T, D, G, P*. Dissimilar letters were drawn from one of two ensembles consisting of *H, K, M, Q, R,* and *X, L, Z, W, Y,* respectively. (The letter *Z* in Australian usage does not rhyme with *B* or *T* but is pronounced "zed.")

All mixed lists were constructed from the similar ensemble and one of the dissimilar ensembles, with the identity of the dissimilar ensemble counterbalanced across participants. For example, one participant might receive an *SDSSSS* list composed of *G, K, P, D, T, B,* in which the *K* could be replaced by any of *M, Q, R,* and *H* on other trials within the same block. Another participant would receive lists of the kind *G, X, P, D, T, B,* in which *X* could be any of *L, Z, W,* and *Y* across trials within the same block. The other mixed lists (*SSSDSS* and *SDSDSD*) followed the same structure, thus ensuring that all mixed lists were drawn from a total vocabulary of 10 potential stimuli for each participant. The pure dissimilar lists (*DDDDDD*) were sampled from both dissimilar ensembles for all participants, thus equalizing vocabulary size for mixed and pure lists. As in Experiments 1 and 2, familiar sequences or acronyms (e.g., *TV, BHP*) were disallowed to prevent chunking.

The results of the bigram analysis are shown in Table 2. In contrast to the preceding studies, predictabilities differed somewhat between the various lists, with the largest difference being around .40. This difference was marginally smaller than the largest difference between list types in the Henson et al. (1996) Experiments 2 and 3, which were specifically designed to equalize predictability. (Their means ranged from .70 to 1.14, so proportionally their spread was considerably larger than ours.) By the same token, our largest difference was less than 25% of the magnitude of the largest difference in the Henson et al. Experiment 1, which they identified as a possible cause for mean differences in accuracy. Therefore, we conclude that list predictability was adequately controlled in Experiment 3.

All lists of each type were presented together in a block of 40 trials, with the order of blocks randomized separately for each participant, subject to the constraint that the block of pure dissimilar lists was always presented last. This ensured that the second dissimilar ensemble that was needed for

---

[3] Henson et al. (1996) attributed the slight isolation effect in their first experiment to the significantly higher bigram predictability of one of their mixed lists. In support, they cited the significantly better overall performance level associated with that list, which they claimed had moved the dissimilar items on mixed lists above the level of the pure list. However, that reasoning remains unconvincing because a main effect of list performance says nothing about the position of particular points on the sawtooth.

Table 2
Average Stimulus Predictability For All List Types in Experiment 3

List type	Dissimilar ensemble	
	H, K, M, Q, R	*X, L, Z, W, Y*
Alternating (*SDSDSD*)	1.88	1.74
Single isolate (e.g., *SDSSSS*)	1.89	1.81
Pure dissimilar (*DDDDDD*)	1.50	

Note. S = similar; D = dissimilar.

the *DDDDDD* lists did not become part of the vocabulary during the earlier mixed-list blocks.

Procedure. The procedure was identical to Experiment 1 except that self-paced breaks were inserted after every 20 trials. Experimental sessions lasted under an hour.

Results and Discussion

Accuracy analysis. Figure 5 shows the accuracy serial-position curves for all list types in Experiment 3, with the single-isolate curve being formed as in Experiments 1 and 2. Because the *SSSSSD* lists were not present in this experiment, factorial analysis of a subset of critical serial positions (2, 4, and 6) across list types was not possible here. Instead, the data were subjected to a 4 (list type) × 6 (serial position) within-subjects ANOVA. The analysis revealed a main effect of list type, $F(3, 48) = 29.84$, $MSE = 0.026$, $p < .01$, and a main effect of serial position, $F(5, 80) = 75.75$, $MSE = 0.022$, $p < .01$, as well as a significant interaction between both variables, $F(15, 240) = 21.53$, $MSE = 0.009$, $p < .01$. The crucial interaction was further explored by paired-samples t tests that revealed significant differences between the pure list (*DDDDDD*) and the corresponding single isolate at Serial Positions 2, $t(16) = 4.283$, $p < .01$, and 4, $t(16) = 2.960$, $p < .01$. Significant differences were also found between the alternating list and the pure list at Serial Positions 2, $t(19) = 2.363$, $p = .03$, and 4, $t(19) = 5.801$, $p < .01$. There was no significant effect involving the alternating list at Serial Position 6, $t(19) < 1$.

These results attest to the success of the efforts to equalize ensemble size across lists: Unlike in Experiment 1, we observed an isolation effect in the accuracy analysis. However, before this conclusion can be fully accepted, one other difference between Experiment 1 and the present experiment must be considered. In Experiment 1, we presented participants with 25 lists of each type, which is roughly comparable to the number of lists per block (18 or 21) used by Henson et al. (1996) in two of their studies. Experiment 3, by contrast, involved 40 lists of each type. It is therefore conceivable that participants had more opportunity for practice and strategy development than in previous studies. This possibility was checked by comparing the results between the first and second half of each block.

Comparison of early and late trials. The preceding analysis was repeated with the addition of another within-subjects variable, position-within-block, which contrasted performance on the first 20 and the last 20 trials within each block. This 2 × 4 × 6 analysis revealed a significant main effect of position-within-block, $F(1, 16) = 15.09$, $MSE = 0.018$, $p = .01$, but, importantly, position-within-block did not interact with any of the remaining variables (list type, serial position), with all Fs < 1. Thus, aside from an overall increase in performance with practice (mean proportion correct for first vs. second half was .656 vs. .695), all relevant effects were unaffected by whether lists were presented early or late in a block.

In further confirmation, three out of four pairwise comparisons involving isolates that were significant in the overall analysis remained significant when repeated for each half of trials separately.

Transposition and intrusion errors. Figure 6 shows the transposition curves underlying the serial-position curves in Figure 5. The corresponding 4 × 6 ANOVA revealed main effects of list type, $F(3, 48) = 53.09$, $MSE = 0.022$, $p < .001$, and serial position, $F(3, 48) = 66.40$, $MSE = 0.010$, $p < .001$, along with an interaction between both variables, $F(3, 48) = 35.12$, $MSE = 0.006$, $p < .001$. Further exploration of the interaction yielded significant differences between the pure list and the single isolate at Serial Position 2, $t(16) = 5.11$, $p < .001$, and at Serial Position 4, $t(16) = 5.26$, $p < .001$. The difference between the pure list and the alternating list was significant at Serial Positions 2, $t(16) = 2.80$, $p < .05$, and 4, $t(16) = 6.23$, $p < .001$, but not at Serial Position 6, $t(16) < 1$. Thus, the pattern of differences for the transposition data was identical to that in the accuracy serial-position curves.

As expected, the frequency of intrusion errors for all lists was approximately the same (omissions formed less than 0.01% of responses). Specifically, the proportion of intrusions was .092 for the pure dissimilar lists, .063 and .076 for the two single-isolate lists (*SDSSSS* and *SSSDSS,* respectively), and .119 for alternating lists.

Experiment 3 showed that the dissimilar immunity effect reported in earlier studies vanishes when ensemble size, and hence

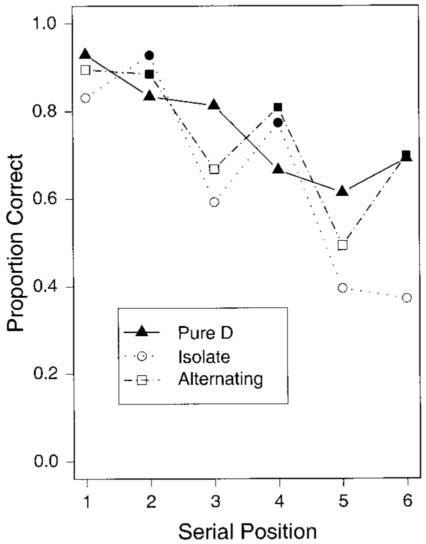

Figure 5. Serial-position curves for correct-in-position recall in Experiment 3. The composite single-isolate curve was formed as for Figure 1, except that the point at Serial Position 6 gives the average of performance on similar items. Solid symbols represent dissimilar items, whereas open symbols correspond to similar items. D = dissimilar.

intrusion rates, are equalized between list types. Using the same serial-recall task that was used in those previous studies, this experiment produced a reliable isolation effect in the standard accuracy serial-position curves.

General Discussion

Summary of Results

We reported three experiments that produced a phonological-isolation effect in short-term serial recall. The isolation effect was reflected in improved order memory for dissimilar items embedded in similar lists, compared with the same items at the same serial positions in pure dissimilar lists. In Experiment 1, we used standard serial recall and the effect was apparent in the transposition probabilities but remained hidden in the accuracy analysis owing to the greater number of intrusions in mixed lists. In Experiment 2, which discouraged intrusions by providing the list items at recall, the isolation effect was observed in the accuracy analysis. In Experiment 3, standard serial recall was again used, but the number of intrusions was indirectly equalized across list types by controlling the size of the stimulus ensemble. Accordingly, an isolation effect was again observed in the accuracy analysis.

To underscore the magnitude and consistency of the effect across the three studies, Figure 7 summarizes the transposition data from all experiments. The figure plots the probability of a transposition for a dissimilar item on a mixed list (either single-isolate or alternating) as a function of the probability of transposing an item in the corresponding serial position in a pure dissimilar list. The diagonal represents idealized dissimilar immunity; from the previous literature, all data points would be expected to lie on or close to this diagonal. However, contrary to that expectation, in the present experiments nearly all observations deviated systematically and considerably from the diagonal, such that there were always more transpositions for an item in a pure dissimilar list than in a mixed list. This was statistically demonstrated by the analyses of the individual experiments, and Figure 7 provides an additional visual summary of the consistency and magnitude of the effect across all studies.

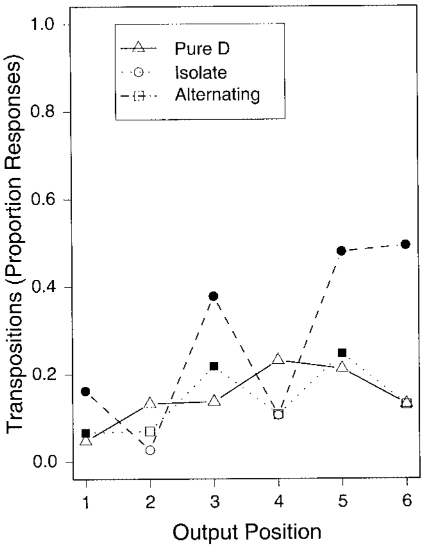

Figure 6. Transposition errors by output position for all list types in Experiment 3. The composite single-isolate curve was formed as for Figure 1. Solid symbols represent dissimilar items, whereas open symbols correspond to similar items. D = dissimilar.

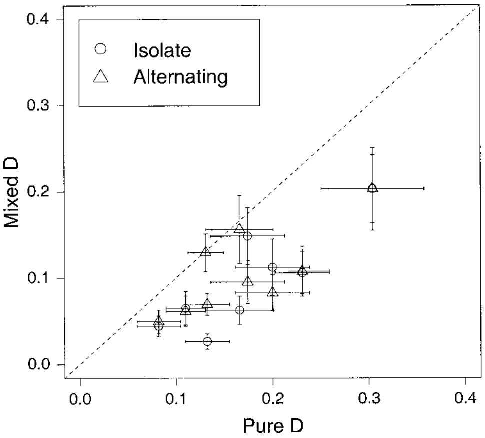

Figure 7. Transposition rates for isolate items in all experiments on mixed lists (mixed *D*) and for items in corresponding serial positions on pure lists (pure *D*). Standard errors are indicated by horizontal and vertical bars and are computed as though all comparisons were between participants. Therefore, they are likely to be overestimates. Mixed lists include alternating (*SDSDSD*) and single-isolate lists (e.g., *SDSSSS*). If dissimilar immunity holds, all observations would be expected to lie on or close to the diagonal. D = dissimilar; S = similar.

Another consistent result was the small size of the isolation effect, or indeed its absence, for the terminal serial position. As we discuss below, this turns out to be qualitatively compatible with the predictions of the SOB model of Farrell and Lewandowsky (2002).[4]

Relationship to Previous Findings

Our experiments are related to two streams of research—investigations of phonological similarity in short-term memory and the isolation effect—that have thus far been pursued in isolation.

Concerning phonological similarity, the discrepancy between our results and previously reported findings underscores the importance of examining error patterns in preference to analyzing accuracy serial-position curves (cf. Saint-Aubin & Poirier, 1999). Nonetheless, our results do not undermine the earlier conclusions of Baddeley (1968) and Henson et al. (1996) regarding the shortcomings of chaining models. If anything, our results further call into question a simple chaining account of serial recall as offered by Wickelgren (1965b) and Lewandowsky and Murdock (1989), because we show enhanced order memory for isolates, where

---

[4] There is one exception to this otherwise consistent effect of serial position. In Experiment 2, the isolation effect was absent in Serial Position 2 and strongest in Serial Position 6. This anomaly, which arose with a reconstruction task, came about for unknown reasons and it disrupts an otherwise consistent effect of serial position on isolation.

chaining models predict a decrement in recall (see, e.g., Baddeley et al., 1991).

Turning to research on isolation effects, our results establish the existence of a phonology-based isolation effect (or "von Restorff" effect) in short-term serial recall. This is a novel finding that extends the generality of other, visually based isolation effects in serial recall (e.g., Kelley & Nairne, 2001). If Experiment 1 were considered in isolation, one might argue that although isolation improved memory for order, it also impaired item memory—quite at odds with the preponderance of relevant results. However, Experiments 2 and 3 clarified that when guessing is experimentally controlled, isolation benefits order memory without a concomitant increase in intrusions.

Another characteristic of isolation effects is that their magnitude appears to decrease as additional isolates are added to the list (Newman & Jennette, 1975). However, the data of Newman and Jennette were not entirely convincing: Even when half the list consisted of isolates[5] they continued to be recalled better than items on a control list—although a single isolate was recalled somewhat better still. In our studies, putting aside performance at Serial Position 6, the isolation advantage appeared to be unaffected by the total number of isolates on the list.

That said, the number of isolates did have a large effect on recall of the surrounding similar items. In Experiments 1 and 2, similar items in single-isolate lists were recalled better than their counterparts on the pure similar lists; and in all experiments, similar items on the alternating list were recalled better than similar items that accompanied a single isolate. This result turns out to be compatible with a "streaming" explanation of the isolation effect.

According to this explanation, people organize an isolate list into two ad hoc categories: one containing the isolate(s) and one containing the remaining list items (e.g., Fabiani & Donchin, 1995; Hunt & Lamb, 2001). Intuition and theory (e.g., cue overload; Watkins & Watkins, 1975) dictate that items from the smaller category are better recalled than items from the larger one—this being the cause of the isolation effect. One implication of this view is that the size of the "background" category, which contains all nonisolates, is smaller for lists containing isolates than for control lists. It follows that the background items on isolate lists should be recalled better than comparable items from the control list. Although this prediction has not always been confirmed (e.g., Kelley & Nairne, 2001), it is supported by the present data, especially in the case of alternating lists, in which half of the list items were isolates.

Implications for Models of Memory

Dual-stage models. The primacy model (Page & Norris, 1998a, 1998b), SEM (Henson, 1998), and the Burgess and Hitch (1999) model, by virtue of their assumption of the independence of item ordering and phonological similarity, predict that recall for dissimilar items on mixed lists will be identical to that of the same items on pure lists. Because phonological similarity affects only a second stage, downstream from the mechanism maintaining order, and because its effect is confined to the set of mutually similar list items, these models have no obvious explanation for the isolation effect presented in this article. Although it is possible that the ordering stage in those models might be modified to capture isolation phenomena, the current results call into question the initial empirical motivation for dual stages of processing.

The isolation effect observed here is not the only result problematic for these models. In Experiments 1 and 2, memory for the order of similar items on mixed lists was sometimes comparable to that of dissimilar items in pure lists. For example, consider the transpositions at Serial Positions 1, 3, and 5 in the alternating list of Experiment 1 (Figure 2), or the (virtually intrusion free) accuracy results in Experiment 2 (Figure 4). Although recall of similar items on mixed lists often falls in between the two types of pure lists (e.g., Baddeley, 1968; Henson et al., 1996; the present Experiment 3), the fact that some similar items were ordered as accurately as some dissimilar items in Experiments 1 and 2 is troublesome for the primacy model and SEM, in which phonological similarity maps into a single parameter that determines the confusability of similar items in the secondary stage. For these models to predict identical transposition probabilities for mixed-list *S* items and pure-list *D* items, this confusion parameter would need to be set to 0 to render all items equally immune to phonological similarity. However, this would cause the predicted serial-position curves for pure *S* and pure *D* lists to coincide, which is inconsistent with all known results.

We suggest that dual-stage models could account for our findings only by allowing interitem similarity to have an effect on the encoding of all items.

Similarity at encoding. We referred to three models at the outset in which similarity plays a role at encoding: SOB (Farrell & Lewandowsky, 2002), SIMPLE (Brown et al., 2002), and the feature model (Nairne, 1990). Although it does not necessarily follow that these models can handle the present results, the models' qualitative predictions mirror our data.

In some applications, the SOB model (Farrell & Lewandowsky, 2002) assumes that the strength with which an incoming item is encoded is inversely proportional to its similarity to the previous contents of memory (Farrell, 2001). Thus, if an incoming item is similar to previous ones, it is encoded with less strength than a dissimilar item would be under identical circumstances. Because retrieval is based on the relative strengths of items, dissimilar items in mixed lists have an advantage over their counterparts on pure lists because of reduced competition from following items. It follows that SOB does not predict an isolation effect for the terminal serial position because the final *D* item (ignoring transpositions) suffers no competition from yet-to-be recalled items in later serial positions, hence it cannot benefit from a reduction in that competition offered by *S* items.

Both SIMPLE (Brown et al., 2002) and the feature model (Kelley & Nairne, 2001; Nairne, 1990) explain isolation effects with reference to the enhanced discriminability of isolated items. Because the isolated items are more distinct (or share fewer features with other list items; Kelley & Nairne, 2001), they will be confused less with other items at recall. Although these models can explain the overall enhanced recall of dissimilar items on mixed

---

[5] When half the list items are "isolated," the label *isolate* may no longer seem appropriate. For simplicity and consistency, we continue to use the term *isolate* even when other list items share its characteristics, although we acknowledge that this may go beyond the original usage of the term.

lists in these experiments, it is less clear how they would account for the absence of an isolation effect at the final serial position.

A potential limitation of our experiments is that lists of different types were blocked together, which might have enabled participants to learn to predict the structure of upcoming lists, and to use this "metaknowledge" to preferentially encode the position of the isolates. However, two points speak against this possibility. First, the nonsignificant interaction of the position of lists within block with any variables, as reported in the split-half analysis of blocks (Experiment 3), does not support the idea that people develop a list-specific encoding strategy with practice. Second, Henson et al. (1996) also presented their lists in blocks, which implies that any argument based on encoding strategy would have to account for the presence as well as the absence of dissimilar immunity, further invalidating this argument.

Conclusion

The results presented here clearly demonstrate enhanced ordered recall for dissimilar items on mixed lists of phonologically similar and dissimilar items. On balance, our data seem to present a challenge to models in which the representation of order of items is independent of their phonological similarity. By the same token, the results support models in which phonological similarity plays a role in the encoding of order. The extent to which these various models can handle our results at a quantitative level is yet to be determined.

References

Anderson, J. A., Silverstein, J. W., Ritz, S. A., & Jones, R. S. (1977). Distinctive features, categorical perception, and probability learning: Some applications of a neural model. *Psychological Review, 84,* 413–451.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology, 18,* 362–365.

Baddeley, A. D. (1968). How does acoustic similarity influence short-term memory. *Quarterly Journal of Experimental Psychology, 20,* 249–263.

Baddeley, A. D., Papagno, C., & Norris, D. (1991). Phonological memory and serial order: A sandwich for TODAM. In W. E. Hockley and S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 175–194). Hillsdale, NJ: Erlbaum.

Bjork, E. L., & Healy, A. F. (1974). Short-term order and item retention. *Journal of Verbal Learning and Verbal Behavior, 13,* 80–97.

Bone, R. N., & Goulet, L. R. (1968). Serial position and the von Restorff isolation effect. *Journal of Experimental Psychology, 76,* 494–496.

Brown, G. D. A., & Chater, N. (2001). The chronological organization of memory: Common psychological foundations for remembering and timing. In C. Hoerl & T. McCormack (Eds.), *Time and memory: Issues in philosophy and psychology* (pp. 77–110). Oxford, England: Oxford University Press.

Brown, G. D. A., Neath, I., & Chater, N. (2002). *A local distinctiveness model of scale-invariant memory and perceptual identification.* Manuscript submitted for publication.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review, 107,* 127–181.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106,* 551–581.

Cimbalo, R. S., Nowak, B. I., & Soderstrom, J. A. (1981). The isolation effect in children's short-term memory. *Journal of General Psychology, 105,* 215–223.

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.

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

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology, 55,* 429–432.

Crowder, R. G. (1979). Similarity and order in memory. In G. H. Bower (Vol. Ed.), *The psychology of learning and motivation: Vol. 13. Advances in research and theory* (319–353). San Diego, CA: Academic Press.

Dunlosky, J., Hunt, R. R., & Clark, E. (2000). Is perceptual salience needed in explanations of the isolation effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 649–657.

Fabiani, M., & Donchin, E. (1995). Encoding processes and memory organization: A model of the von Restorff effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 224–240.

Fallon, A. B., Groves, K., & Tehan, G. (1999). Phonological similarity and trace degradation in the serial recall task: When CAT helps RAT, but not MAN. *International Journal of Psychology, 34,* 301–307.

Farrell, S. (2001). *Similarity-sensitive encoding, redintegration, and response suppression in serial recall.* Unpublished doctoral thesis, University of Western Australia, Perth, Australia.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review, 9,* 59–79.

Henry, L. A. (1991). The effects of word length and phonemic similarity in young children's short-term memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 43*(A), 35–52.

Henson, R. N. A. (1998). Short-term memory for serial order: The start–end model. *Cognitive Psychology, 36,* 73–137.

Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49*(A), 80–115.

Hirshman, E., & Jackson, E. (1997). Distinctive perceptual processing and memory. *Journal of Memory and Language, 36,* 2–12.

Hockley, W. E., & Cristi, C. (1996). Tests of encoding tradeoffs between item and associative information. *Memory & Cognition, 24,* 202–216.

Hunt, R. R. (1995). The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review, 2,* 105–112.

Hunt, R. R., & Lamb, C. A. (2001). What causes the isolation effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1359–1366.

Kelley, M. R., & Nairne, J. S. (2001). Von Restorff revisited: Isolation, generation, and memory for order. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 54–66.

Lewandowsky, S., & Murdock, B. B. (1989). Memory for serial order. *Psychological Review, 96,* 25–57.

Lian, A., Karlsen, P. J., & Winsvold, B. (2001). A re-evaluation of the phonological similarity effect in adults' short-term memory of words and nonwords. *Memory, 9,* 281–299.

Lippman, L. G. (1980). Perceptual isolation and position judgement. *American Journal of Psychology, 93,* 683–695.

Lippman, L. G., & Lippman, M. Z. (1978). Isolation and similarity effects in a serial reconstruction task. *American Journal of Psychology, 91,* 35–50.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subjects designs. *Psychonomic Bulletin & Review, 1,* 476–490.

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

Newman, S. E., & Jennette, A. D. (1975). The effects of set size on learning an item in the set. *American Journal of Psychology, 88,* 117–124.

Page, M. P. A., & Norris, D. (1998a). The primacy model: A new model of immediate serial recall. *Psychological Review, 105,* 761–781.

Page, M. P. A., & Norris, D. (1998b). Modeling immediate serial recall with a localist implementation of the primacy model. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 227–255). Hillsdale, NJ: Erlbaum.

Poirier, M., & Saint-Aubin, J. (1996). Immediate serial recall, word frequency, item identity, and item position. *Canadian Journal of Psychology, 50,* 408–412.

Saint-Aubin, J., & Poirier, M. (1999). The influence of long-term memory factors on immediate serial recall: An item and order analysis. *International Journal of Psychology, 34,* 347–352.

Solso, R. L., & Juel, C. L. (1980). Positional frequency and versatility of bigrams for two through nine-letter English words. *Behavior Research Methods & Instrumentation, 12,* 297–343.

Wallace, W. P. (1965). Review of the historical, empirical, and theoretical statutes of the von Restorff phenomenon. *Psychological Bulletin, 63,* 410–424.

Watkins, O. C., & Watkins, M. J. (1975). Build up of proactive inhibition as a cue overload effect. *Journal of Experimental Psychology: Human Learning and Memory, 1,* 442–452.

Whiteman, H. L., Nairne, J. S., & Serra, M. (1994). Recognition and recall-like processes in the long-term reconstruction of order. *Memory, 2,* 275–294.

Wickelgren, W. A. (1965a). Acoustic similarity and retroactive interference in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 4,* 53–61.

Wickelgren, W. A. (1965b). Short-term memory for phonemically similar lists. *American Journal of Psychology, 78,* 576–574.

Winters, J. J., & Hoats, D. L. (1989). On the development of the isolation effect in free recall: Isolation by color. *Bulletin of the Psychonomic Society, 27,* 219–222.

Received February 11, 2002
Revision received January 20, 2003
Accepted January 27, 2003 ■