

## Temporal Grouping Effects in Immediate Recall: A Working Memory Analysis

Graham J. Hitch

*Lancaster University, Lancaster, U.K.*

Neil Burgess

*University College, London, U.K.*

John N. Towse

*Royal Holloway, London, U.K.*

Vicki Culpin

*Lancaster University, Lancaster, U.K.*

The presence of temporal pauses during list presentation can markedly improve immediate memory for a sequence of verbal items. A series of experiments analysed this effect using Baddeley's (1986) model of working memory. Experiment 1 showed that the effect of temporal grouping on memory for visual sequences was removed by either articulatory suppression or reciting random digits. Experiment 2 indicated that effects of temporal grouping were insensitive to the word length of the items. Experiment 3 showed that articulatory suppression did not remove the temporal grouping effect for auditory lists. Experiment 4 showed that the temporal grouping effect was insensitive to the phonemic similarity of the items. The effects of concurrent articulation suggest that grouping influences the phonological loop component of working memory. However, the working memory model is insufficiently well specified to account for the insensitivity of grouping effects to word length and phonemic similarity. The main findings could be simulated by a connectionist model of the phonological loop, which invokes a context timing signal (Burgess & Hitch, 1992, in press). This assumed that pauses during list presentation affect the timing signal in a similar way to the pause before list presentation and made some novel predictions.

---

Requests for reprints should be sent to Graham J. Hitch, Department of Psychology, Lancaster University, Lancaster LA1 4YF, U.K. Email: G.Hitch@lancaster.ac.uk

We thank Peter Walker for helpful comments on an earlier version of the manuscript, and we acknowledge useful discussions with Mike Malloch, John O'Keefe, and Jon Shapiro, as well as suggestions made in review by Nelson Cowan, Fergus Craik, and Bob Logie. Part of this work was supported by a grant to GJH and Jon Shapiro under the Joint Research Councils' Initiative on Cognitive Science. NB was supported by a Royal Society University Research Fellowship.

---

Organizational processes exert large effects on memory, both in remembering novel information over short intervals and in learning material over the longer term (Bower, 1972). In a major contribution towards understanding the effects of organization on serial recall, Miller (1956) analysed variations in memory span for different types of materials. He observed that span was remarkably constant when measured in "chunks", higher-order units in which the elements are organized according to a familiar pattern or rule. Thus, span for random words is approximately the same as for random letters, even though each word contains several letters. According to Miller, the span of immediate memory is equal to the "magic number" of approximately seven chunks, and this constraint reflects the involvement of a limited-capacity system. Despite problems of precise definition (see, e.g., Simon, 1974), the concept of chunking has been highly influential, particularly in the analysis of skilled memory (Chase & Ericsson, 1982). The present investigation is concerned with temporal grouping, a form of organization that consists of inserting extra pauses into an otherwise regular sequence of items. Temporal grouping is related to chunking in that it involves segmenting a sequence into smaller subsequences, and has large effects on both immediate recall and learning (Bower & Winzenz, 1969; Ryan, 1969a, 1969b). However, there is an important conceptual difference between these two forms of organization, in that temporal grouping enhances memory for unfamiliar subsequences, whereas for chunking to occur, the subsequences must form familiar patterns.

In a re-examination of Miller's concept of the magic number seven, Broadbent (1975) noted that on the criterion of totally accurate recall, memory span is typically only about three items, and three is also the optimal size of group in memory for temporally grouped lists (Ryan, 1969a). Largely because of this correspondence, he suggested that chunking and temporal grouping share a common basis. Subsequently, Broadbent (1984) identified this with "abstract working memory", a central, limited-capacity buffer that holds the elements of a serially presented chunk or temporal group while they are recoded into a higher order unit. Chunking and grouping were assumed to improve immediate recall by freeing up this buffer. Broadbent (1975, 1981) took the typical clustering of items in small groups in retrieval from long-term memory (Graesser & Mandler, 1978) as evidence that the same buffer constrains the way new information is acquired.

The aim of the present investigation was to reconsider the hypothesis that temporal grouping affects a central buffer. For example, another suggestion is that it affects rehearsal strategies (Ryan, 1969b). This alternative is supported by effects of instructions to rehearse temporally ungrouped lists in subjectively defined groups (Wickelgren, 1964, 1967). Subjective and objective (i.e. temporal) grouping have remarkably similar effects in that both types of organization enhance recall, with grouping by threes being optimal. A different though related hypothesis is that temporal grouping affects rehearsal for visual lists, but that it also has an effect on perceptual storage in the case of spoken lists, by influencing parsing of the speech input (Frankish, 1985, 1989).

The theoretical framework for the present study was the model of working memory proposed by Baddeley and Hitch (1974), as modified by Baddeley (1986). This is a modular account in which working memory is assumed to consist of separate but interconnected subsystems with limited capacities. The two most important here are the central executive and the phonological loop. Although there is no direct translation between this conceptualization and Broadbent's (1984) Maltese Cross model, his abstract

store can be located within the central executive of the working memory model. The central executive acts as a general workspace for control processes, including access to other working memory subsystems and activation of information in long-term memory (Baddeley, this volume). The phonological loop is a separate store in which speech-motor information can be maintained by subvocal rehearsal. This theoretical account is supported by a robust and extensive body of experimental evidence and has led to the development of sets of experimental manipulations that have selective effects on the different subsystems (Baddeley, 1986). For example, articulatory suppression (the repetition of a redundant word), word length, and phonemic similarity are all assumed to affect the phonological loop and can be used as converging operations for this subsystem. The method adopted here was to employ such techniques to explore how working memory subsystems are affected by the temporal grouping of inputs to the system. In a final section, limitations of the original working memory model in explaining the experimental observations lead to a new account based on a computational model (Burgess, 1995; Burgess & Hitch, 1992, in press).

## EXPERIMENT 1

Much of the support for the abstract store account of temporal grouping is based on an experiment reported by Broadbent and Broadbent (1981). They suggested there was an influence of grouping on memory for sequences of random digits even when subjects were required to suppress articulation. It was argued that use of a visual, sequential presentation procedure minimized the influence of sensory memory and that suppression blocked subvocal rehearsal. However, an obvious weakness in this argument is that the involvement of a central store is inferred by default. More direct evidence based on a manipulation that selectively affects the central component of working memory would clearly be more convincing. A second difficulty is that the evidence for a grouping effect rested on comparisons of within and between-group transition error probabilities in the recall of grouped lists. There are a number of problems with such comparisons, not least the inevitability of confounding due to serial position effects. A more straightforward test is to compare the overall levels of recall of grouped and ungrouped lists.

Experiment 1 therefore re-examined the basis of the effect of temporal grouping on memory for visual lists using (a) an ungrouped control condition, and (b) concurrent tasks placing different loads on the central and phonological components of working memory. Articulatory suppression was assumed to load selectively on the phonological loop, and the more complex task of repeating a novel multi-digit number was assumed to load both the central executive and the phonological loop (Hitch & Baddeley, 1976). If temporal grouping influences only the central workspace, the grouping effect should be disrupted by reciting random numbers but not by articulatory suppression. Alternatively, if temporal grouping affects only the phonological loop, its effect should be disrupted by both suppression and reciting random numbers. If both the workspace and the phonological loop contribute to the grouping effect, it should be interfered with by both tasks, with reciting random digits having the greater effect. Following Broadbent and Broadbent (1981), visual lists were presented sequentially in the same spatial location to exclude any possible component of the grouping effect arising in sensory memory.

## Method

### Subjects

Eighteen volunteers, most of them female, from the subject panel maintained by the Psychology Department, Manchester University, served as subjects. They were each paid £3 for participation. Nine subjects were allocated to each grouping condition.

### Design and Materials

The design was a two-way mixed factorial. The presence or absence of temporal grouping was varied between-subjects, and concurrent task was varied within-subjects. Concurrent task had three levels corresponding to control, articulatory suppression, and reciting random digits. The order in which the concurrent task conditions were completed was rotated across subjects.

Lists of 9 letters were constructed by random sampling without replacement from a pool of 12 phonologically dissimilar consonants (*C D F H J L N Q R S Y Z*), taking care to exclude meaningful clusters (e.g. *SLY*). Three sets of 23 lists were prepared, for use in the three concurrent task conditions. A Macintosh LCII computer controlled by Psychlab software was used to present the visual lists. In addition, the digits 1–5 were recorded digitally using Audio Media software, and 23 random permutations were prepared for playing back through an amplified external loudspeaker.

### Procedure

Subjects were tested individually. Typed instructions indicated that the stimuli would be separated by even or uneven pauses, depending on the condition, and asked subjects to attempt written serial recall of each letter sequence immediately after its presentation. Each trial was initiated by the subject by a key press. For ungrouped lists, each letter was displayed for 300 msec, with a blank ISI of 300 msec. For grouped lists, each letter was displayed for 300 msec, with blank ISIs of 750 msec after Items 3 and 6 and 150 msec after the remainder. Thus grouped lists were perceived as three groups of three letters, and the total presentation time matched that for ungrouped lists. Immediately following the ninth letter, three dots were displayed as a cue for recall. All stimuli appeared at the centre of the visual display. Subjects wrote on a prepared response sheet, with separate spaces for each of the 9 serial positions. They were encouraged to guess or omit items if they were uncertain. The first three trials in each concurrent task condition were treated as practice, leaving a total of 20 experimental trials per condition.

Subjects completed the three task conditions in blocks of trials separated by short breaks. In the control and suppression conditions, presentation of the letter sequence began approximately 2 sec after subjects initiated the trial. In the suppression condition, subjects were instructed to repeat the word “blah” at an even pace (approximately 2 words per sec) at the same time as initiating the trial. They continued to suppress until the signal for recall. In the digit recitation condition, initiation of each trial was followed by auditory presentation of a random 5-digit permutation, which took approximately 2 sec. This was followed by a 2-sec interval and the letter stimuli, as previously, with subjects repeating the number sequence throughout presentation, stopping at the signal for recall. Performance of the concurrent tasks was monitored by the experimenter, who ensured that subjects maintained an even pace. There were virtually no errors on the digit recall task. Each experimental session lasted approximately 20–30 min.

## Results

In all the present experiments, recall was scored correct only when an item was reported in the correct position. The mean numbers of letters recalled are shown in Figure 1, from which it is clear that the advantage of temporal grouping was markedly reduced by the requirement to perform either concurrent task. The data were subjected to a two-way analysis of variance (ANOVA) with factors grouping and concurrent task. There was a significant Grouping  $\times$  Concurrent Task interaction,  $F(2, 32) = 9.37, p < 0.01$ . Analysis of simple effects showed that although grouping improved recall in the control condition,  $F(1, 32) = 47.4, p < 0.01$ , there was no reliable grouping effect with either articulatory suppression,  $F(1, 32) = 4.07$ , or reciting random digits,  $F(1, 32) = 1.58$ .

Serial position curves exhibited the normal pattern for visual presentation, namely strong primacy coupled with weaker recency (see Figure 2). In the control condition the curves for grouped lists showed the typical scalloped profile. However, the scalloping was less marked than with auditory lists (Ryan, 1969a), consistent with evidence that grouping effects are weaker for visual than for auditory presentation (Frankish, 1985). The serial position curves also show that recall performance approached floor levels at some positions, particularly when there was a concurrent task. However, the Grouping  $\times$  Concurrent Task interaction seems unlikely to be a simple floor effect, as it was clearly evident over early serial positions, where performance in all conditions was relatively good.

## Discussion

Articulatory suppression removed the bulk of the recall advantage for temporally grouped lists, and there was no additional effect of reciting random digits. This suggests that the effect of temporal grouping on memory for visual lists depends on the phonological loop, with no major involvement of the central workspace, in contrast to the position taken by

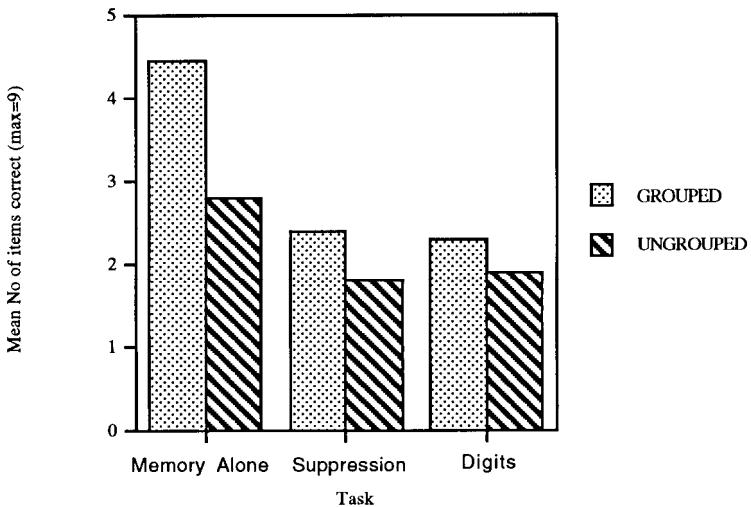


FIG. 1. Effects of concurrent tasks on recall of temporally grouped and ungrouped lists.

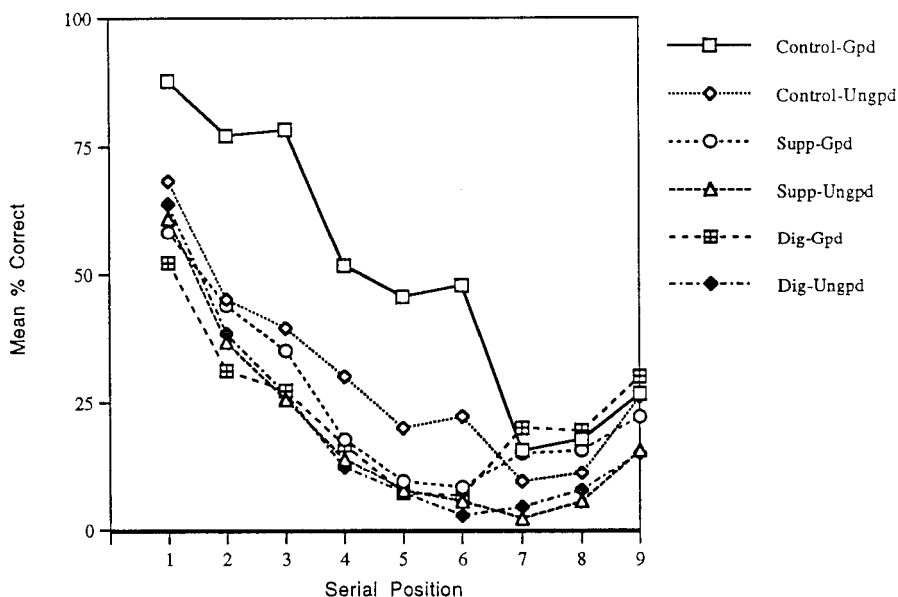


FIG. 2. Serial position curves for Experiment 1.

Broadbent and Broadbent (1981). Note, however, that the two sets of data are not necessarily incompatible, as the phonological loop could conceivably be involved in the improvement in recall due to grouping, but not in the encoding of grouped structure. Notice also that the present results do not argue strongly against any contribution from the central executive to the grouping effect, as reciting random digits did not have the expected effect of disrupting memory for ungrouped lists more than articulatory suppression.

In conclusion, the effects of concurrent tasks implicate the phonological loop in the temporal grouping effect and argue against a major contribution from the central executive. However, the results provide little clue as to how the phonological loop is involved. Experiment 2 was designed to explore this.

## EXPERIMENT 2

According to Baddeley's (1986) account, the phonological loop can be fractionated into a passive phonological store and a control process of subvocal rehearsal. Memory traces in the phonological store decay rapidly but can be refreshed by rehearsal. Thus, the longer words take to rehearse, the fewer can be refreshed in the time available before trace decay has gone too far. This analysis is supported by the effect of word length on immediate recall, which is systematically related to the time taken to articulate the items (Baddeley, Thomson & Buchanan, 1975). The evidence indicates that the capacity of the phonological loop is not a fixed number of items, but is time-based, corresponding to approximately 2 sec of inner speech. Further studies show that the word length effect is

abolished by articulatory suppression, consistent with it depending on subvocal rehearsal (Baddeley et al., 1975; Baddeley, Lewis, & Vallar, 1984).

If the temporal grouping effect is mediated by subvocal rehearsal, it is reasonable to assume that the optimal group size is not a fixed number of items (Ryan, 1969a) but is, instead, a fixed time corresponding to the capacity of the rehearsal loop. If so, the optimal group size should be numerically smaller for lists of longer words. Relatedly, the benefits of temporal grouping should be reduced if word length is increased such that subgroups are too long to be held within the loop. Experiment 2 explored these predictions by varying both the word length of the list items and the size of the temporal groups.

## Method

### Subjects

Forty-eight volunteer students from the University of Lancaster served as subjects and were each paid £3; 16 subjects were allocated to each of the 3 grouping conditions.

### Design and Materials

The design was a two-way mixed factorial. The between-subjects factor was temporal grouping with three levels: no grouping, 2 groups of 3 words ( $2 \times 3$ ), and 3 groups of 2 words ( $3 \times 2$ ). The within-subjects factor was word length, which had three levels (1, 3, and 5 syllable words).

Stimuli were 6-item lists constructed from pools of 1, 3, or 5-syllable words used by Baddeley et al. (1975, Expt. 6). These pools comprised one word from each of 10 semantic categories matched for word frequency (e.g. *stoat*, *gorilla*, *hippopotamus*, *mumps*, *leprosy*, *tuberculosis*). Ten lists of 6 words were produced by random sampling without replacement from each pool. The resulting 30 lists were then put into two separate random orders. Half the subjects completed one order, half the other. A Macintosh LCII computer controlled by Superlab software was used to present the lists.

### Procedure

Subjects were tested individually. Written instructions requested spoken recall of each sequence and emphasized the importance of recalling the words in the same serial order as presentation. Subjects were encouraged to guess but were allowed to omit answers. Each word was presented at the centre of the screen, in upper case, for 2 sec. In the ungrouped condition, each word was followed immediately by the next word. In the two grouped presentation conditions, there was an ISI of 1 sec between groups (i.e. after the third word in the  $2 \times 3$  condition and after the second and fourth words in the  $3 \times 2$  condition). Immediately following the sixth word in each list, a question mark appeared as the cue for recall. The experimenter marked down the subject's recall on a prepared answer sheet. Subjects completed the 30 lists of 6 words in one session lasting about 30 min.

## Results

Figure 3 shows the mean number of items recalled for each word length and each grouping condition. A two-way ANOVA confirmed that there was a significant main effect of word length,  $F(2, 42) = 218.3$ ,  $p < 0.01$ , such that recall declined systematically as word length increased. There was also a significant main effect of grouping,  $F(2, 21) =$

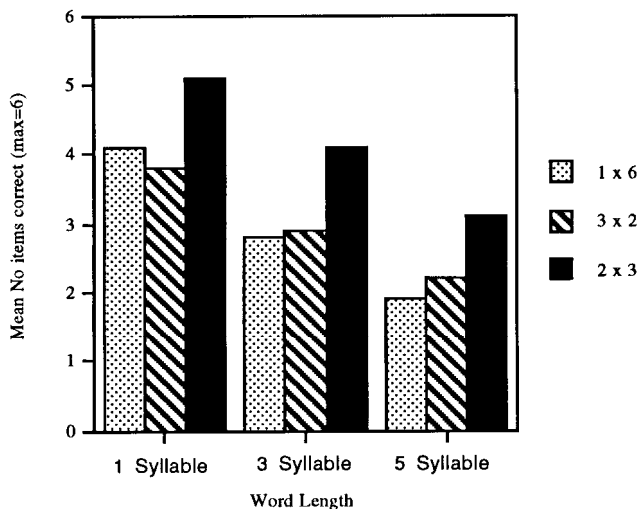


FIG. 3. Effect of different patterns of temporal grouping as a function of word length.

22.5,  $p < 0.01$ . Newman-Keuls tests indicated that grouping by threes led to significantly higher recall than did the other two conditions, and that grouping by twos was no better than no grouping at all. The Grouping  $\times$  Word Length interaction was non-significant,  $F(4, 42) = 2.42$ , though there was a slight tendency for the benefit of grouping to be bigger for lists of longer words.

## Discussion

The results were clear, though not as predicted by the hypothesis that grouping affects subvocal rehearsal. Word length had its anticipated effect on the overall level of recall, but the size of the grouping effect was not reduced and was even slightly bigger for the longest words. The major determinant of the grouping effect was evidently the distribution of pauses across items, independently of item length. Thus, for all word lengths, grouping by threes improved recall, whereas grouping by twos had no detectable effect. Given that the number of groups was not constant across conditions, there must be some caution in interpreting this result solely in terms of optimal group size. It is nevertheless a striking observation.

It might be argued that word length was not varied over a large-enough range to provide a strong test of the phonological loop hypothesis. However, estimates based on previous data (Baddeley et al., 1975) suggest that groups of three 5-syllable words would take about 2.5 sec to articulate, compared with estimates of approximately 2 sec for the capacity of the phonological loop, whereas three single-syllable words would take only about 1.5 sec. A more subtle argument is that temporal grouping affects the organization of a rehearsal plan, an aspect of rehearsal that would not be expected to depend on word length (and which, incidentally, lies outside the standard model of the phonological loop). Strictly, therefore, the results refute only the specific hypothesis that temporal grouping effects reflect the limited capacity of the rehearsal loop.



## EXPERIMENT 3

It is interesting to note that disrupting subvocal rehearsal is only one way in which articulatory suppression might have reduced the temporal grouping effect in Experiment 1. Suppression also disrupts the process of verbally recoding visual stimuli in order to access the phonological store, as indicated by its ability to abolish effects of the phonemic similarity of items on recall (Baddeley et al., 1984). However, suppression does not do this for auditory lists (Baddeley et al., 1984; Murray, 1968), consistent with evidence that auditory inputs access the phonological store automatically (Salamé & Baddeley, 1982). If, therefore, the temporal grouping effect is dependent only on information from the list accessing the phonological store and not on it being rehearsed, then the effect should remain under articulatory suppression for auditory lists. If, on the other hand, the conclusion drawn from Experiment 2 is incorrect and the grouping effect does depend on subvocal rehearsal, it should be removed by suppression for both auditory and visual lists. Experiment 3 set out to test these predictions.

In a previous study, Frick (1989) reported improved recall of temporally grouped auditory lists even when subjects suppressed articulation. However, this is not convincing evidence as suppression was only required during list presentation, and in order to block rehearsal of auditory lists effectively it is necessary to require suppression during both presentation and recall (Baddeley et al., 1984). Experiment 3 therefore required suppression throughout presentation and recall of auditory lists. The experiment also included a visual presentation condition, in order to replicate the observation that suppression removes the temporal grouping effect for visual lists.

## Method

### Subjects

Thirty-two students from the University of Lancaster served as subjects, and they were each paid £2; 16 subjects were allocated to each of the two grouping conditions.

### Design and Materials

A three-way factorial design was used with factors grouping  $\times$  articulation  $\times$  presentation modality. Temporal grouping (no grouping vs. grouping by threes) was manipulated as a between-subjects factor. All subjects served in the two concurrent articulation conditions (no suppression vs. suppression), for each of the two presentation modalities (auditory vs. visual).

Four sets of 10 lists each of 9 letters were constructed by random sampling without replacement from a pool of 12 consonants with low phonemic confusability (*C D F H J L N Q R S Y Z*). Any meaningful clusters of letters were replaced by further random sampling. Visual lists were presented on a Macintosh II screen controlled by Psychlab, as in Experiment 1. Auditory presentation involved playing a tape-recording of the experimenter reading out grouped or ungrouped lists from the visual display. During recording, an effort was made to ensure a flat intonation and to keep strictly to the timing of the display.

## Procedure

Subjects were tested individually. Written instructions emphasized the importance of recalling items in their correct serial positions, but due to an oversight did not emphasize that recall was to be in serial order. At the start of each trial a *Press space bar to begin* command was given on the computer screen for visual lists and a spoken “Ready?” command was given for auditory lists. Immediately following the final letter of visual lists, subjects were cued to begin recall by presenting three dots centre-screen. For auditory lists there was no cue. Subjects recorded their answers on a response sheet.

Subjects completed four blocks of 10 trials, one for each presentation modality (visual vs. auditory) and concurrent task (no suppression vs. suppression). The order of blocks and the assignment of lists to conditions was counterbalanced across subjects. Articulatory suppression was achieved by asking subjects to repeat the word “the” at an even pace (approximately 2 words per sec) at the start of each trial and to continue throughout list presentation and recall. Testing sessions lasted approximately 30 min.

## Results

Figures 4 and 5 illustrate the mean number of items recalled in the various experimental conditions. Considering auditory lists first, recall was clearly improved by grouping and impaired by suppression (Figure 4). However, suppression did not affect the tendency for grouping to enhance recall. A two-way (Grouping  $\times$  Suppression) ANOVA revealed significant main effects of grouping,  $F(1, 30) = 35.2, p < 0.01$ , and suppression,  $F(1, 30) = 99.0, p < 0.01$ , and no Grouping  $\times$  Suppression interaction,  $F < 1$ . The serial position curves for auditory lists are illustrated in Figures 6 and 7, from which it is clear that the multiple bowing typical of temporally grouped lists was present both with and without suppression.

For visual lists, recall was also enhanced by grouping and impaired by suppression (see Figure 5). However, the grouping effect appeared to be rather weaker than in Experiment 1. A two-way ANOVA indicated significant main effects of grouping,  $F(1, 30) = 5.74, p <$

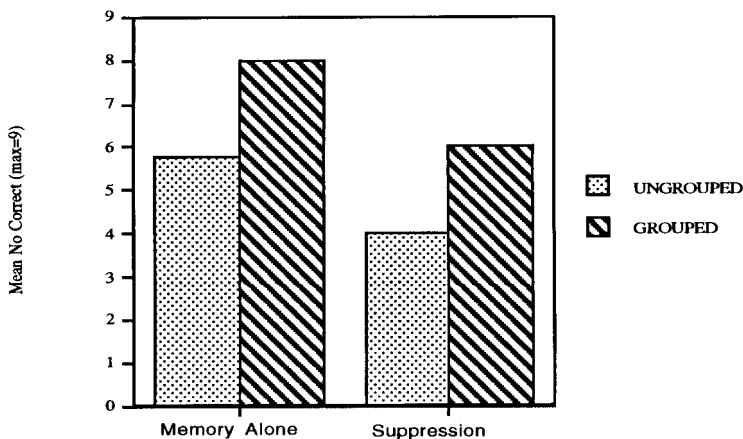


FIG. 4. Effects of articulatory suppression and temporal grouping for auditory lists.

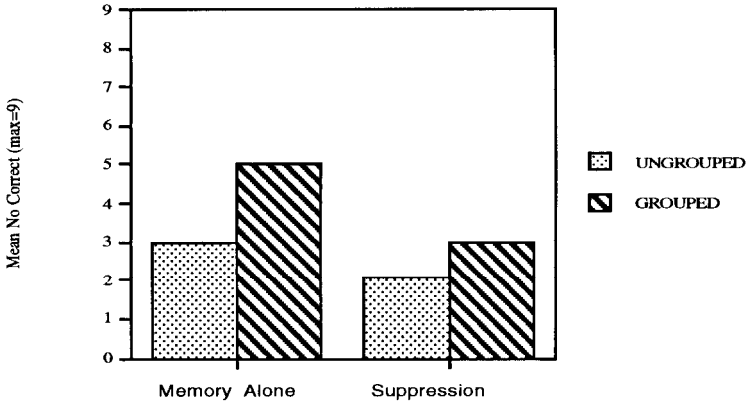


FIG. 5. Effects of articulatory suppression and temporal grouping for visual lists.

0.05 and suppression,  $F(1, 30) = 53.3$ ,  $p < 0.01$ , and a significant interaction,  $F(1, 30) = 4.28$ ,  $p < 0.05$ . The serial position curves (see Figures 8 and 9) show signs of multiple bowing in the control condition, and some suggestion of bowing for the first triplet under articulatory suppression.

## Discussion

For auditory lists, articulatory suppression did not bring about any significant reduction in the effect of temporal grouping, nor did it remove the multiple bowing of the serial position curve that is characteristic of encoding by groups. Given that suppression throughout presentation and recall disrupts rehearsal of auditory lists (Baddeley et al., 1984), these results are further evidence against a rehearsal-based interpretation of the temporal grouping effect. As suppression does not prevent auditory items being phonemically encoded, the results are consistent with the involvement of the phonological store in the grouping effect. However, it should be noted that the results are equally consistent with the hypothesis that perceptual storage is responsible for the auditory grouping effect (Frankish, 1985, 1989).

For visual lists, the effect of grouping was not particularly strong, possibly because overall performance was rather low. Articulatory suppression reduced the temporal grouping effect, as in Experiment 1, but there are once again problems of floor effects at some serial positions. Looking just at the first three serial positions, where performance was highest, confirms that suppression did, indeed, reduce the grouping effect but suggests that the effect may not have been completely abolished. This implies that although the influence of temporal grouping on visual lists has a phonological component, it is premature to rule out a non-phonological contribution (see also discussion of Experiment 1).

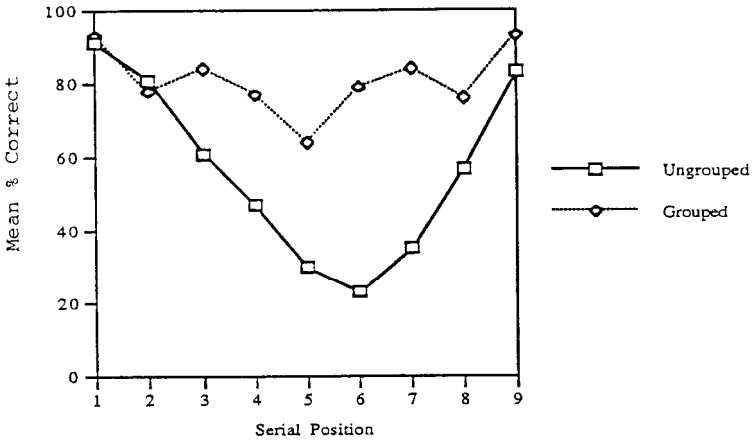


FIG. 6. Serial position curves for Experiment 3. Auditory presentation, memory alone.

### EXPERIMENT 4

The final experiment explored the role of the phonological store further by manipulating the phonemic similarity of the items. It is well known that immediate memory is impaired by phonemic similarity (Conrad, 1964), and, according to the working memory model, this is because recall involves recovering item information from memory traces in the phonological store. It is assumed that due to noise in the system, the partially decayed traces of phonemically similar items are less discriminable from one another and hence are more difficult to recover from the store (Baddeley, 1986). However, predicting how phonemic similarity will influence the effect of temporal grouping is not straightfor-

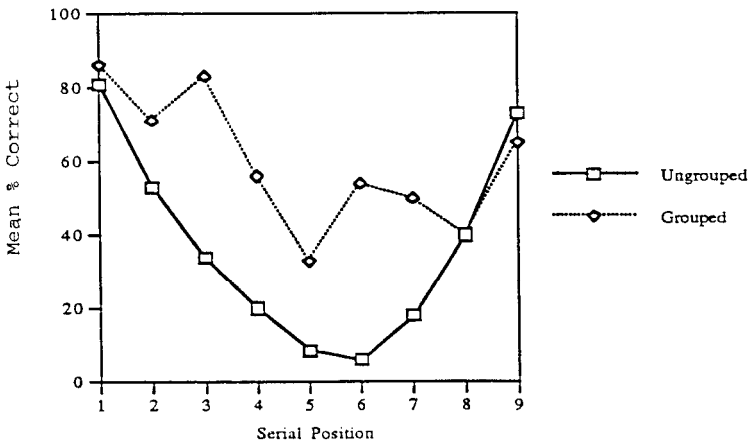


FIG. 7. Serial position curves for Experiment 3. Auditory presentation, suppression.

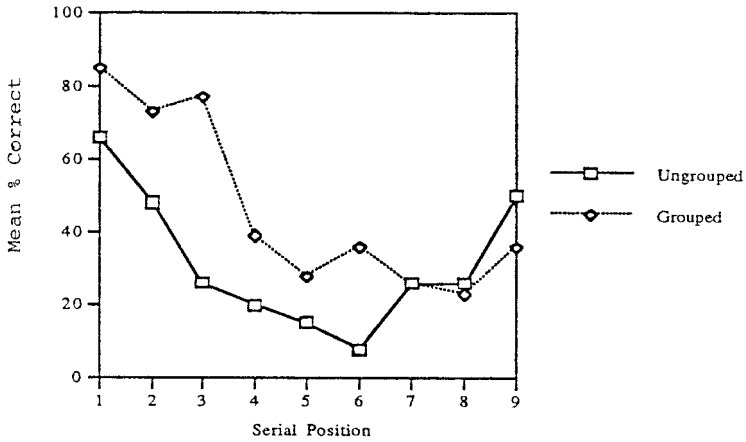


FIG. 8. Serial position curves for Experiment 3. Visual presentation, memory alone.

ward. If grouping enhances the discriminability of items within the phonological store, it should lead to a greater improvement in recall for phonemically similar items. But if subject tend to abandon the phonological store when lists contain phonemically similar items, recall of such lists would benefit less from grouping. Thus, different predictions can be made depending on the precise assumptions made.

## Method

### Subjects

There were 26 subjects, the majority of whom were female. They were either members of the Royal Holloway College subject panel or undergraduate students and were paid £2; 13 subjects were randomly allocated to each grouping condition.

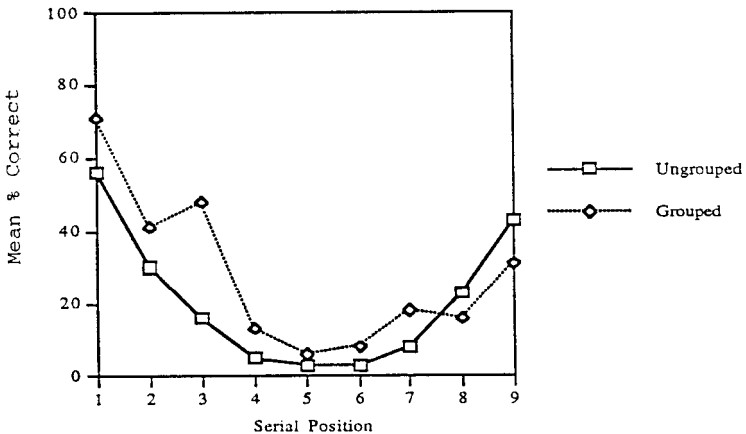


FIG. 9. Serial position curves for Experiment 3. Visual presentation, suppression.

## Design and Materials

The design was a two-way mixed factorial with temporal grouping (no grouping vs. grouping by threes) as the between-subjects factor and phonemic similarity (similar vs. dissimilar words) varied within subjects.

Word lists were constructed from pools of phonemically similar (e.g. *can, cad*) and dissimilar (e.g. *cow, day*) words matched for frequency, as used by Baddeley (1966). Random sampling without replacement was used to generate 15 lists of 6 words from each pool. Obvious two- or three-word collocations were avoided by resampling. The resulting 30 lists were put into two separate random orders with the similar and dissimilar lists intermingled. Half the subjects completed one order, half the other.

Word lists were presented sequentially on a monitor controlled by an Archimedes 310 micro-computer. Each word appeared centre-screen for 700 msec. In the ungrouped condition, the ISI was 200 msec for each item. In the grouped condition, the ISI between the third and fourth item was extended to 1200 msec. Immediately following the sixth item, three dots appeared centre-screen as a recall cue.

## Procedure

Subjects were tested individually or in pairs. Instructions indicated that there would be six words followed by a cue for written recall of the words in the correct serial order using the answer sheet provided. Subjects were given three practice lists followed by 30 experimental lists. The experiment took approximately 15 min to complete.

## Results

Figure 10 shows the mean numbers of phonemically similar and dissimilar words recalled for grouped and ungrouped lists. A two-way ANOVA, with similarity and grouping as factors, showed that more words were remembered when presentation was grouped,

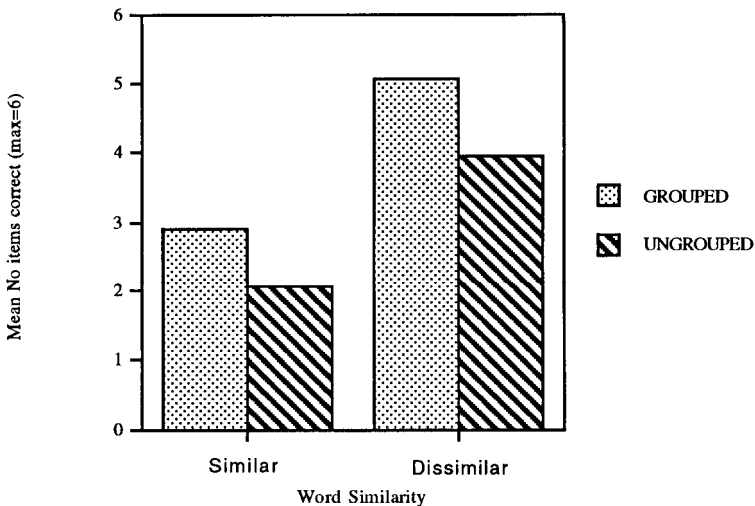


FIG. 10. Effect of temporal grouping as a function of phonemic similarity of the items.

$F(1, 24) = 6.33, p < 0.05$ . There was also a large effect of similarity,  $F(1, 24) = 124.0, p < 0.01$ , with highly similar words much harder to recall. There was no Grouping  $\times$  Similarity interaction,  $F(1, 24) = 1.47$ . Figure 11 illustrates the serial position curves, which show modest scalloping in the grouped condition for both similar and dissimilar lists.

## Discussion

The lack of any significant alteration in the temporal grouping effect for phonemically similar items suggests that grouping does not affect item discriminability in the same way as does phonemic similarity. This result also goes against the idea that phonemic similarity of the materials encourages subjects to rely on non-phonological forms of storage that are insensitive to grouping. The general conclusion, therefore, seems to be that temporal grouping and phonemic similarity affect separate processes. However, given that this conclusion seems to be at odds with the idea that grouping affects the phonological store, it is appropriate at this point to review the evidence from all four experiments.

The suggestion that the temporal grouping effect is dependent on material entering the phonological store came from the finding that articulatory suppression interacted with the grouping of visual lists (Experiments 1 and 3). Other evidence suggested that the grouping effect is nevertheless independent of phonological characteristics of the items such as their similarity (Experiment 4) or length (Experiment 2). This is a puzzling pattern of results for the working memory model given its assumption that articulatory suppression, phonemic similarity, and word length are converging operations for the phonological loop. Even considering the fractionation of the loop into a phonological store and a subvocal rehearsal process fails to suggest an obvious explanation. However, the problem seems to be not with the general concept of the phonological loop, for which there is much independent support (see e.g. Baddeley & Hitch, 1994), but

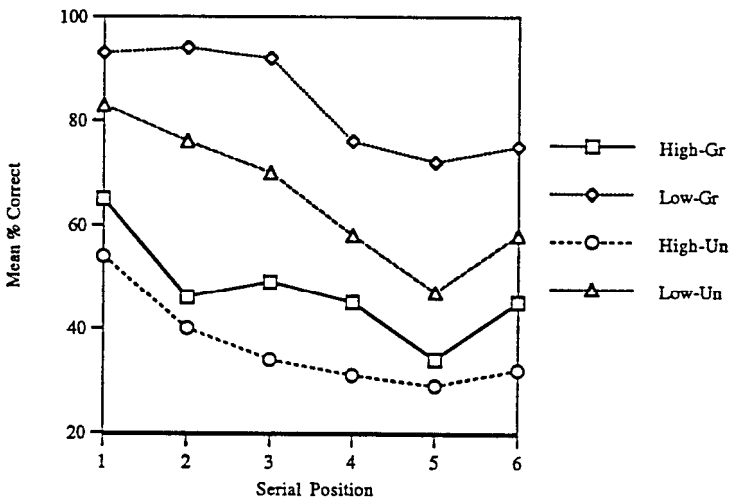


FIG. 11. Serial position curves for Experiment 4.

with the fact that the concept is not sufficiently well specified. This was particularly evident in Experiment 4 with regard to the difficulty of making a firm prediction for the effect of phonemic similarity on grouping. We therefore considered whether a more detailed, computational model of the phonological loop (Burgess, 1995; Burgess & Hitch, 1992, in press) could lead to a clearer theoretical interpretation of the present results. One advantage of this type of model is explicitness, which allows predictions to be derived by simulation.

## A CONNECTIONIST INTERPRETATION OF TEMPORAL GROUPING

Burgess (1995, see also Burgess & Hitch, in press) describes a simple extension of an earlier connectionist model of the phonological loop (Burgess & Hitch, 1992). The extended version maintains the core attributes of the previous model and accounts for a wide range of phenomena, such as the effects of word length and phonemic similarity, patterns of order errors, the shape of the serial position curve, effects of list length, and influences from previously learned lists such as position-specific intrusions and the Hebb effect.

The extended model involves three types of information: (1) a representation of items, (2) a representation of phonemes, and (3) a repeatable timing signal referred to as “context”. These are modelled by three layers of nodes linked by Hebbian connections (see Figure 12). Presentation of an item generates activity in the phoneme layer, which feeds through to produce activity in the item layer. The item node with the biggest input is selected, and temporary associations are then formed between this selected item and currently active nodes in the context layer. The context signal corresponds to a “moving window” such that each node is active for a short time period, creating an overlap between the contexts for adjacent items (see Figure 13A). This is the basic mechanism for encoding serial order information in the model. During recall, context representations are reproduced in their original order, feeding activity back into the item layer via the temporary associations formed during presentation. At each step, the most active item node is selected using a process analogous to “competitive queuing” (Houghton, 1990) and feeds activity into its constituent phoneme nodes for output. Finally, the selected item is given an inhibitory input that decays with time, preventing consecutive recall of the same item.

Recall errors arise from two sources: (1) decay of context–item associations over time, and (2) the impact of noise in the activation of item nodes. These influences lead to a predominance of order errors involving proximate items and phonemic confusions. As shown previously (Burgess & Hitch, 1992), the model exhibits limited span and shows word-length effects because context–item associations have more time to decay for longer words. There is a simple mapping onto the phonological loop idea in that the phoneme layer corresponds to the phonological store and the selection mechanism corresponds to articulation. An important difference, however, is that the network model incorporates an additional component within the system—the context signal.

Within the model, the occurrence of a pause during list presentation could be treated in three possible ways: (1) as presentation of a silent item—that is, like the presentation of



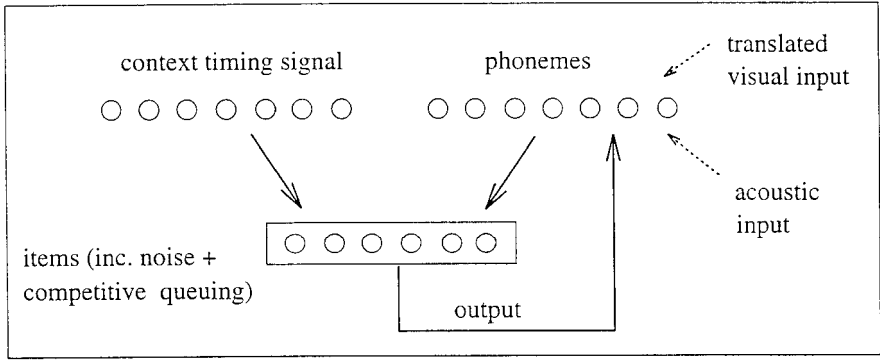


FIG. 12. The architecture of the model, adapted from Burgess (1995), which is a slight modification of that in Burgess & Hitch (1992). Full lines are connections with short and long term plasticity; dashed lines are routes by which information enters the model.

any other item, except that the silent item is associated with a “phoneme node” representing a short period of silence; (2) as activation of a start-of-group or end-of-group marker in the phonological representations of the items adjacent to the pause; or (3) as disturbing the context signal. In deciding among these alternatives, we considered two effects: (a) temporal grouping typically improves overall recall performance; and (b) there is at the same time a relative increase in the likelihood of errors in which an item from one group is incorrectly recalled at the corresponding position within another group (Ryan, 1969b).

If implemented in our model in the above way, option (1) would cause effect (a) by making the context signals for the items on either side of the pause less similar. However, it would not cause effect (b), as the model contains no item–item associations. Option (2) would have an effect similar to increasing the phonemic similarity among the first and last items in different groups, so helping to cause effect (b). However, it would only cause effect (a) in so far as it led to a reduction in the similarity of the phonemic representations

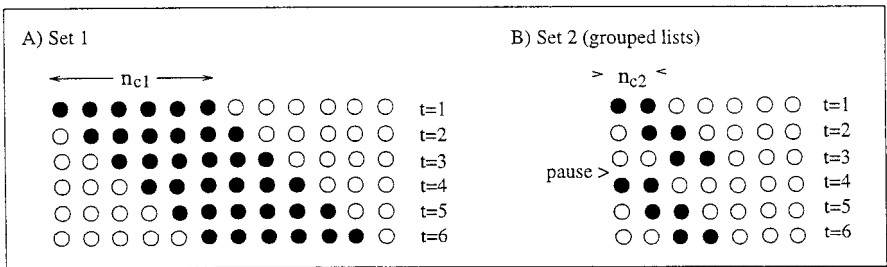


FIG. 13. The “context” timing signal: filled circles are active nodes, empty circles are inactive nodes, the index  $t$  represents serial position. (A) The usual representation for ungrouped lists, a moving window of activation supported by a set of  $n_{c1}$  temporal oscillators (set 1) that are reset at the start of recall (adapted from Burgess, 1995; see also Burgess & Hitch, 1992). (B) The second set of context nodes, supported by a set of  $n_{c2}$  temporal oscillators (set 2) that are reset after each pause, reflecting the presence of a pause between  $t = 3$  and  $t = 4$ .

of items at different group positions. It therefore predicts no increase in overall performance when lists of phonemically dissimilar items are used, in contrast to the human data, which show no interaction between similarity and grouping effects. We preferred to explore option (3) for the following reasons: The moving window of the context signal is triggered by the start of presentation, and it seems natural to view a pause as having a similar, though weaker effect to that of starting a new list. Further, the context signal varies with serial position—that is, we are already interpreting it as reflecting the rhythm of presentation. Notice also that the context signal functions in a way similar to attaching serial position “tags” to each item, and the idea that temporal grouping may modify such tags has been postulated previously (McNicol & Heathcote, 1986).

By taking option (3), we would not expect to see a strong interaction between temporal grouping and phonemic similarity, because the model’s context and phonemic components supply activation to item nodes independently of one another. Nor would we expect a strong interaction between grouping and word length, because we are not modelling grouping as an effect of rehearsal. However, these expectations must be validated, and the success of the model judged against more detailed aspects of temporal grouping, such as the effect of group size on performance, the “scalped” serial position curves, and the relative increase in confusions between items with the same within-group position.

## A Simple Model for Temporal Grouping

There are a number of ways in which temporal grouping could affect the context signal in our model. According to Treisman, Cook, Naish, and McCrone (1994), there is a central timing system, which is underpinned by sets of oscillators, each set varying around a modal frequency (see also Church & Broadbent, 1990, 1992). Frequency variability allows entrainment of the oscillators to external rhythms, as has been shown in effects on time estimation and the EEG (Treisman et al., 1994). We suggest that a central timing mechanism of this sort furnishes the context signal in our model, and we explain temporal grouping effects by assuming that the context signal becomes entrained to the inter-group frequency. This proposal has the dual appeal that it links serial ordering in short-term memory with timing in general, and that it specifies a precise mechanism for the temporal grouping effect. It is also directly related to the idea that temporal grouping modifies serial order tags (McNicol & Heathcote, 1986). According to this idea, tags for ungrouped lists are represented by points along a line, whereas tags for grouped lists are represented by points on a plane whose second dimension reflects within-group position. In an analogous way, oscillators repeating at the start of the list can be regarded as providing a signal along the first dimension, and oscillators entrained by temporal grouping provide a signal along the second dimension.

In terms of implementation, the context nodes can be thought of as a set of  $N$  thresholded oscillators, all with the same period  $T$  but with each lagging the node on its left by the time taken to present or recall one item. A node is “active” whenever the oscillation it represents exceeds a threshold  $\Theta$ , and at any one time  $n_{cl}$  of these nodes are active. The start of presentation (and the start of recall) sets (and resets) the initial phase of the oscillators such that the leftmost  $n_{cl}$  nodes are the active ones. The “window” of active nodes then shifts one place to the right as the phase of each oscillator advances with

the presentation of each successive item. Note that the threshold  $\Theta$  must be large enough for only a small proportion of nodes to be active (i.e.  $n_{c1} \ll N$ ) and that only oscillators with a period greater than the duration of the list are used, so that nodes active at the start of presentation will not be reactivated towards the end of presentation.

In the event of temporal grouping, we suppose that a second set of oscillators (of which  $n_{c2}$  are active at any one time) become entrained such that they reset after every pause and at the start of recall (see Fig. 13B). The product of the two sets of timing signals is then used to activate item nodes during recall. Thus, at time  $t$  in recall, each item node receives activation from the context nodes proportional to the fraction of the nodes active during its presentation that are again active. The fraction of nodes in Set 1 that are active during the presentation of item  $i$  and are also active at time  $t$  in recall is called  $F^1_{i}(t)$  and  $F^2_{i}(t)$  is defined similarly for the nodes in Set 2.<sup>1</sup> The activation that reaches item  $i$  at time  $t$  from the context timing signal is proportional to  $F^1_{i}(t) \times F^2_{i}(t)$ , and the constant of proportionality is the factor by which the connection weights have decayed since presentation.

## Performance, Discussion, and Predictions

We expect primacy and recency within each group for the same reason as primacy and recency in the whole list (namely, that fewer items are associated to similar states of the context signal with which to be confused). We also expect temporal grouping to produce an increase in overall performance and an increase in the relative number of confusions between items with the same within-group position, because the similarity between context states for items with the same within-group position is unchanged but is decreased between all other items. The model's performance was investigated in greater detail by numerical simulation. Its main successes are as follows.

Grouping a 6-item list in threes produces modestly scalloped serial position curves, and the effect of grouping is relatively independent of the phonemic similarity of the items (see Figure 14). Also, though not shown here, the grouping effect results from a reduction in order errors, consistent with experimental evidence (Ryan, 1969a). Figure 15 shows that recall of 6-item lists is best when they are grouped in threes rather than grouped in twos or ungrouped, and the size of the grouping effect is largely independent of the word length of the items. Analysis of errors shows that there is an increased probability of confusion between items that have the same within-group position relative to those that do not. For example, in a simulation of the recall of lists of 6 items grouped in threes, 33.8% of order errors were items recalled in the wrong group but the same within-group position. In comparison, the percentage of order errors involving corresponding serial positions in ungrouped lists was 7.5%.

---

<sup>1</sup> For Set 1, this fraction is 1 for the current item [i.e.  $F^1_{i}(t) = 1.0$ ] and is successively less for neighbouring items [i.e.  $F^1_{i+1}(t) = (n_{c1} - 1)/n_{c1}$ , etc.]. For Set 2, the fraction is 1 for the current item and for all other items with the same within-group position [i.e.  $F^2_{i}(t) = F^2_{i+g}(t) = 1.0$ , etc, where  $g$  is the number of items per group] and is successively less for items with different within-group positions [i.e.  $F^2_{i+1}(t) = F^2_{i+g+1}(t) = (n_{c2} - 1)/n_{c2}$ , etc.].

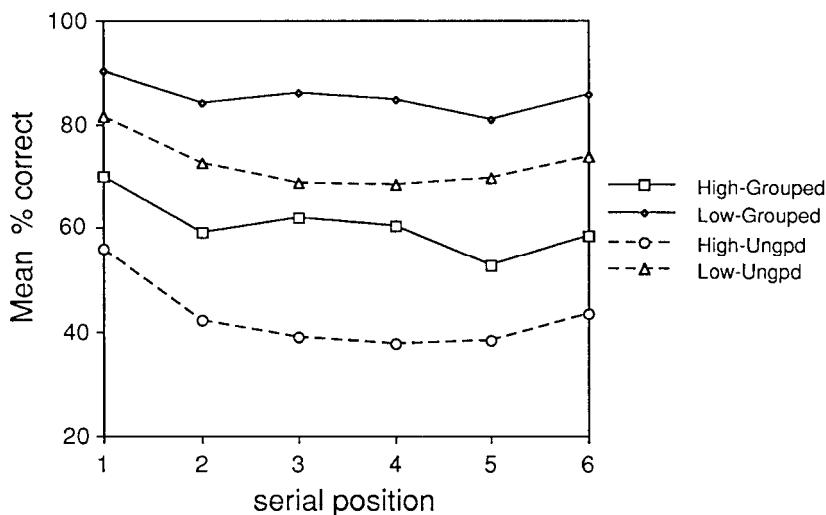


FIG. 14. Simulated recall performance of 6-item lists as a function of temporal grouping and phonemic similarity. Items were composed of 3 phonemes; there were either 0 or 2 phonemes common to all of the items in a list, labelled by “dissimilar” and “similar” respectively. Items were either ungrouped or grouped in threes. All simulations are the average of 5000 trials, each phoneme has duration 0.2 sec,  $n_{c1} = 6$ ,  $n_{c2} = 2$ , all other parameters are as in Burgess (1995).

One limitation of the model is that it gives a grouping effect for lists grouped in twos, which disagrees with the results from Experiment 2. However, in recent unpublished experiments we have found benefits of grouping by twos, so it remains to be seen whether this is a substantive problem. A more definite limitation is that we have not modelled the effects of presentation modality, which include extra recency at the end of each group and a greater enhancement of performance for auditory relative to visual presentation (Frankish, 1985, 1989; see also Experiment 3). It is intuitively clear why pauses might provide a stronger entraining stimulus during auditory presentation than during visual presentation, as the auditory system is highly tuned to detect temporal variations in its input. However, it remains to be seen how this would be modelled. Another question is how the encoding of a grouped list could change before the first pause occurs, i.e. before it is clear that the list will be grouped. In the experiments in this paper the subjects know when lists will be grouped, as grouped and ungrouped trials occur in different blocks. One prediction of the model is that recall of the first item in the first group will not be enhanced if subjects do not know in advance that the list will be grouped. Another, more striking prediction is that the context timing signal ought to be entrainable to a suitable external rhythm (Treisman et al., 1994), causing interference with short-term memory for serial order and impairing our ability to take advantage of temporal grouping to improve recall.

In summary, we have previously proposed that there is a third functional sub-component of the phonological loop in addition to the phonological store and the process of subvocal rehearsal: a repeating context signal that reflects the rhythm of the items' presentation (Burgess, 1995; Burgess & Hitch, 1992, in press). Here we further propose

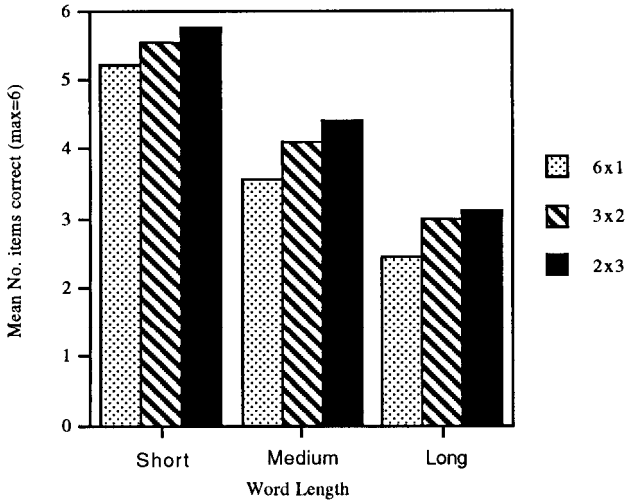


FIG. 15. Simulated recall performance as a function of temporal grouping and word length. Items were composed of 2, 4, or 6 dissimilar phonemes and were either ungrouped ( $6 \times 1$ ), grouped in twos ( $3 \times 2$ ), or grouped in threes ( $2 \times 3$ ).

(1) that this signal is generated by a central timing system of temporal oscillators, and (2) that the phenomenon of temporal grouping is caused by modification of the timing signal, via entrainment of oscillators to the inter-group frequency. Simulations embodying these proposals reproduced some of the main experimental findings, principally a temporal grouping effect that is insensitive to either phonemic similarity or word length of the items.

## GENERAL DISCUSSION

This series of experiments began by exploring whether a central buffer store mediates the improvement in short-term memory brought about by the temporal grouping of visually presented verbal sequences (Broadbent & Broadbent, 1981). Although a concurrent task designed to load central and phonological components of working memory disrupted the grouping effect, so too did articulatory suppression, a task that is assumed to interfere selectively with the phonological subsystem of working memory (Baddeley, 1986; Baddeley & Hitch, 1974). These results therefore implicate the phonological loop and not a central buffer in the temporal grouping effect for visual lists.

The results also provided evidence that, contrary to earlier suggestions (Ryan, 1969b), the temporal grouping effect is not critically dependent on rehearsal. Thus the effect of temporal grouping was insensitive to word length (Experiment 2) and persisted despite articulatory suppression when lists were presented auditorily (Experiment 3). Experiment 2 also suggested that the optimal group size does not reflect the capacity of the phonological loop.

However, while the working memory model clearly provided a useful framework for guiding the investigation of temporal grouping, the concept of the phonological loop was insufficiently well specified to provide a satisfactory account of all the data. This was most apparent with respect to the pattern of interactions between temporal grouping and manipulations assumed to affect the loop. However, these effects could be addressed by a simple adaptation of a connectionist model, which proposes that, in addition to the phonological store and the process of articulation, the phonological loop involves a context signal (Burgess, 1995; Burgess & Hitch, 1992, in press). By assuming that the context signal involves a set of temporal oscillators that reflect the rhythm of list presentation, the model reproduced many aspects of the experimental data on temporal grouping, including the scalloping of the serial position curve and the absence of major interactions involving temporal grouping and the phonological characteristics of the items. An essential feature of the model is that grouping affects the organization of information at the item rather than the phonological level, and it is interesting to note that this may have some overlap with the idea of distinguishing a rehearsal plan from subvocalization (see Discussion of Experiment 2). As stated earlier, some such overlap is suggested by the parallels between the effects of temporal grouping and instructions to rehearse ungrouped lists in groups (Wickelgren, 1964, 1967).

As should be clear, the present interpretation of grouping is post hoc and requires further testing, both in terms of its novel predictions and its underlying assumptions. A major issue is to test the proposal that immediate memory for serial order is mediated by a system of temporal oscillators. A related issue is whether the proposed context signal is unique to tasks involving phonological short-term memory or is, indeed, general. The assumption that it is general may offer some insight into why suppression interacts with the grouping effect but word length and phonemic similarity do not. Thus, given that suppression requires output at a regular rate, it may be controlled by the same timing system as the memory task. This does not immediately explain why suppression affects temporal grouping with visual but not auditory presentation. However, as argued earlier, the inherently temporal nature of the auditory modality may have something to do with this asymmetry. This could be tested experimentally by looking for differences in the timing of suppression responses for visual and auditory grouped lists.

Finally, and more generally, the present model suggests the possibility that the chunking of a sequence into familiar subsequences may involve a similar effect on the context signal to that caused by temporal grouping. However, any such similarity should be viewed in the context that chunking phenomena clearly involve much more complex processes than are involved in temporal grouping.

## REFERENCES

- 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, 249–264.
- Baddeley, A.D. (1986). *Working memory*. Oxford: Oxford University Press.
- Baddeley, A.D., & Hitch, G.J. (1974) Working memory. In G. Bower (Ed.), *The psychology of learning and motivation*, Vol. 8 (pp. 47–90). New York: Academic Press.

- Baddeley, A.D., & Hitch, G.J. (1994). Developments in the concept of working memory. *Neuropsychology*, 8, 485–493.
- Baddeley, A.D., Lewis, V.J. & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology*, 36, 233–252.
- Baddeley, A.D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 14, 575–589.
- Bower, G.H. (1972). Organizational factors in memory. *Cognitive Psychology*, 1, 18–46.
- Bower, G.H., & Winzenz, D. (1969). Group structure, coding and memory for digit series. *Journal of Experimental Psychology Monograph*, 80 (No. 2, Part 2).
- Broadbent, D.E. (1975). The magic number seven after fifteen years. In A. Kennedy & A. Wilkes (Eds.), *Studies in long-term memory* (pp. 3–18). New York: Wiley.
- Broadbent, D.E. (1981). Association lecture: From the percept to the cognitive structure. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 1–24). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Broadbent, D.E. (1984). The Maltese Cross: A new simplistic model of memory. *Behavioral & Brain Sciences*, 7, 55–94.
- Broadbent, D.E., & Broadbent, M.H. (1981). Articulatory suppression and the grouping of successive stimuli. *Psychological Research*, 43, 57–67.
- Burgess, N. (1995). A solvable connectionist model of immediate recall of ordered lists. In: G. Tesauro, D. Touretzky, & T.K. Leen (Eds.), *Neural information processing systems, Vol. 7*, Cambridge, MA: MIT Press.
- Burgess, N., & Hitch, G.J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language*, 31, 429–460.
- Burgess, N., & Hitch, G.J. (in press). A connectionist model of STM for serial order. In S. Gathercole (Ed.), *Models of short-term memory*. Hove: Lawrence Erlbaum Associates, Ltd.
- Chase, W.G., & Ericsson, K.A. (1982). Skill and working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation, Vol. 16*, New York: Academic Press.
- Church, R.M., & Broadbent, H.A. (1990). Alternative representations of time, number and rate. *Cognition*, 37, 55–81.
- Church, R.M., & Broadbent, H.A. (1992). A connectionist model of timing. In: M.L. Commons, S. Grossberg & J.E.R. Staddon (Eds.), *Neural network models of conditioning and action* (pp. 225–240). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Conrad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology*, 55, 75–84.
- Frankish, C. (1985). Modality-specific grouping effects in short-term memory. *Journal of Memory and Language*, 24, 200–209.
- Frankish, C. (1989). Perceptual organization and precategorical acoustic storage. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 469–479.
- Frick, R.W. (1989). Explanations of grouping in immediate ordered recall. *Memory and Cognition*, 17, 551–562.
- Graesser, M., & Mandler, G. (1978). Limited processing capacity constrains the storage of unrelated sets of words and retrieval from natural categories. *Journal of Experimental Psychology: Human Learning & Memory*, 4, 86–100.
- Hitch, G.J., & Baddeley, A.D. (1976). Verbal reasoning and working memory. *Quarterly Journal of Experimental Psychology*, 28, 603–621.
- Houghton, G. (1990). The problem of serial order: a neural network model of sequence learning and recall. In: R. Dale, C. Mellish, & M. Zock (Eds.), *Current Research in Natural Language Generation* (pp. 287–319). London: Academic Press.
- McNicol, D., & Heathcote, A. (1986). Representation of order information: An analysis of grouping effects in short-term memory. *Journal of Experimental Psychology: General*, 115, 76–95.
- Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits of our capacity for processing information. *Psychological Review*, 63, 81–97.
- Ryan, J. (1969a). Grouping and short-term memory: Different means and patterns of grouping. *Quarterly Journal of Experimental Psychology*, 21, 137–147.

- Ryan, J. (1969b). Temporal grouping, rehearsal and short-term memory. *Quarterly Journal of Experimental Psychology*, 21, 148–155.
- Salamé, P., & Baddeley, A.D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior*, 21, 150–164.
- Simon, H.A. (1974). How big is a chunk? *Science*, 183, 482–488.
- Treisman, M., Cook, N., Naish, P.L.N., & McCrone, J.K. (1994). The internal clock: Electroencephalographic evidence for oscillatory processes underlying time perception. *Quarterly Journal of Experimental Psychology*, 47A, 241–289.
- Wickelgren, W.A. (1964). Size of rehearsal group and short-term memory. *Journal of Experimental Psychology*, 68, 413–419.
- Wickelgren, W.A. (1967). Rehearsal grouping and hierarchical organization of serial position cues in short-term memory. *Quarterly Journal of Experimental Psychology*, 19, 97–102.

*Manuscript received 31 October 1994*

*Accepted revision received 16 August 1995*