Chapter 3: Positional Theory

Two experiments testing Positional Theory

The previous chapter failed to find any evidence for chaining in serial recall, throwing doubt on models in which order is stored via item-item associations. The present chapter examines evidence for positional information in serial recall, in order to address models that assume some type of position-item association. In Experiment 2, this evidence comes from transpositions between groups that maintain their position within groups (interpositions). In Experiment 3, this evidence comes from intrusions between trials that maintain their position within trials (protrusions). Both experiments also continue to examine other errors in serial recall, such as omissions and intrusions, for the purposes of modelling in Chapter 5.

Grouping

Grouping a sequence into smaller subsequences improves retention of that sequence (Ryan, 1969a; 1969b; Wickelgren, 1967). Ryan (1969a), for example, presented lists of nine digits auditorily, which were grouped into three groups of three by a number of means: by a pause after every third digit, by a tone pip after every third digit, or simply by instruction. Grouping generally improved recall, compared to an ungrouped condition, though the advantage only proved significant when grouping was achieved by pauses (temporal grouping). Grouping of auditory stimuli can also be achieved by alternating the laterality or the voicing of groups, which is as effective as, but not additive with, temporal grouping (Frankish, 1989). In fact, even accenting the pitch of the last item in each group is sufficient to produce an advantage equal to that of temporal grouping (Frankish, 1995). Effects of temporal and spatial grouping are also found with visual presentation, though they are much smaller than with auditory presentation (Frankish, 1985).

Nevertheless, all methods of grouping have similar and striking effects on the distribution of errors. In serial position curves, these effects are sometimes revealed as mini-primacy and mini-recency effects within groups, resulting in "scalloped" shapes for each group (e.g., Figure 3-1). More generally, grouping reduces the number of transpositions

between groups, except those that maintain their position within groups (i.e., interpositions).

Ryan (1969b) found that presentation rate had little effect on the size of the temporal grouping effect. Frankish (1989) showed further that, providing the pause was perceptible, auditory temporal grouping was unaffected by pause length. Frick (1989) showed that a similar grouping effect remained under concurrent articulatory suppression, which prevents rehearsal (Baddeley, 1986). These results suggest that, though extra rehearsal during pauses may play a role, particularly in the visual modality (Frankish, 1989), differences between temporally grouped and ungrouped lists mainly reflect alternative representations in memory.

One alternative representation is the recoding of groups into single units or *chunks* (Miller, 1956/1994). The advantage of grouping is then to reduce the number of chunks that must be recalled, from nine in Ryan's ungrouped case, to three in her grouped case. However, as single units, chunks are recalled in an all-or-none fashion (Johnson, 1972). While this might explain the overall reduction in transpositions between groups, it has problems explaining why interpositions between groups remain. Most interpositions occur singly, and are not an artefact of whole groups swapping (Lee & Estes, 1981; Nairne, 1991). Moreover, chunking usually implies the preexistence of long-term memory codes, which is not the case for the novel groupings of items in the above experiments. Grouping is more likely to reflect a "reordering" rather than "recoding" of items (Frankish, 1974).[1]

Grouping is not necessarily contingent on objective organisation of a list: People will often spontaneously group lists. Such *subjective grouping* was noticed by Frankish (1974): "organisational strategies are widely used even in comparatively straightforward tests of short-term memory, such as the digit-span task." (PhD abstract). For example, people will often tend to group digits into threes, as they might do for a telephone number. Surprisingly however, subjective grouping in serial recall is often overlooked. Closer inspection of serial position curves nearly always reveals some scalloping, particularly over the third and fourth positions. Though this might be dismissed as random variability in one study, the pattern is consistent

---

1. Though the notion of chunks and groups are distinct, they are related, as demonstrated by Bower and Springston (1970). They showed lists like *ITVRSVPBT* were recalled well when the group structure was consistent with the chunk structure (e.g., *ITV RSVP BT*), but not when it was inconsistent (e.g., *IT VRSV PBT*). Indeed, a constant grouping structure may be necessary for the development of new chunks (Bower & Winsenz, 1969). Thus chunks might be viewed as groups that have become crystallised in long-term memory.

over numerous studies (Madigan, 1980). Given the powerful effects grouping has on performance, not only in terms of lists correct, but also serial position curves and underlying error distributions, it is unwise to ignore subjective grouping. The prevalence of such grouping suggests that it plays an important role in storing serial order in short-term memory.

Experiment 2

The first aim of Experiment 2 was to examine the detailed effects on error patterns of temporal grouping of visually presented digits. The second aim was to examine error patterns in lists of seven, eight and nine digits, without such objective grouping. Most people have considerable difficulty in serial recall of lists of seven or more items (Miller, 1956/1994) and many studies have shown how the percentage of lists recalled correctly decreases with list length in an inverse sigmoidal fashion (e.g., Crannell & Parrish, 1957). However, it is unclear from previous studies how this difficulty is manifest in error patterns. Some studies (e.g., Drewnowski & Murdock, 1980) have shown serial position curves which "stretch out" with increasing list length, maintaining characteristic primacy and recency effects, but few, if any, have examined any change in the nature of these increasing errors. Does recall of longer lists produce more omissions, more intrusions, or simply many more transpositions? Do people tend to give up after recalling the first few items, leave a series of omissions in the middle, before having a stab at the last one or two items (to give a recency effect)? This is an important consideration when recall of the middle items of long lists falls as low as 20% correct (e.g., Murdock, 1968; Madigan, 1971). Moreover, the exact effects of list length and grouping on detailed error patterns is important for testing models of serial recall (Chapter 5).

The issues of supraspan recall and grouping are closely related. It is possible that people are only able to recall lists of more than five or six items by resorting to subjective grouping of lists. This possibility has been neglected in previous studies of supraspan recall.

Method

Subjects

Eighteen subjects from the APU Subject Panel were tested, seven male and six female, with a mean age of twenty-seven years.

Materials

Stimuli were lists of seven, eight or nine digits, drawn without replacement from vocabularies of eight, nine or ten digits respectively, in the range 0-9. The order of digits was random, except for no pairs of stepsize one (e.g., "34" or "43") or triplets of stepsize two (e.g., "246" or "579"). Each digit appeared approximately equally often at each position.

Procedure

In the three ungrouped conditions of seven, eight and nine digits (conditions U7, U8 and U9 respectively), digits were presented at the rate of one every 600-ms (400-ms on; 200-ms off). The fourth condition (condition G9) had nine digits grouped as three groups of three (3-3-3 grouping). This condition had faster presentation of digits within groups, one every 450-ms (400-ms on, 50-ms off), and a pause of 450-ms between groups, so that the total presentation times for grouped and ungrouped nine-item lists was equated. Subjects were instructed to use pauses to group the digits in threes, as they might do for a telephone number.

Each digit was presented in the centre of a VDU, replacing the previous one. Subjects read the digits in silence, before attempting written recall immediately after the last digit had disappeared. Responses were written on a sheet containing the appropriate number of boxes for each list-length. Subjects were instructed to recall in a strictly forward manner, writing from left to right on the response sheet. They were encouraged to guess if they were unsure, but if no digit came to mind, to put a line through the box and proceed to the next box.

Subjects were tested in all four conditions, each as a block of 20 lists. Conditions were always attempted in the order: U7, U8, U9, G9, to reduce the chance of subjects spontaneously grouping the ungrouped lists, as might occur if the G9 condition occurred before the others. Subjects received ten practice trials. The whole experiment took about 40 minutes.

Results

In brief, all conditions showed evidence of grouping in error position curves, though the effects were strongest in the objectively grouped condition. Longer lists produced more transpositions and omissions, though the distribution of these errors differed: Transpositions showed a recency effect whereas omissions did not. Grouping decreased both transpositions and omissions, and also affected their distribution. In particular, most transpositions

decreased, except for three- and six-apart transpositions between groups, which increased. These interpositions were most common between the middle of groups.

Overall Performance

The percentage of lists recalled correctly was approximately 38% in the U7 condition, 21% in the U8 condition, 9% in the U9 condition, and 15% in the G9 condition.

Error Position Curves

The effects of list-length and grouping were apparent in error position curves (Figure 3-1, upper and lower panel respectively). All error position curves showed some scalloping, suggesting spontaneous grouping by subjects: seven-item lists by 3-4 grouping, eight-item lists by 4-4 (or even 2-2-2-2) grouping and nine-item lists by 3-3-3 grouping. In other words, lists in conditions U7, U8 and U9, though not grouped objectively, were nevertheless grouped subjectively. Indeed, most subjects reported using some form of grouping in these conditions, even though objective grouping was only introduced in the last condition, G9. The subjective grouping in condition U9 was unfortunate, in that it no longer provided a truly ungrouped baseline with which to compare condition G9. Nevertheless, it remains unclear whether truly ungrouped recall of supraspan lists is possible.

Longer list-lengths increased errors on nearly all positions, including the first, as revealed by a one-way ANOVA on the log-odds of an error on Position 1 in conditions U7, U8 and U9, *F(2,34)=7.78, MSE=0.55, p<.005*. A two-way ANOVA on the log-odds of an error in conditions U9 and G9 showed a significant effect of objective grouping, *F(1,323)=21.47, MSE=0.63, p<.001*, and position, *F(8,323)=61.52, p<.001*, but no significant interaction, *F(8,323)=1.12, p=.35*. The lack of an interaction confirmed that the majority of subjects grouped the nine items as three groups of three in both conditions.[2] Nevertheless, the main effect of objective grouping showed that the explicit pause in condition G9 made 3-3-3 grouping more effective and/or more consistent across trials and subjects. Comparisons across conditions U9 and G9 were still useful indices of grouping therefore.

---

2. The presence of subjective group boundaries in the U9 condition was confirmed by a conditional analysis of first errors (Chapter 1). Though the conditional probability of a first error tended to decrease across Positions 2-3, 4-6 and 7-9 (Appendix 1), there was a significant increase in this probability across Positions 3 and 4, *Z(18)=5.14, p<.0001*, and Positions 6 and 7, *Z(18)=3.19, p<.005*, indicating the start of a new group.
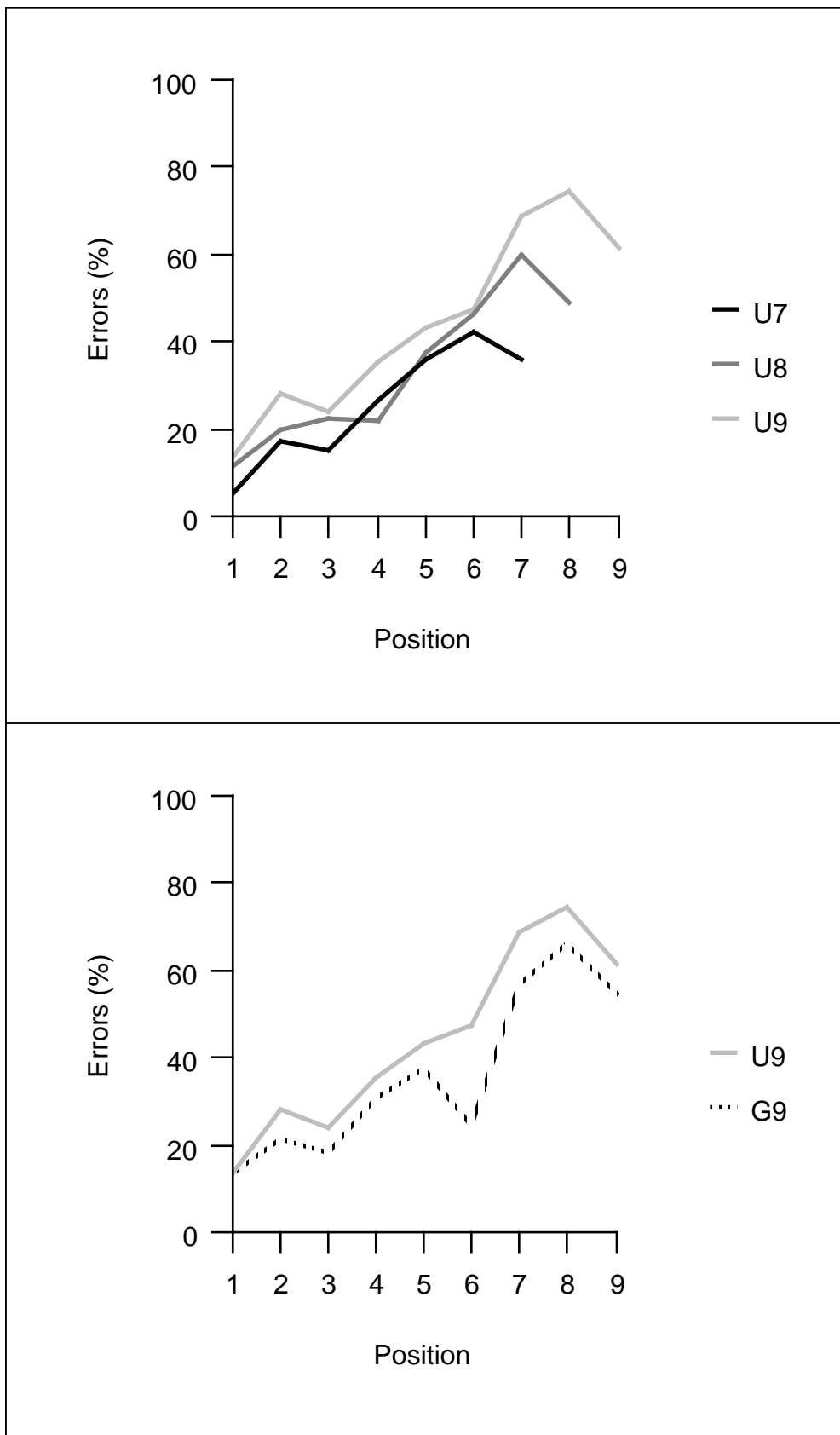
Figure 3-1: Errors by position for ungrouped lists (upper panel) and nine-item lists (lower panel) in Experiment 2.

Error Types

The proportion of responses that were omissions or transpositions was calculated for each condition (Table 3-1). Intrusions were rare (less than 5% of responses), as expected from the limited experimental vocabulary, as were repetitions. (Intrusions are analysed in more detail in Experiment 3; repetitions are analysed in more detail in Chapter 4.) A two-way ANOVA on the log-odds of an error for conditions U7, U8 and U9 showed a significant effect of error type (omission or transposition), $F(1,107)=48.41, p<.001$, list length, $F(2,136)=7.20, p<.005$, but no significant interaction, $F(2,107)=1.87, p=.16$. Transpositions were more common than omissions and both increased with list length. A two-way ANOVA on the log-odds of an error for conditions U9 and G9 also showed significantly more transpositions than omissions, $F(2,85)=11.87, p<.005$, but no significant effect of grouping or interaction, $F<1$ in both cases (though more specific tests below did show differences between these conditions).

Condition	Omissions	Transpositions
U7	.05 (.06)	.19 (.10)
U8	.10 (.08)	.22 (.10)
U9	.18 (.15)	.25 (.14)
G9	.14 (.11)	.20 (.12)

Table 3-1: Frequency of omissions and transpositions in Experiment 2.

Omissions

When split by output position, omissions in the ungrouped conditions showed a monotonic increase towards the end of recall (Figure 3-2). A similar increase was found for condition G9 at the level of groups, such that whole groups tended to be omitted towards the end of recall. One way of explaining this pattern of omissions is through a knock-on effect, where as soon as subjects forget an item, they "give up" on the remaining items, resulting in omissions for all subsequent positions (Experiment 1). However, this is not always the case,
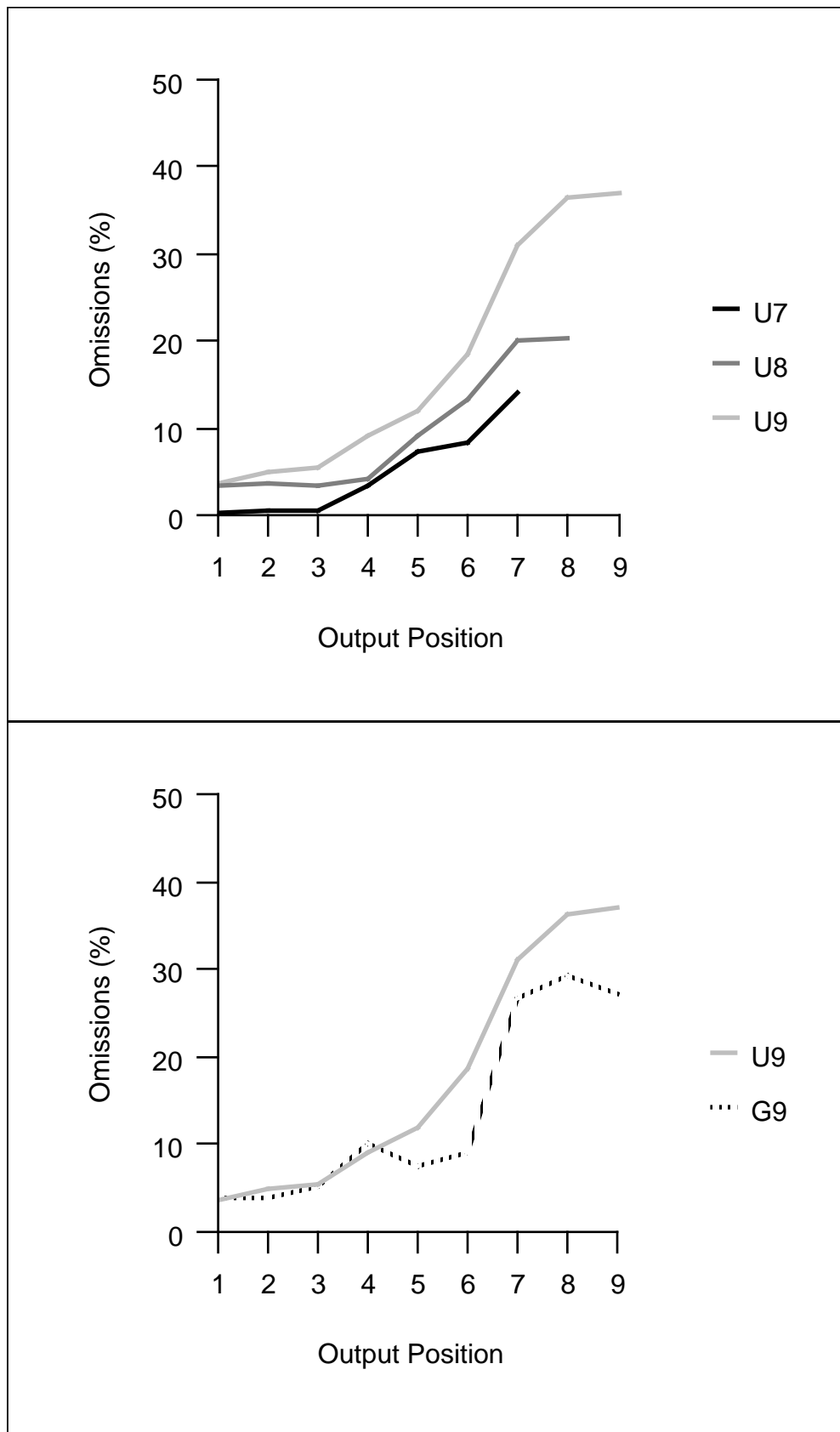
Figure 3-2: Omissions by output position for ungrouped lists (upper panel), and nine-item lists (lower panel) in Experiment 2.

because approximately 25% of the time, subjects who omitted one response (drew a line through the appropriate box) went on to make further responses (writing in subsequent boxes). Indeed, over 50% of these responses following omissions were correct. In sum, subjects could omit one item before proceeding to recall the next one correctly, but in most situations, an omission signalled that the subject could not recall the rest of the list.

Transpositions

The monotonic increase in omissions with output position meant that the recency in overall errors must have arisen from a much stronger recency effect in transpositions. This was confirmed, with all conditions showing a strong, last-item recency effect (Figure 3-3). There was also mini-recency at the end of groups, particularly in condition G9.

Transpositions in conditions U9 and G9 were further split by transposition distance, collapsing across subjects (upper panel of Figure 3-4). Unlike the transpositions in Figure 2-3, there was no monotonic decrease in transpositions with increasing transposition distance, particularly for the grouped condition. The decrease was punctuated by peaks for three- and six-apart transpositions: These were (necessarily) interpositions. Objective grouping in condition G9 not only increased three- and six-apart transpositions, but also decreased other transpositions between groups, as well as one- and two-apart transpositions within groups (Table 3-2). Indeed, tests of weighted log-odds showed that objective grouping significantly increased the proportion of transpositions that were interpositions, $Z(18)=4.61, p<.0001$, yet significantly decreased both the proportion between groups that were not interpositions, $Z(18)=2.35, p<.05$, and the proportion within groups, $Z(18)=2.18, p<.05$.

Condition	Within Groups	Between Groups	
		Interpositions	Other
U9	.38	.27	.39
	(.09)	(.09)	(.09)
G9	.33	.39	.30
	(.09)	(.10)	(.09)

Table 3-2: Proportion of transpositions within and between groups in Experiment 2. (Calculated from weighted log-odds.)
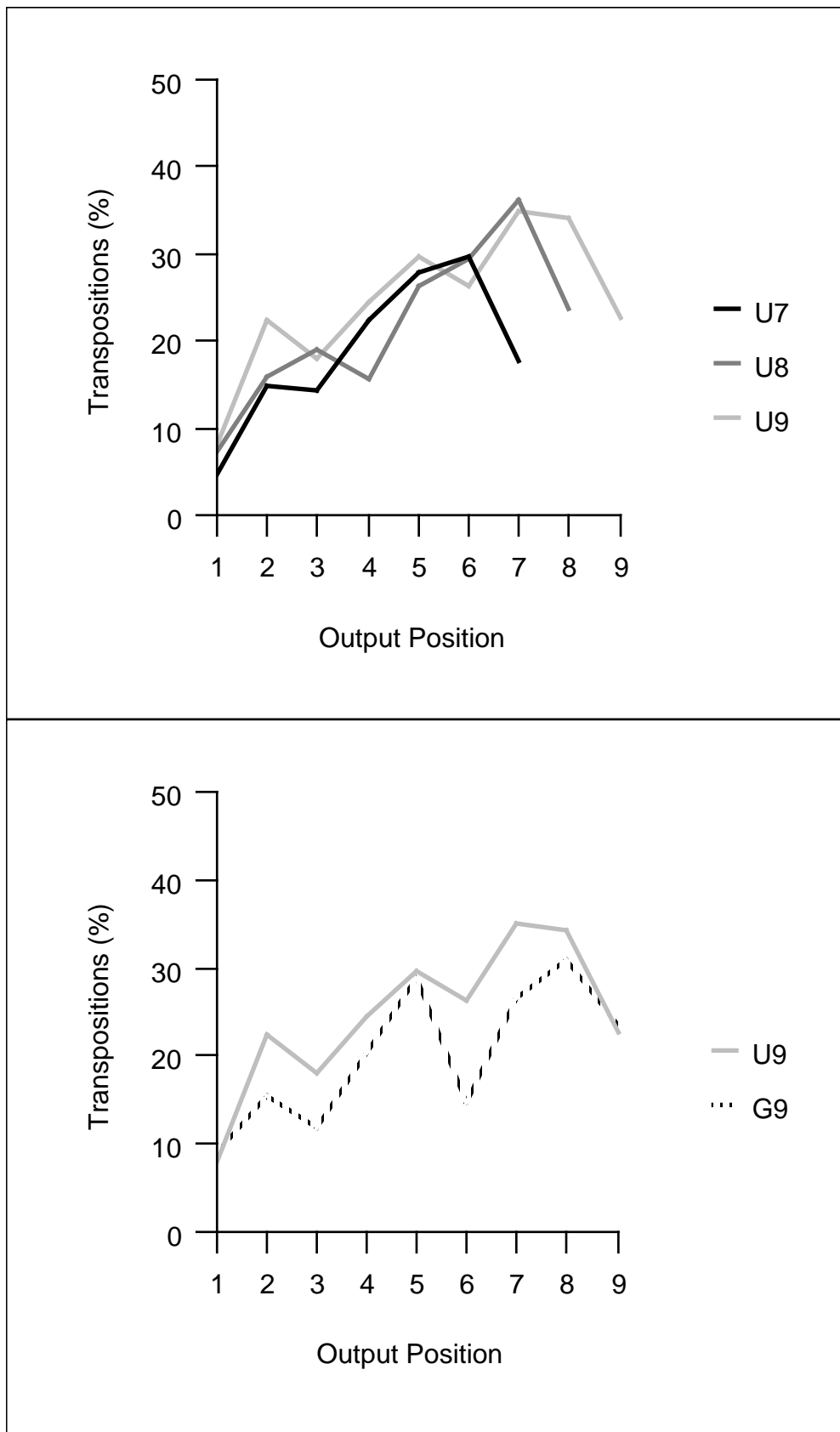
Figure 3-3: Transpositions by output position for ungrouped lists (upper panel) and nine-item lists (lower panel) in Experiment 2.
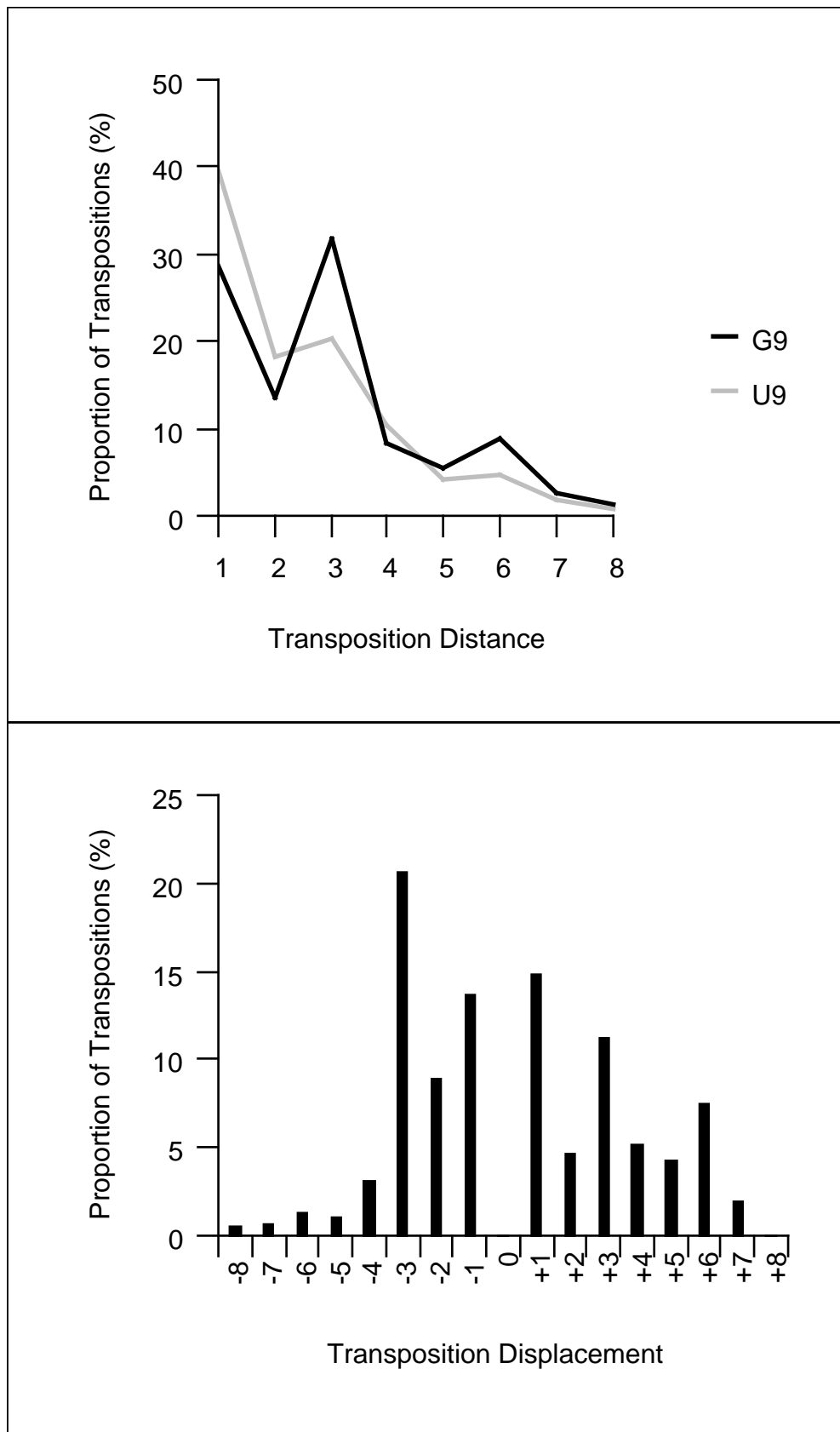
Figure 3-4: Proportion of transpositions by transposition distance in conditions G9 and U9 (upper panel), and by displacement in condition G9 (lower panel) in Experiment 2.

Further analysis of condition G9 distinguished transpositions moving forwards from transpositions moving backwards (lower panel of Figure 3-4; negative displacements represent items recalled too early, or *anticipations*; positive displacements represent items recalled too late, or *perseverations*). Most interpositions were anticipations from the immediately following group. Of the more remote six-apart interpositions between nonadjacent groups, most were perseverations (and often repetitions). The majority of interpositions were between the middle of groups (*M=.44, SD=.17*) rather than the start (*M=.33, SD=.16*) or end (*M=.28, SD=.16*) of groups, differences that were significant under weighted log-odds, *Z(17)>2.55, p<.05* (excluding one subject who made no interpositions).

Approximately 40% of the responses following an interposition in condition G9 were correct, and 18% were associates (i.e., a further interposition). Thus interpositions were not the result of whole groups swapping (which would predict 66% of interpositions to be followed by associates). The majority of interpositions occurred singly, with the correct response often following in spite of the error. Nevertheless, some dependency between recall of groups and the items within groups was suggested by the fact that only 11% of interpositions were followed by associates in the condition U9, close to that expected by chance (and considerably greater than expected from the locality constraint; Chapter 1).

Discussion

The most important finding of Experiment 2 was that people will spontaneously group items presented at a constant rate without any instruction to that purpose, with a choice of group size dependent on list-length. This is hardly surprising in one sense, given that many everyday sequences such as telephone numbers are explicitly grouped, and yet it is surprising in the sense that spontaneous grouping is normally overlooked in psychological studies. Groups of three were most common, though a constant group size appeared preferable, in that lists of eight items were grouped in twos or fours. These "natural choices" of grouping strategy are those that are optimal when grouping strategy is imposed (Wickelgren, 1967).

Spontaneous grouping is rarely reported in studies of serial recall, and yet it is often evident in serial position curves (Madigan, 1980). This means that failures to find significant differences between objectively ungrouped and grouped conditions (e.g., Ryan's tone pips)

may be an artefact of subjective grouping in the "ungrouped" condition. In fact, the notion of an ungrouped, supraspan list may be a myth, and people have to resort to grouping in order to recall more than six or seven items in order (otherwise they fail completely). A need to group supraspan lists, but not subspan lists, may begin to explain some of the differences between these cases (Brooks & Watkins, 1990). The model in Chapter 5 provides a rationale for this grouping hypothesis, in terms of the limited resolution of positional codes. That spontaneous grouping is not always apparent in serial position curves may be an artefact of averaging over subjects using different grouping strategies. Moreover, serial position curves are not the best indicators of grouping, given that grouping can be evident in conditional analyses and inter-response times without necessarily being evident in serial position curves (Frankish, 1974).

The size of the grouping effect depends on factors such as method of grouping (Ryan, 1969a) and modality (Frankish, 1974). The objective, temporal grouping in the present experiment exerted effects beyond those of subjective grouping in the ungrouped condition, through either stronger or more consistent grouping across subjects. The effects of objective grouping were threefold: It 1) decreased transpositions within groups, 2) increased interpositions between groups, and 3) decreased all other transpositions between groups. The interpositions generally arose singly, without whole groups swapping (though not necessarily completely independently). Such interpositions are important because they imply that people encode the position of an item within a group. This supports positional models of serial recall.

In addition to the effects of grouping, the second aim of the present experiment was to examine the effects of list length on different error types. Intrusions and repetitions were rare. Omissions were more common, particularly for longer lists, where they were almost as common as transpositions. Omissions increased monotonically towards the end of recall, showing no recency effect. The recency effect in overall errors came from the marked reduction in transpositions on the last position. The fact that people can omit before going on to recall the next item is important for this pattern of transpositions (Chapter 4). When people are not given the option of omitting, serial position curves often fail to show any recency effects (e.g., Drewnowski & Murdock, 1980). This pattern of omissions and transpositions proves important for testing models of serial recall (Chapter 5).

Experiment 3

The interpositions of the previous experiment demonstrated the existence of position-in-group information in recall of grouped lists. The present experiment examined the evidence for position-in-list information in ungrouped lists. This evidence came from position-specific intrusions between trials, that is, protrusions.

Conrad (1960) demonstrated protrusions in immediate serial recall of eight digits (though he called them serial order intrusions; Chapter 1). He showed erroneous responses of items that occurred at the same position in the previous list were more common than predicted by chance. He also found that increasing the intertrial interval decreased the incidence of such protrusions without affecting overall performance. He used the latter to argue that protrusions do not cause errors in recall, but arise after recall has already failed.

One way protrusions might have arisen in Conrad's experiment is through guessing strategies. When people forget an item and resort to guessing, they might tend to guess what they recalled last time. If the frequency of forgetting and hence guessing increases towards the end of recall (as might be expected from corresponding increases in omissions; Experiment 2), then later responses are likely to be guesses of the most recent items from the last trial (i.e., those from end of that trial). This would produce a greater incidence of protrusions on later positions than expected by chance, but only for short intertrial intervals (when people remember what they recalled last time). This is consistent with Drewnowski and Murdock's (1980) observation that intrusions are "overwhelmingly derived from the terminal location of the preceding list" (p. 329). Because protrusions are assumed to be guesses, this hypothesis is also consistent with Conrad's argument that protrusions do not play a causal role in forgetting.

The present experiment was an attempt to confirm Conrad's results and test the above guessing hypothesis. The present experiment had three important differences however:

1. Intertrial intervals of 2 and 20 seconds were used, rather than Conrad's intervals of 15, 25 and 40 seconds, in order to test a shorter intertrial interval. Also, Conrad did not report giving any instruction to subjects for the unfilled intertrial interval, whereas subjects in the present experiment were required to shadow a random sequence of digits between trials. This was to prevent subjects dwelling on (or even "rehearsing") previous lists.

2. Lists of five words were used, rather than Conrad's lists of eight digits. The use of such short lists was to reduce the need for subjects to group the lists subjectively (Experiment 2), a factor overlooked by Conrad. In order to produce significant numbers of errors for such short lists however, recall was delayed slightly: Subjects were required to shadow three further digits during the retention interval.

3. Lists were constructed such that no word appeared in two successive trials, unlike Conrad, who reused the same items on each trial. This was to ensure that any protrusions from previous lists were, by necessity, intrusions. The proportion of intrusions that were protrusions can therefore be compared to that expected by chance (one fifth), without needing to control for any artefactual correlations between item positions across lists.

It was hoped that these procedural differences would allow a clearer demonstration of proactive interference of position-in-list information.

Method

Subjects

Eighteen students from Cambridge University were tested, nine male and nine female, with a mean age of twenty-five years.

Materials

Stimuli were lists of five phonologically nonconfusable, single-syllable, low-frequency words. All words contained five letters, had a Kucera-Francis frequency between four and five, and possessed both a unique vowel sound and a unique first letter. Half the lists had words drawn from the set: *yacht*, *goose*, *verve*, *psalm*, *wedge*, *haunt* and *clump*. The other half had words drawn from the set: *kneel*, *midge*, *latch*, *shine*, *bathe*, *flown* and *trout*. These two sets were alternated across trials, so that no word appeared in two successive trials. The order of words within lists was randomised with the constraint that, over all trials, each word appeared equally often at each of the five positions.

Procedure

Each word was presented in the centre of a VDU, replacing the previous one, at a rate of one every second (500-ms on, 500-ms off). The fifth word was followed by a further

sequence of three digits (drawn randomly without replacement from the set 1-9), presented at the same rate as the words. Subjects vocalised each word and digit as it appeared, but recalled only the words, by speaking them aloud in the same order that they saw them. Their responses were written down by the experimenter. Subjects were encouraged to guess if they were unsure, or to say "blank" if no word came to mind. After finishing recall, subjects vocalised a further two random digits (the Short condition) or twenty random digits (the Long condition), presented at the same rate of one a second, before pressing a key to commence the next trial.

Subjects received six practice trials, before being tested on a block of 28 lists for each condition (though the first two trials per block were not analysed in order to allow appreciable proactive interference to emerge). The order of conditions was counterbalanced across subjects. The whole experiment took about 40 minutes.

Scoring Protrusions

Two types of protrusions can be identified: intrusions of items at the same position in the previous list (*input protrusions*) and intrusions of items at the same position in the previous report (*output protrusions*). Given that recall on the previous trial may not be veridical, distributions of input and output protrusions can differ. These two types are compared below.

Results

In brief, overall performance was significantly worse in the Short than Long condition, mainly owing to an increase in intrusions, and the incidence of protrusions was significantly above chance in both conditions. These results are contrary to Conrad's (1960). Furthermore, output protrusions were the most common intrusion over all positions. This result contradicts the guessing account of protrusions outlined above.

Overall Performance

The proportion of lists correct in the Short condition (*M=.31*, *SD=.25*) was less than in the Long condition (*M=.42, SD=.27*), a difference that was significant, *Z(18)=3.60, p<.001*.

Error Types

The main difference between the two conditions was a greater incidence of omissions and intrusions in the Short condition than Long condition (Table 3-3). Tests of weighted, log-

Condition	Omissions	Intrusions	Transpositions
Short	.06 (.06)	.09 (.07)	.22 (.13)
Long	.04 (.05)	.05 (.07)	.22 (.12)

Table 3-3: Frequency of omissions, intrusions and transpositions in Experiment 3.

odds showed these differences were significant in both cases, $Z(18)=2.67, p<.01$, and $Z(18)=4.76, p<.0001$, respectively. There was no significant difference in the incidence of transpositions, $Z(18)=0.06, p=.95$, and repetitions were negligible.

The greater error rate in the Short condition was spread mainly over middle positions (upper panel of Figure 3-5). There was no evidence of spontaneous grouping in these curves, or in transposition gradients. The distribution of omissions and transpositions (lower panel of Figure 3-5) was similar to that in Experiment 2. The distribution of intrusions showed a small recency effect, with most intrusions occurring on the penultimate position.

Intrusions

Extravocabulary intrusions were rare, and those that did occur were normally phonologically related words (e.g., *verge* for *verve*, *or shown* for *flown*). Of the nine possible intravocabulary intrusions on each trial, the majority were immediate intrusions of one of the five items in the previous list (*input intrusions*) or previous report (*output intrusions*).[3]

The frequency of immediate intrusions, and the proportion that were protrusions, was calculated for the 15 subjects who made at least one immediate intrusion (Table 3-4). A two-way ANOVA on the log-odds of an immediate intrusion showed a significant effect of condition, $F(1,42)=142.97, p<.001$, though no significant effect of scoring by input or output, or interaction, $F<1$. The effect of condition reflected a greater incidence of immediate intrusions in the Short than Long Condition. There was a trend for a greater proportion of output than input intrusions, but the lack of significance in this case was not surprising, because the two measures were highly correlated (given that most responses were correct).

---

3. A small proportion (24%) of output intrusions were also transpositions with respect to the current trial.
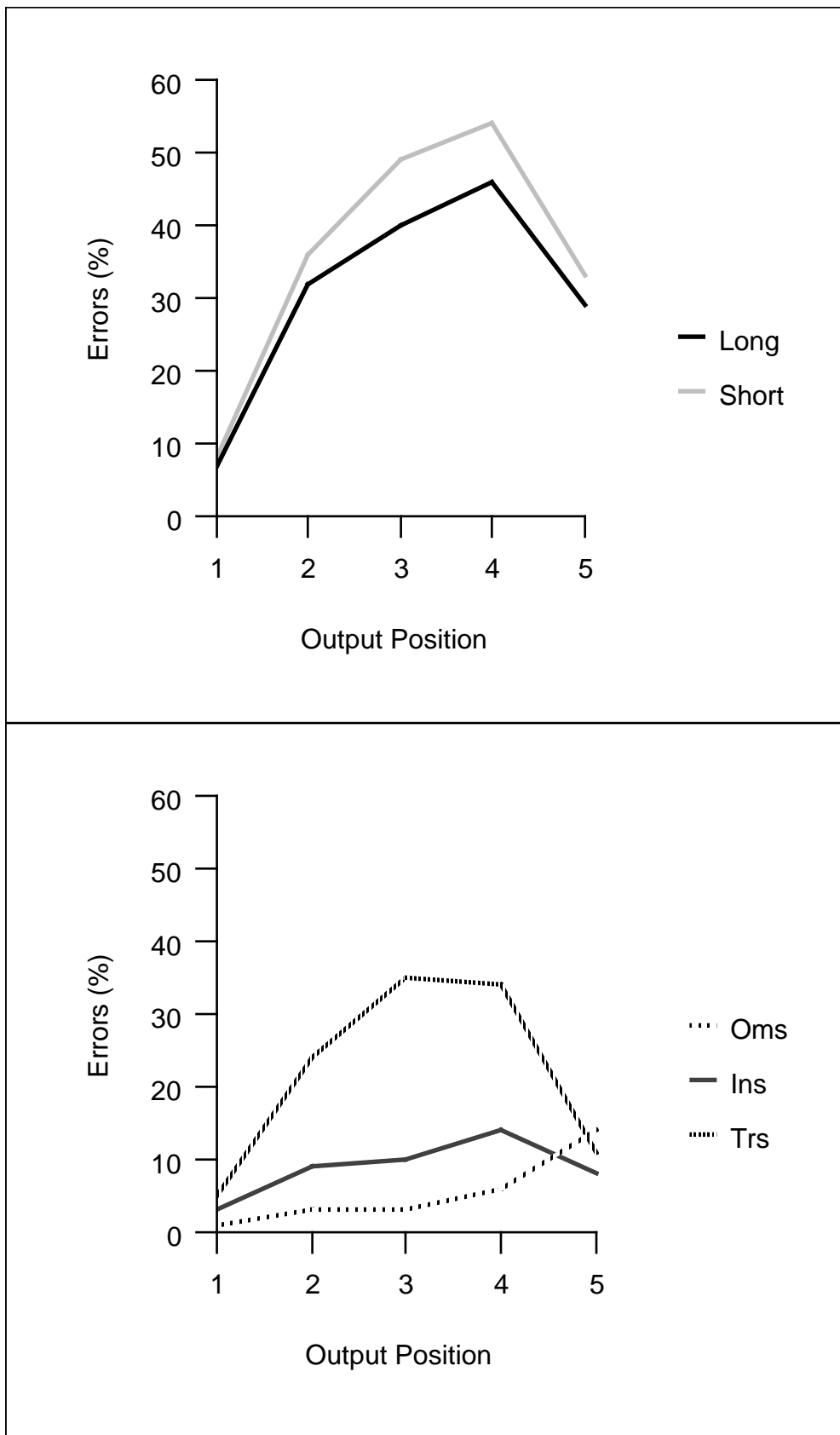
Figure 3-5: Errors by position for Long and Short conditions (upper panel), and error types by output position for the Short condition (lower panel) in Experiment 3. (Oms=omissions, Ins=intrusions, Trs=transpositions.)

Condition	Immediate intrusions		Protrusions	
	Input	Output	Input	Output
Short	.10	.15	.31	.37
	(.04)	(.05)	(.19)	(.17)
Long	.07	.09	.32	.38
	(.04)	(.05)	(.28)	(.30)

Table 3-4: Frequency of immediate intrusions and proportion that were protrusions in Experiment 3.
(Calculated from weighted log-odds, *n=15*.)

A two-way ANOVA on the log-odds that an immediate intrusion was a protrusion showed no significant effects of condition, scoring, or interaction, *F(1,42)<1.97, p>.17*. Given that there were more immediate intrusions in the Short than Long condition, this implies that there were also more protrusions in the Short than Long condition, contrary to Conrad (as confirmed by an ANOVA on the proportion of responses that were protrusions, which showed a significant effect of condition, *F(1,51)=30.93, MSE=.052, p<.001*). Most importantly, the proportion of immediate intrusions that were protrusions was significantly above chance (*.20*) for both input, *Z(18)>2.82, p<.005,* and output protrusions, *Z(15)>4.38, p<.0001.*

Output intrusions were examined by position, collapsing across subjects and condition. The resulting intrusion gradients showed peaks for output positions that corresponded to the same output position in the previous report (i.e., protrusions; upper panel of Figure 3-6). In other words, an intrusion was likely to have come from the same or nearby position in the previous report. Though the number of output intrusions was greatest for middle positions, the proportion that were protrusions was greatest for the first position (lower panel of Figure 3-6).

Protrusions could have arisen because subjects mistakenly repeated their entire previous report in recall of the current list. This would result in a protrusion being scored on every position. However, further analysis shows such perseveration of whole sequences was rare: Only 7% of the output intrusions on Positions 1 to 4 were followed by a further output protrusion, whereas 34% were followed by a correct response. Thus protrusions, like interpositions, occur singly, and not simply from recalling the wrong list on the wrong trial.
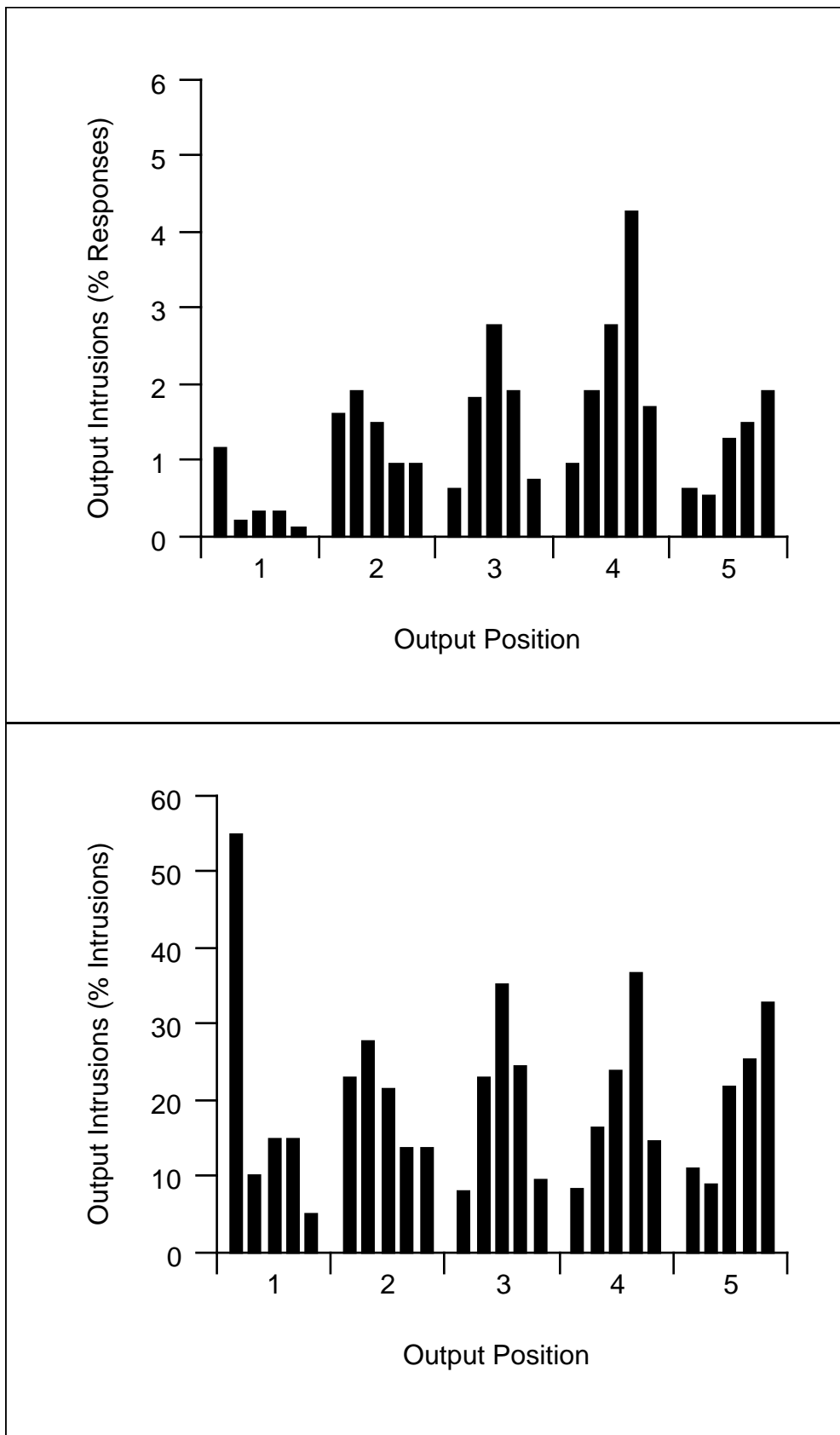
Figure 3-6: Output intrusions as a proportion of responses (upper panel) and as a proportion of intrusions per output position (lower panel) in Experiment 3.

Discussion

In agreement with Conrad (1960), the present experiment showed a significant incidence of protrusions, which decreased as the intertrial interval increased. Furthermore, protrusions were the most common intrusions across all positions in recall of ungrouped, five-item lists. The triangular-shaped intrusion gradients were therefore a graphic illustration of proactive interference of position-in-list information.

Unlike Conrad however, the greater number of protrusions with a short intertrial interval was mirrored by significantly worse overall performance. The most likely reason why Conrad failed to find such a difference in overall performance is that his shortest intertrial interval was 15 seconds, rather than the 2 seconds of the present experiment. If the amount of proactive interference is inversely related to the intertrial interval (e.g., Turvey, Brick & Osborn, 1970), the present experiment is a more powerful test of the effects of intertrial interval, by virtue of a greater range of proactive interference.

Given that the poorer performance in the short intertrial interval owed mainly to a greater number of intrusions (with a small increase in omissions, but hardly any change in transpositions), it seems reasonable to conclude, contrary to Conrad, that proactive interference does play a causal role in forgetting in short-term serial recall. A similar conclusion was reached by Sanders and Willemsen (1978b). In particular, one noncausal explanation for protrusions was ruled out by present results: the hypothesis that protrusions are an artefact of a guessing strategy. This hypothesis predicts that protrusions should be confined towards later positions, when recall falters, and yet protrusions were found across all positions in the present experiment (and contrary to Drewnowski & Murdock, 1980, intrusions were not overwhelmingly from the last position of the previous trial). Further evidence that protrusions are not simply guesses is given in Experiment 5. More likely, protrusions result from response competition at each position of recall (Chapter 5).

Also unlike Conrad's data, the proportion of immediate intrusions that were protrusions was still significantly above chance after 20 seconds between trials. One reason may be that Conrad employed immediate serial recall, rather than delaying recall by 3 seconds as in the present experiment. Longer retention intervals will tend to increase proactive

interference (e.g., Crowder, 1993). Another reason may be because Conrad's design meant he could not distinguish intrusions from transpositions, making classification of protrusions uncertain. Even longer intertrial intervals therefore, such as Conrad's 40 second delay, may be required before protrusions fall to chance levels. As such, the proactive interference in the present experiment demonstrates a surprising longevity of short-term memory for positional information. Indeed, Nairne found evidence for positional information after two minutes of distraction following incidental learning (Nairne, 1991).

Conrad only measured input protrusions. However, the present experiment showed that output protrusions were more common than input protrusions (a trend confirmed in Chapter 4). In other words, if an item is recalled in the wrong position on one trial, it appears more likely to protrude on the next trial in its position of recall rather than its position of presentation. This suggests that recall is itself a learning episode, such that, in the case of errors, an item is relearned in a different position. The previous report may be a greater source of proactive interference because it represents a more recent learning episode than the previous list. Alternatively, recall may be a stronger learning episode than presentation, as suggested by the fact that the Hebb effect requires multiple recall attempts; multiple vocalised presentations are insufficient (Cunningham, Healy & Williams, 1984). Another possibility is that responses enter a separate rehearsal store, which becomes a secondary source of interference between trials (Estes, 1991). The questions of proactive interference and recoding during recall are covered more fully in Chapter 6.

Finally, the intrusion gradients in Figure 3-6 demonstrated that intrusions were more common in the middle than the start or end of reports. However, the proportion that were protrusions was greatest at the start, suggesting that the first position is coded more precisely. In other words, proportional intrusion gradients give an idea of the *positional uncertainty* associated with each position: Shallower gradients indicate greater positional uncertainty. Nevertheless, intrusion gradients are not a perfect reflection of positional information in short-term memory. There are several reasons why the positional information used in serial recall may be considerably more precise (giving the sharper transposition gradients in Figure 2-2). Firstly, intrusion gradients necessarily index positional information from the previous trial,

which is likely to become less accurate over time. Secondly, there may be several sources of proactive interference, such as that from even earlier trials, which will introduce additional noise to the extent that the sources are uncorrelated. Thirdly, there are extraneous reasons for intrusions, such as people's predisposition to guess certain words. (One subject for example recalled the word "shine" on nearly every trial.) This additional noise will blur intrusion gradients even further. These points are relevant to the question of whether positional information is sufficient to underlie serial recall (below).

General Discussion

The present experiments demonstrated two types of positional errors: interpositions between groups that maintain their position within group (Experiment 2), and protrusions between trials that maintain their position within trials (Experiment 3). These errors are evidence of positional information in short-term memory, supporting positional theory. They cannot be explained by ordinal or chaining theory.

The intrusion gradients in Experiment 3 resemble the transposition gradients in Experiment 1. As Page and Norris (1996b) demonstrate however, transposition gradients do not, on their own, necessitate a positional model. Their ordinal model produced similar transposition gradients for each position. This is because errors in the relative order of nearby items also produce peaked transposition gradients. However, the fact that intrusions show peaked gradients does necessitate a positional model. This is because intrusions that maintain positions between trials cannot be attributed to errors of relative order within a trial (Chapter 4). Protrusions indicate that items are coded for position independently of surrounding items. A similar argument applies to interpositions between groups.

Page and Norris (1996b) made several arguments against interpreting positional errors as evidence for positional models of immediate serial recall. Firstly, they argued that positional information might be limited and therefore insufficient to support serial recall. The limitation of positional information was based on the argument that group sizes of three are optimal, in which position can be characterised as *start, middle,* and *end* (Wickelgren, 1967). These codes only require specification of the first and last item of each group, since the middle item can be defined by exclusion. Such codes are sufficient to explain the interpositions in

Experiment 2. For larger groups of items however, the codes *start, middle* and *end* would not be sufficient to order nonterminal items, suggesting that positional coding is limited to three positions at the most. This suggestion is refuted by the intrusion gradients in Experiment 3. These five peaked gradients demonstrate that positional information extends beyond start, middle and end. Subjects in Experiment 3 must have possessed at least five positional codes, perhaps even *first*, *second*, *third*, *fourth*, and *fifth* (though Chapter 6 argues for a somewhat different representation of position).

This prompts the question of whether subjects in Experiment 2 possessed codes for "first" through to "ninth" in the nine-item, ungrouped condition. This is possible, but the fact that they preferred to spontaneously group such lists suggests that there may well be a limit to the number of positions people can distinguish, as Page and Norris suggest. By breaking a list into groups, a smaller range of positional codes may suffice, by using *first*, *second* and *third* to code both an item's position within a group, and the group's position within the list. This is illustrated in Chapter 5. The model developed in that chapter has a limited range of positional coding, yet one that is more fine-grained than the *start*, *middle* and *end* of Page and Norris. Moreover, the model demonstrates that this information is sufficient to support serial recall.

Another argument offered by Page and Norris is that positional errors have typically been demonstrated with delayed rather than immediate serial recall. The phonological similarity effect, which Page and Norris use as a signature of their model of the phonological loop, decreases as the recall delay increases (Baddeley, 1986). A corresponding increase in positional errors would produce a double dissociation that might suggest two different sources underlying serial recall, an ordinal (phonological) one and a positional (nonphonological) one. However, while it is true that delayed recall was employed in Experiment 3 (to increase overall error rates), the meta-analyses in Chapter 4 reveal that positional errors also arise in immediate serial recall of span-length lists. Moreover, the model developed in Chapter 5 explains the trade-off between positional and phonological errors without appealing to two different theories of serial order. The increase in positional errors with delay is attributed to a ratio-rule of proactive interference (e.g., Crowder, 1993), applying to positional information, and the decrease in phonological errors is attributed to rapidly-decaying, phonological traces

(e.g., Tehan & Humphreys, 1995), constituting only supplementary item information.

A third argument by Page and Norris is that positional errors are epiphenomenal rather than causal. This might be suggested by the rarity of positional errors like protrusions (Experiment 3). There are several counterarguments. Firstly, interpositions are a far more common example of positional error. Indeed, they were more common than adjacent transpositions in Experiment 2. Secondly, Conrad's belief in a noncausal role of protrusions was contradicted by Experiment 3, which did suggest a causal role. Finally, positional errors are not restricted to guesses (Experiments 4 and 5). Indeed, anecdotal evidence suggests that people often make interpositions without even being aware of having made an error.

A final argument offered by Page and Norris was whether positional errors are an artefact of subsidiary processes in serial recall. For example, are they an artefact of subjects copying down responses from previous trials during written recall? This cannot be true of the spoken recall in Experiment 3. Are they an artefact of output processes operating in a speech buffer, like those assumed to underlie speech errors (Chapter 8)? This seems unlikely in the case of protrusions in the Long condition of Experiment 3, which remained above chance even after at least 28 seconds of shadowing digits and letters between the termination of recall in one trial and the initiation of recall in the next. (The speech buffer account also has difficulty explaining why interpositions remain under articulatory suppression, Page & Norris, 1996a, and why protrusions do not increase with degree of vocalisation, Murray, 1965.) Are they an artefact of visuospatial strategies such as imagery? This also seems unlikely, given that there was no objective spatial information in the sequential presentation of Experiments 2 and 3 (or in the spoken recall of Experiment 3), and given that imagery was rarely reported during debriefing. Moreover, others have argued that the role of visuospatial information in such experiments is minimal (Hitch & Morton, 1975).

Thus none of the arguments offered by Page and Norris (1996b) appears to hold in the light of Experiments 2 and 3. These experiments demonstrate positional information that extends over at least five positions and plays an important, causal role in serial recall from short-term memory. The exact nature of this information is deferred to Chapters 5, 6 and 8.

Chapter Summary

This chapter described two experiments providing evidence for positional errors in serial recall from short-term memory. These errors are explicable by positional theory, but not chaining or ordinal theories. Nevertheless, Chapter 4 continues to entertain all three theories and tests their predictions in a more detailed analysis of error distributions. This takes the form of meta-analyses over a number of experiments, complementing and confirming the analyses in Experiments 1-3, and providing a fuller set of constraints for the model in Chapter 5.