Examining the relationship between free recall and immediate serial recall: the effect of list length and output order

Geoff Ward[1], Lydia Tan[2], and Rachel Grenfell-Essam[1]

[1]University of Essex

[2] City University

Correspondence to:
Geoff Ward,
Department of Psychology,
University of Essex,
Wivenhoe Park,
Colchester,
Essex
CO4 3SQ, UK

Telephone: +44 1206 873799
Fax: +44 1206 873590
E-mail: gdward@essex.ac.uk

Running header: serial and free recall

Abstract

In four experiments, participants were presented with lists of between 1 and 15 words for tests of immediate memory. In each experiment, participants tended to initiate recall with the first word on the list for short lists, but as the list length was increased so there was a decreased tendency to start with the first list item; and, when free to do so, participants showed an increased tendency to start with one of the last four list items. In all conditions, the start position strongly influenced the shape of the resultant serial position curves: when recall started at serial position 1, elevated recall of early list items was observed; when recall started towards the end of the list, there were extended recency effects. These results occurred under free recall, and different variants of immediate serial recall (ISR) and reconstruction of order tasks. We argue that these findings have implications for the relationship between recall and rehearsal and free recall and ISR.

161 words (986 characters)

The overall aim of this research is to promote greater theoretical integration between two highly important and widely-used tests of immediate memory: immediate serial recall (ISR) and immediate free recall. The main claim of the paper is that greater theoretical integration between these tasks can be achieved, if only researchers understood fully the effects of increasing list length on both tasks. To this end, we report the data from four experiments looking at the effects of increasing list length on the output order and serial position curves for the immediate free recall task, and the ISR task and its variants.

At first glance, one might think that there should be no need to promote greater theoretical integration between ISR and free recall. On the face of it, the methodologies of the two immediate memory tests are remarkably similar: in both tasks, participants are presented with a sequence of items and at the end of the list they must try to recall as many items as possible, in either the same order as that presented (ISR) or in any order (free recall). In addition, both tasks share a common theoretical heritage. Both tasks have provided empirical evidence taken as key ‘signature findings’ supporting the establishment of a short-term memory store (STS) of limited capacity: in ISR, the memory span limitation has been taken to reflect the limited capacity of verbal STS (whether this be measured in items, chunks, or time, e.g., Baddeley, 1986; Miller, 1956) and the advantage in recall of the last items known as the recency effect in free recall has also been taken as evidence for the direct output of items at test from a short-term buffer store (e.g., Atkinson & Shiffrin, 1971; Glanzer, 1972)

It is therefore perhaps surprising that most current theories of ISR do not provide a detailed account of free recall. For example, currently influential accounts of ISR include: the phonological loop model of working memory (Baddeley & Hitch, 1974; Baddeley, 1986, 2000), the Burgess and Hitch (1992, 1999, 2006) model, the primacy model of Page and Norris (1998), the Start-End model of Henson (1998), the feature model (Nairne, 1988, 1990), OSCAR (Brown, Preece, & Hulme, 2000), SOB (Farrell & Lewandowsky, 2002;

Lewandowsky & Farrell, 2008), and the accounts by Botvinick and Plaut (2006) and by Oberauer and Lewandowsky (2008). None of these current accounts of ISR provide a detailed account of free recall.

Moreover, most current theories of the free recall task do not provide a detailed account of ISR. Influential accounts by Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, and Usher (2005), the Temporal Context Model (Howard & Kahana, 2002) and its recent variants (e.g., Polyn, Norman & Kahana, 2009; Sederberg, Howard & Kahana, 2008), the account by Laming (2006, 2008, 2009), the SAM model (Raaijmakers & Shiffrin, 1981), and indeed our own account of free recall (Tan & Ward, 2000) do not explain ISR.

There are a few exceptions, such as the SIMPLE model of Brown, Neath and Chater (2007), the ACT-R model of Anderson, Bothell, Lebiere, and Matessa (1998), and the LIST PARSE model of Grossberg and Pearson (2008). However, it should be noted that some of these models differ in the type and number of mechanisms required to underpin the two tasks.

We propose three reasons for the current lack of theoretical integration between these two tasks in the literature. First, there are capacity difficulties in entertaining a limited-capacity STS account of both ISR and the recency effect in free recall. For example, Baddeley and Hitch (1974, 1977) asked participants to study lists of 16 words for free recall whilst concurrently maintaining lists of 6 digits for ISR. They found that the magnitude of the recency effect in free recall was unaffected by a concurrent 6-digit load for ISR, suggesting that active maintenance of 6 digits for ISR could not be competing for the same limited-capacity STS as the maintenance of the last few words in the list for free recall (a finding since replicated and extended by Bhatarah, Ward & Tan, 2006). Different theoretical approaches to recency and ISR is therefore more likely to the extent that authors appeal to STS as an explanation of one of these findings. For example, Baddeley and Hitch concluded that ISR was underpinned by STS, but “working memory, which in other respects can be regarded as a modified STS, does not provide the basis for recency" (Baddeley and Hitch (1974, p.81).

Second, early reviews of the free recall and ISR literatures found that ISR was affected by variables such as phonological similarity, word length, presentation rate, and memory load (suggesting a speech-based STS), but the recency effect in free recall was not particularly sensitive to these variables (see e.g., Baddeley, 1976, p. 182). This differential affect of different variables on ISR and recency further suggested a division between explanations of the recency effect in free recall (not underpinned by STS) and ISR (underpinned by STS).

A third and final reason for the current lack of theoretical integration between theories of free recall and ISR is that the two tasks give rise to highly dissimilar serial position curves. In free recall, explanations of the serial position curve are dominated by explanations of recency effects (e.g., Davelaar et al, 2005; Howard & Kahana, 2002; Tan & Ward, 2000), whereas in ISR, explanations of the serial position curve are dominated by explanations of primacy effects (e.g., Lewandowsky & Farrell, 2008; Page & Norris, 1998). These contrasting serial position curves offer a difficult challenge to accounts of memory that try to model both tasks, and this difficulty is made all the more transparent in computational models where the calculations and predictions of the models are most explicit.

In short, it is fair to say that it is currently controversial to even consider a common theoretical framework for free recall and ISR (for a recent exchange of views on this issue, see e.g., Brown, Chater, & Neath, 2008; Murdock, 2008).

However, we believe that most previous comparisons between the free recall and ISR tasks have been hindered by the use of different list lengths on the two tasks. Typically, short lists of 5-6 words are used for ISR, whereas much longer list lengths of 10-40 items are used for free recall. Therefore, apparent differences between the two tasks may have been reflecting differences in the typical list lengths that were used.

Consistent with this possibility, many similarities have recently been found between ISR and free recall when the two tasks have been compared under identical presentation conditions and identical list lengths of 6 or 8 words (Bhatarah, Ward & Tan, 2008; Bhatarah, Ward, Hayes, & Smith, 2009). For example, (1) words on both tasks are rehearsed and encoded in the same way with little or no effect of test expectancy, even though the output order of the two tasks can be very different, (2) the degree of forwards ordered recall can be similar on the two tasks, even though free recall does not require forwards ordered recall, (3) both tasks are affected by word length, and articulatory suppression, variables traditionally associated with ISR, and (4) both tasks are affected similarly by presentation rate, a variable traditionally associated with free recall. These recent findings suggest that (in contrast to many contemporary accounts) common memory mechanisms may underpin both free recall and ISR.

In this research, we seek to examine further the case for greater theoretical integration between ISR and immediate free recall. We will again examine both tasks under identical methodological conditions, but in these experiments we will additionally manipulate list length from very short lists well within what is typically used in ISR through to longer lists typically reserved for free recall. Central to our analyses will be how increasing the list length affects the patterns of output and the serial position curves in different immediate memory tasks.

EXPERIMENT 1

In Experiment 1, we examined only free recall. On each trial, participants were presented with lists of between 1 and 15 words, but did not know the length of the list in advance of the cue to recall. The lists were presented visually and read silently at a reasonably fast rate of 1 word per second. We were interested in whether the length of the list affected the order of the words recalled. Specifically, we were interested in knowing whether list length affects the words with which participants would initiate their recall and the extent to which these initial recalls affected the resultant serial position curves.

Our hypothesis was that if the dissimilarities between free recall and ISR are due at least in part to differences in list length, then the output order and serial position curves of free recall might more closely resemble those of ISR at shorter list lengths. In summary, to the extent that participants started their free recall with the first list item and exhibited extended primacy effects with shorter lists, so we would find evidence for increased similarities between free recall and ISR.

Method

Participants. Fifty-five participants took part in this experiment from the University of Essex and City University.

Materials and Apparatus. The materials consisted of a set of 480 words taken from the Toronto Word Pool (Friendly, Franklin, Hoffman & Rubin, 1982). Subsets of 360 words were randomly selected to be the materials for each individual. The materials were presented in 52-point Times New Roman font in the centre of a computer monitor.

Design. The experiment used a within-subjects design. There were two within-subjects independent variables: list length, with fifteen levels (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15) and serial position (SP) with up to 15 levels. The dependent variable was the proportion of words recalled (in any order).

Procedure. Participants were tested individually and informed that they would be shown one practice list of 7 words followed by 45 experimental lists of words. The experimental trials were arranged into three blocks of fifteen trials. In each block, participants received one trial of each of the 15 different list lengths, but the order of the list lengths within each block was randomised. Each trial started with a warning tone, followed after 3s by a sequence of between 1 and 15 words presented one at a time in the centre of the screen. The presentation rate was 1 word every second, with each word displayed for 0.75s with an additional 0.25s inter-stimulus interval in which the stimulus field was blank. Participants were instructed to read each word silently as it was presented. At the end of the list there was an auditory cue and participants wrote down as many words as they could remember in any order that they wished.

Results

First, we examined the serial position curves using all the data. Figure 1 shows the proportion of words recalled at each of the 15 different list lengths. Consistent with list length effects in free recall, the mean proportion of words recalled decreased monotonically, from .99 (list length 1) to .30 (list length 15), F (14, 756) = 419.3, MSE = .008, p <.001. Furthermore, the mean number of words recalled increased from .99 (list length 1) to 4.49 (list length 15), F (14, 756) = 124.4, MSE = 0.461, p <.001.

--Figure 1 about here--

At short lists, such as list length 4, the proportion of words recalled was at a consistently high level (M = .81), with no significant effect of serial position, F (3, 162) = 0.802, p > .05. At longer lists, such as at list length 8, there was a significant effect of serial position, F (7, 378) = 10.21, MSE = .090, p <.001, reflecting significant recency and primacy effects of approximately equivalent magnitude. At still longer lists, such as at list length 15, there were again significant effect of serial position F (14, 756) = 20.50, MSE = .069, p <.001, reflecting significant recency and primacy effects, but there was significantly greater recall of the recency items than the primacy items.

Second, we examined which words from a list were the first to be recalled on each trial (for related analyses, see Hogan, 1975; Howard & Kahana, 1999; Laming, 1999). Table 1 shows that at short lists, participants tended to initiate their recall with the first words in the list (those presented at serial position 1), but as the list length increases, so there is a tendency to initiate recall with one of the last 4 words in the list.

--Table 1 about here--

Figure 2 clearly illustrates participants' tendency to initiate free recall with the first list item for shorter lists. As the list length increases, so there is an increasing tendency to initiate free recall with one of the last four list items for longer lists. The proportion of trials in which recall started with serial position 1 decreased with increasing list length, F (14, 756) = 127.5, MSE = .048, p <.001. For each participant, we calculated the slope of the function relating how the proportion of trials in which recall started with serial position 1 decreased with increasing list length, across sets of list lengths 1-3, 4-6, 7-9, 10-12, and 13-15. The mean slopes across these five sets differed significantly from each other, F (4, 216) = 10.12, MSE = .021, p <.001, and were -0.05, -0.17, -0.07, -0.03, and -0.01, respectively, and confirmed that the function declines most steeply across list lengths 4-6, where the addition of each extra list item reduced the proportion of trials starting with serial position 1 by on average 17%.

--Figure 2 about here--

We then examined the effect of first recall on the resultant serial position curves. Figure 3A shows the serial position curves for the free recall trials in which the first word recalled was from serial position 1 plotted using ISR scoring. Figure 3B shows the serial position curves for those same trials plotted using free recall scoring. Finally, Figure 3C shows the serial position curves for the free recall trials in which the first word recalled was from one of the last four serial positions plotted using free recall scoring. The shapes of the serial position curves in Figure 3A, 3B, and 3C are strikingly different.

For illustrative purposes, consider performance at list length 6. Thirty-five participants initiated recall on one or more trials with serial position 1 at this list length. Using ISR scoring, the mean recalls across the six serial positions in Figure 3A, are 1.00, .63, .42, .33, .17, and .13, and discounting serial position 1 (whose mean by definition is 1.00), serial positions 2-6 differed significantly, F (4, 136) = 16.47, MSE = .074, $p < .001$; pairwise comparisons confirming significant extended primacy effects with ISR scoring.

The corresponding mean recalls across the same trials using free recall scoring (Figure 3B) are 1.00, .67, .59, .59, .46, and .53. Discounting serial position 1, serial positions 2-6 did not differ significantly, F (4, 136) = 1.15, MSE = .185, $p > .05$, but remain at a relatively elevated level of recall.

Finally, consider performance by the 43 participants who initiated recall on one or more trials with one of the last four serial positions. Figure 3C shows that the mean recalls across these six serial positions are .48, .32, .44, .54, .79, and .88; means which differed significantly, F (5, 168) = 16.61, MSE = .120, $p < .001$; pairwise comparisons confirming 1-item primacy, and significant extended recency effects.

--Figure 3 about here--

Finally, we examined the extent to which the output orders showed evidence of forwards ordered recall. Figure 4 shows the lag-CRP curves (for full details, see Kahana, 1996; Howard & Kahana, 1999) for each list length. The x-axis plots the lag between successive pairs of words recalled, which is calculated by subtracting the serial position of the first word of each pair from the serial position of the second word of each pair. Smaller lag values therefore represent recall transitions between words from more similar serial positions; whereas larger lag values represent recall transitions between words from less similar serial positions. Similarly, positive lag values represent recall transitions proceeding in a forwards direction; negative lag values represent recall transitions proceeding in a backwards direction. Of critical interest is the frequency of recall transitions with lag of +1, the lag at which the output order of successively recalled pairs is the same as the input order. The y-axis plots the conditionalized response probability (CRP), which is calculated by taking the number of transitions actually made of a given lag during output and dividing this total by the number of opportunities that a participant might reasonably be expected to have had to make such a lag transition. The CRP-values control for the reduced opportunities to make transitions at extreme lags (and the increased opportunities to make transitions at small lags), and also assume that it is unreasonable for a participant to recall an item has already been recalled. Figure 4 indicates that there is evidence for what is known as the asymmetric lag recency effect (see Howard & Kahana, 1999; Kahana, 1996): there is a preference for transitions to be nearer neighbours than remote neighbours, and there is a greater tendency for transitions to proceed in a forwards rather than a backwards order. A close inspection also suggests that there are also high CRP-values at extreme positive and negative lag values.

In order to examine these observations statistically, the lags (and opportunities to output items at different lags) were categorised into 8 different lag values: extreme negative lags (the lowest possible lag value at each list length, equivalent to 1-list length), remote negative lags (lags 2-list length through to -3), -2, -1, 1, 2, remote positive lags (lags 3 to list length-2) and extreme positive lags (the highest possible lag value at each list length, equivalent to lag list length-1). Fifty participants had CRP values for all eight categories, and their mean CRPs for the eight respective categories were .21, .06, .13, .19, .50, .16, .09, and .08, which differed significantly, F (7, 343) = 105.9, MSE = .009, p <.001. Pairwise comparisons showed that there was a clear tendency for recall to follow the order at study, as demonstrated by higher CRP-values for lag +1 than all other lags, including lag -1, demonstrating significant asymmetry. There were also significant lag recency effects, the CRP values for lag -1 were greater than those for lag -2, which in turn were greater than the values for the remote negative lags, and the CRP values for lag +1 were greater than those for lag +2, which in turn were greater than the values for the remote positive lags. However, the extreme negative lag CRP-values were also significantly higher than all but the closest transitions (see Farrell & Lewandowsky, 2008), indicating a high tendency to transition between the last to the first list item.

--Figure 4 about here--

An inspection of Figure 4 also shows that the CRP-values for lag +1 decrease with increasing list length, a finding shown more clearly in Figure 5. Fifty-one participants had valid lag+1 CRP values at all 14 list lengths (list lengths 2 to 15), and these values differed significantly from each other, F (13, 650) = 46.12, MSE = .017, p <.001. At shorter lists, the majority of transitions between successive words output are in the same order as that in which the words were presented (lag +1), even though the participants were free to recall in any order. The lag +1 value decreases with increasing list length, but does not reduce to near chance levels (approximately the recipricol of the list length -1) but stays at a relatively high value of between .3 and .4.

--Figure 5 about here--

Discussion

When participants were presented with short lists of up to 3 words for free recall, they started their recall with the first word in the list and continued in forward serial order; that is, their free recall resembled ISR. As the list length was increased from lengths of 4 through to 9 words, so there was an increased tendency to initiate recall with one of the last four words (as is more typical in immediate free recall).

In addition, the first word recalled helped determine the shape of the serial position curve: heightened early list performance and reduced recency effects were observed when the first word recalled was from serial position 1, and more extended recency effects and reduced primacy effects were observed when the first word recalled was from one of the last four serial positions. In all recall, there was a clear tendency to recall in a forwards direction, but this tendency was greater for shorter lists.

It appears therefore that at least some of the apparent differences between the output orders and serial position curves observed in free recall and ISR might indeed reflect the typical differences in the list lengths that are used in the two tasks - when free recall was performed with long lists typically used in free recall, participants recalled the last items first and there was enhanced recency and reduced primacy; but when free recall was performed with shorter lists typically used in ISR, then ISR-like findings were observed, with an increased tendency to recall spontaneously from the start of the list leading to enhanced early list performance on these trials.

EXPERIMENT 2

Experiment 1 provided initial evidence supporting the call for greater theoretical integration between free recall and ISR: some “ISR-like” features of recall were observed when free recall was performed with shorter, “ISR-like” list lengths. In Experiment 2, we attempted to replicate these findings and extend them to examine whether a variant of ISR might additionally show some “free recall-like” features of recall when that variant of ISR was performed under “free recall-like” list lengths.

An immediate concern with such an endeavour was to consider the best way to examine ISR at such a wide range of list lengths. Our solution in Experiment 2 was to use a variant of the ISR task used by Tan and Ward (2007, see also Crowder, 1969) that we will refer to here as the “ISR-free” task. In the ISR-free task, participants are presented with a list of words one at a time, and at the end of the list, they are required to recall the list items in the correct serial position by writing the list items in the appropriate position in a lined response grid. In Experiment 2, this paper response grid always consisted of two columns of 15 rows, the first column contained the numbers 1 to 15 in increasing order, and the second column was empty. Thus, the participants were required to write down the words from the list in the rows corresponding to each items’ serial positions, such that the first word in the list should be written in the first row of the grid, the second word should be written in the second row, and so on. Unlike standard ISR, participants in the ISR-free task were free to write their spatially-ordered responses in the response grids in any temporal order that they like. For example, they were free to start their recall by writing the last three words from the list in the corresponding rows before returning to the top of the grid to continue recall with earlier list items.

In Experiment 2, one group of participants performed free recall, a second group performed ISR-free and all were presented with a total of 66 lists of between 1 and 15 words. In order to maximise the number of trials per list length within an experimental session, list lengths 9, 11, 13 and 14 were excluded, such that participants undertook 6 trials each of 11 different list lengths (lists of 1, 2, 3, 4, 5, 6, 8, 10, 12, and 15 words). As in Experiment 1, the words were presented visually at a reasonably fast rate of 1 word per second. Unlike Experiment 1, the participants were required to read each word aloud as the words were presented. At test, an empty response grid appeared on the screen that had the exact number of rows (between 1 and 15) appropriate for that trial, so that participants could see the number of items that had been presented and so knew which of the 15 rows on their paper response grid they should complete. In the free recall condition, participants were free to recall the words in any order – they simply wrote down the words from the top to the bottom of their response grids. In the ISR-free condition, the participants wrote the words in the positions in the grid corresponding to each word's serial position, but they were free to write down the words in any temporal order that they liked. In both conditions, the participants spoke out loud the words that they were writing down and their responses were recorded using a tape recorder to code the output order of the items.

As in Experiment 1, we were interested in the order in which words participants would initiate their recall at different list lengths, and the extent to which these initial recalls affected the resultant serial position curves. Our hypothesis was that if the dissimilarities between free recall and ISR are due primarily to differences in list length, then when the list length in ISR-free was increased to list lengths typically associated with free recall, then the output orders and serial position curves found in both free recall and ISR-free might both be affected similarly by increasing list length. To the extent that participants switched from initiating recall with the first word to initiating recall with one of the last list items on both tasks with increasing list length, so we would find evidence for increased similarities between free recall and this variant of ISR.

Method

Participants. Forty volunteers participated in this experiment from the University of Essex. None had participated in Experiment 1.

Materials and Apparatus. The materials for each participant consisted of 438 words randomly selected from a set of 480 words taken from the Toronto Word Pool (Friendly, Franklin, Hoffman & Rubin, 1982). The materials were presented in 52-point Times New Roman font in the centre of a computer monitor.

Design. The experiment used a mixed design. The type of task was manipulated between subjects. There were two within-subjects independent variables: list length, with eleven levels (1, 2, 3, 4, 5, 6, 7, 8, 10, 12, and 15) and serial position (SP) with up to 15 levels. The dependent variable for free recall was the proportion of words recalled (in any order) and the dependent variable for ISR-free was the proportion of words recalled (in the correct serial position).

Procedure. Participants were presented with a total of 66 trials consisting of 2 blocks of free recall trials, or two blocks of trials using the ISR-free task. In each block there was 33 experimental trials, which consisted of three trials of each of the eleven different list lengths. The order of the list lengths within each block was randomised. Participants were tested individually, and the first block was preceded by task-specific instructions and a practice list on that task of list length 7.

Each trial started with a warning tone, followed after 3s by a sequence of between 1 and 15 words presented one at a time in the centre of the screen. The presentation rate was 1 word every second, with each word displayed for 0.75s with an additional 0.25s inter-stimulus interval in which the stimulus field was blank. Participants were instructed to read each word aloud as it was presented. At the end of the list there was an auditory cue and an empty grid was displayed on the screen that contained the same number of rows as there were number of words on the current trial.

The participants wrote down as many words as they could remember in their lined response grids whilst saying out loud what they were writing down. In free recall, the participants were free to write the words down in any order that they wished, and filled up their response grids from the top of the grid. In ISR-free, the participants were free to write down the words in any temporal order that they wished but they were instructed to write down the words in the position in the response grid that corresponded to each word's serial position.

Results

The data for the two tasks were analysed separately, as each task used a different dependent variable.

Free Recall. Figure 6A shows the proportion of words recalled at each of the 11 different list lengths. Consistent with list length effects in free recall, the mean proportion of words recalled decreased monotonically, from .98 (list length 1) to .30 (list length 15), F (10, 190) = 387.2, MSE = .004, p <.001, and the mean number of words recalled increased from .98 (list length 1) to 4.52 (list length 15), F (10, 190) = 37.32, MSE = .848, p <.001.

--Figure 6 about here--

At short lists, such as list length 4, the proportion of words recalled was consistently at a high level (M = .94), with no significant effect of serial position, F (3, 57) = 2.26, p > .05. At longer lists, such as at list length 8, there was a significant effect of serial position, F (7, 133) = 17.27, MSE = .049, p <.001, reflecting significant primacy and extended recency effects, but there was significantly greater recall of the recency items than the primacy items.

Figure 7A shows that the first word recalled tended to be from serial position 1 when the list was short, but as the list length increased, there was an increased tendency that it would be one of the last four serial positions. The proportion of trials on which recall started with the first list item decreased with list length, F (10, 190) = 100.8, MSE = .028, p <.001. The mean slopes across list lengths 1-3, 4-7, and 8-15 differed significantly from each other, F (2, 38) = 34.35, MSE = .005, p <.001, and were -0.02, -0.17, and -0.03, respectively, and confirmed that the function declines most steeply across list lengths 4-7, where the addition of each extra list item reduced the proportion of trials starting with serial position 1 by on average 17%.

--Figure 7 about here--

We then examined the effect of first recall on the resultant serial position curves. Figure 8A shows the serial position curves for the free recall trials in which the first word recalled was from serial position 1 plotted using ISR scoring. Figure 8B shows the serial position curves for those same trials plotted using free recall scoring. Finally, Figure 8C shows the serial position curves for the free recall trials in which the first word recalled was from one of the last four serial positions plotted using free recall scoring. Replicating the findings of Experiment 1, the shapes of the serial position curves in Figure 8A, 8B, and 8C are strikingly different.

We again limit our analyses to performance at a representative list length, list length 6. Nineteen participants initiated recall on one or more trials with serial position 1 at this list length. Using ISR scoring, the mean recalls across the six serial positions in Figure 8A, are

1.00, .44, .26, .19, .12, and .08, and discounting serial position 1 (whose mean by definition is 1.00), serial positions 2-6 differed significantly, $F(4, 72) = 9.51$, $MSE = .040$, $p < .001$; pairwise comparisons confirming that there is significant extended primacy effects with ISR scoring.

The corresponding mean recalls across the same trials using free recall scoring (Figure 8B) are 1.00, .59, .61, .55, .63, and .72. Discounting serial position 1, serial positions 2-6 did not differ significantly, $F(4, 72) = 0.79$, $MSE = .096$, $p > .05$, but remain at a relatively elevated level of recall.

Finally, consider performance by the 17 participants who initiated recall on one or more trials with one of the last four serial positions. Figure 8C shows that the mean recalls across these six serial positions are .48, .16, .67, .71, .84, and .86; means which differed significantly, $F(5, 80) = 11.15$, $MSE = .104$, $p < .001$; pairwise comparisons confirming 1-item primacy, and significant extended recency effects.

--Figure 8 about here--

ISR. Figure 6B shows the proportion of words recalled at each of the 11 different list lengths. The mean proportion of words recalled decreased monotonically from .99 (list length 1) to .18 (list length 15), $F(10, 190) = 389.9$, $MSE = .006$, $p <.001$, but the mean number of words recalled first increased from .99 (list length 1) to 3.58 (list length 4) and then declined to 2.52 (list length 15).

At short lists, such as list length 4, the proportion of words recalled was at a high level ($M = .89$), but there was nonetheless a significant effect of serial position, $F(3, 57) = 3.51$, $p < .05$, reflecting a significant difference between serial positions 1 and 3. At longer lists, such as at list length 8, there was a significant effect of serial position, $F(7, 133) =$

24.45, *MSE* = .051, *p* <.001, reflecting significant primacy and extended recency effects, but there was significantly greater recall of the recency items than the primacy items.

Figure 7B shows that the first recall data in the ISR task resembles that from the free recall task. In ISR, the first word recalled also tended to be from serial position 1 when the list was short, but as the list length increased, so there was an increased tendency that it would be one of the last four serial positions. The proportion of trials on which recall started with the first list item decreased with list length, *F* (10, 190) = 101.1, *MSE* = .029, *p* <.001. The mean slopes across list lengths 1-3, 4-7, and 8-15 differed significantly from each other, *F* (2, 38) = 37.35, *MSE* = .005, *p* <.001, and were -0.02, -0.18, and -0.02, respectively, and confirmed that the function declines most steeply across list lengths 4-7, where the addition of each extra list item reduced the proportion of trials starting with serial position 1 by on average 18%.

Figure 8D shows the serial position curves for the ISR-free trials in which the first word recalled was from serial position 1, Figure 8E shows the same data scored by free recall, and Figure 8F shows the serial position curves for the ISR trials in which the first word recalled was from one of the last four serial positions (using ISR scoring). The shapes of the serial position curves in Figure 8D, 8E, and 8F are again very different.

Figure 8D shows that there is extended primacy with reduced recency when the first word is recalled first using the ISR scoring. Considering a representative list length, list length 6, the means were .96, .57, .40, .23, .25, and .55, across the six serial positions (note that the value at serial position 1 need not be exactly 1.00 in this analysis, since although the first word was always output first, it need not always be placed in the correct serial position). A within-subjects ANOVA on the 17 participants who started one or more trials with serial position 1 showed that these means differed significantly, *F* (5, 80) = 14.87, *MSE*

= .082, $p < .001$, pairwise comparisons confirmed that there were significant extended primacy effects and 1-item recency.

Figure 8E shows that there is consistently elevated recall of early list items when these same ISR-free data are re-scored using free recall scoring. The means at list length 6 were 1.00, .61, .57, .43, .52, and .67, means of serial positions 2-6 did not differ significantly F (4, 64) = 1.19, MSE = .123, $p > .05$.

Figure 8F shows that when recall was initiated with one of the last four words, the ISR-free data show extended recency with reduced primacy using the ISR scoring. For example, at list length 6, the means across the six serial positions were .21, .20, .30, .42, .70, and .90, and the presence of extended recency effects were confirmed by a within-subjects ANOVA (and subsequent pairwise comparisons) on the 18 participants who started one or more trials with one of the last four serial positions, F (5, 80) = 23.99, MSE = .062, $p < .001$.

Lag analyses

Finally, we examined the extent to which the output orders showed evidence of forward ordered recall. Figure 9 plots the CRP for lag +1 responses as a function of list length for free recall and ISR-free. Note that neither task requires participants to output successive items in the same order as at input. The CRP-values for lag+1 responses were analysed by a 2 (task: free recall and ISR-free) x 10 (List length: 2, 3, 4, 5, 6, 7, 8, 10, 12, 15) mixed ANOVA, which revealed a non-significant main effect of task, F (1, 38) = 0.61, MSE = .069, $p > .05$, a significant main effect of list length, F (9, 342) = 121.5, MSE = .021, $p < .001$, and a significant interaction, F (9, 342) = 2.04, MSE = .021, $p < .05$. An exploration of the interaction revealed that there non-significant differences between the two tasks for nine of the list lengths, but a significant difference at a single list length in favour of free recall over ISR-free at list length 12, $t(38) = 2.16$, $p <.05$. Thus there is a similar tendency for forward ordered recall to decrease with increasing list length for both tasks, with a similar degree of forwards order on both tasks, with the exception of list length 12 where there is a slightly greater tendency for forwards recall in the free recall conditions. Note that even at the longer list lengths there remains considerable forwards order recall (CRP values for lag +1 remain greater than .30) for both tasks, even though there is no formal requirement to output in forwards order in the two tasks.

--Figure 9 about here--

Discussion

Experiment 2 provided further evidence for the need for greater theoretical integration between free recall and ISR. Replicating the main findings of Experiment 1, Experiment 2 found “ISR-like” features in free recall when it was performed with shorter “ISR-like” list lengths. In addition, Experiment 2 found some “free recall-like” features in serial recall when ISR-free was performed under “free recall-like” list lengths.

Specifically, in both tasks, participants started their recalls with the first word on the list when the list was short, but as the list length increased so there was an increased tendency to start recall with one of the last four words. Moreover, in both tasks, the initial response had a large effect on the subsequent serial position curves: heightened recall of early list positions occurred with limited recency when recall began with serial position 1, but extended recency with limited primacy was observed when recall began with one of the last four items. Furthermore, in both tasks the output order reflected the input order: there was a high probability of Lag +1 transitions in both tasks, especially for shorter list lengths.

EXPERIMENT 3

Experiment 2 compared free recall performance with performance on ISR-free, a variant of ISR. An advantage of using ISR-free is that we can see which words participants would choose to output first in a serial recall test. It therefore speaks to which words are most accessible at the time of test during ISR (see Tan & Ward, 2007). However, it remains to be seen whether the list length effects observed on the output orders and serial position curves using ISR-free would also be found when the participants were instructed to perform a more standard ISR task.

In Experiment 3, the effects of list length on the output orders and serial position curves were examined using a between-subjects design that examined a standard ISR task and the ISR-free task (used by Tan & Ward, 2007; and Experiment 2). A total of forty participants were presented with 66 lists of words for serial recall. Half the participants were instructed using the ISR-free instructions, and the other half were instructed using more standard ISR instructions. Specifically, we adopted the instructions used by Golomb, Peelle, Addis, Kahana, and Wingfield (2008) who examined serial recall of 10-word lists. Golomb et al encouraged participants to begin recall with the first list item; however, because of the supraspan list length, they instructed participants that if they were unable to retrieve the first item, they should begin their recall with the earliest item that they could remember.

Method

Participants. Forty volunteers participated in this experiment from the University of Essex. None had participated in any of the earlier two experiments.

Materials and Apparatus. The materials were the same as those used in Experiment 2.

Design. The experiment used a mixed design. The type of task was manipulated between subjects with two levels: standard ISR and ISR-free. There were also two within-subjects independent variables, list length, with eleven levels (1, 2, 3, 4, 5, 6, 7, 8, 10, 12, and 15) and serial position (SP) with up to 15 levels. The dependent variable was the proportion of words recalled in the correct serial position.

Procedure. The procedures were identical to those used in Experiment 2, with the exception that half the participants received two blocks of the ISR-free task and the other half received two blocks of the standard ISR task. Participants in the ISR-free task spoke their responses out loud, so that their output order could be recorded; participants in the standard ISR task wrote down their responses in forwards serial order and they did not speak their responses out loud.

Results

Figure 10 shows the proportion of words recalled in the correct serial position for the standard ISR task (Figure 10A) and the ISR-free task (Figure 10B) at each of the 11 different list lengths. A 2 (task) x 11 (list length) mixed ANOVA was performed on the mean ISR scores for the two groups at each list length. This revealed a significant main effect of list length, F (10, 380) = 666.5, MSE = .016, p <.001, a non-significant main effect of task, F (1, 38) = 2.80, MSE = .033, p >.05, and a significant 2-way interaction, F (10, 380) = 666.5, MSE = .016, p <.001. The two way interaction occurred because at list lengths 2 to 4 there were small but non-significant advantages for participants performing standard ISR compared with ISR-free, but the pattern was reversed at list lengths greater than 4, and at list lengths of 10 and greater, there was a small but significant increase in performance in ISR-free relative to standard ISR.

--Figure 10 about here--

We examined serial position curves on the two tasks at three representative list lengths: 4, 8, and 15. A 2 (task) x 4 (serial position) mixed ANOVA revealed a non-significant main effect of task, F (1, 38) = 0.94, MSE = .047, $p > .05$, a significant main effect of serial position, F (3, 114) = 7.15, MSE = .019, $p > .05$, and a non-significant interaction, F (3, 114) = 0.88, MSE = .019, $p > .05$. For both tasks there was significant primacy and recency effects at list length 4, and there was little difference in the shapes of the serial position curves of the two tasks.

A 2 (task) x 8 (serial position) mixed ANOVA revealed a non-significant main effect of task, F (1, 38) = 0.83, MSE = .082, $p > .05$, a significant main effect of serial position, F (7, 266) = 65.81, MSE = .041, $p < .001$, and a significant interaction, F ((7, 266) = 3.60, MSE = .041, $p < .01$. There were significant primacy and recency effects in both tasks, with greater recency than primacy. However, there were small but non-significant recall advantages at serial positions 1-4 in favour of the standard ISR task, there were small but non-significant recall advantages at serial positions 5-7 in favour of the ISR-free task, and there was a significant recall advantage for the ISR-free task at serial position 8.

A 2 (task) x 15 (serial position) mixed ANOVA revealed a significant main effect of task, F (1, 38) = 8.14, MSE = .065, $p < .01$, a significant main effect of serial position, F (14, 532) = 128.97, MSE = .022, $p < .001$, and a non-significant interaction, F (14, 532) = 1.42, MSE = .022, $p > .05$. There were significant primacy and recency effects in both tasks, with greater recency than primacy, and greater recall performance overall with ISR-free than standard ISR.

Second, we examined which words from a list were the first to be recalled on each trial. Figure 11A (standard ISR) and Figure 11B (ISR-free) show that in the vast majority of all trials with short list lengths, recall started with the first word in the list, but as the list length increases so this tendency decreases for both tasks, but more rapidly with ISR-free than standard ISR. This pattern was confirmed in a 2 (task: ISR-free and standard ISR) x 10 (list lengths: 2-8, 10, 12, 15) mixed ANOVA performed on the proportion of trials starting with serial position 1. This revealed a significant main effect of task F (1, 38) = 26.25, MSE = .123, p < .001, a significant main effect of list length, F (9, 342) = 214.7, MSE = .025, p < .001, and a significant interaction, F (9, 342) = 6.09, MSE = .025, p < .001.

Figures 11A and 11B also show that at longer list lengths there is an increased tendency to initiate recall with words from one of the last four serial positions, and that this tendency increases more rapidly with ISR-free than standard ISR. This pattern was also confirmed in a 2 (task: ISR-free and standard ISR) x 10 (list lengths: 2-8, 10, 12, 15) mixed ANOVA performed on the proportion of trials starting with one of the last four serial positions. This revealed a significant main effect of task F (1, 38) = 36.48, MSE = .189, p < .001, a significant main effect of list length, F (9, 342) = 124.2, MSE = .026, p < .001, and a significant interaction, F (9, 342) = 14.1, MSE = .026, p < .001.

--Figure 11 about here--

We split the serial position curve data by whether participants initiated recall with the first word or one of the last 4 words. Figure 12A shows the serial position curves for the standard ISR trials in which the first word recalled was from serial position 1, Figure 12B shows the same data plotted by free recall scoring, and Figure 12C shows the serial position curves for the standard ISR trials in which the first word recalled was from one of the last four serial positions. As in earlier studies, the shapes of the serial position curves in Figure 12A, 12B, and 12C are different: there is an increased tendency for heightened primacy with reduced recency when the first word is recalled first in Figure 12A, elevated levels of recall on early items I Figure 12B, but extended recency with reduced primacy in Figure 12C when one of the last 4 words is first recalled.

Figures 12D, 12E and 12F show the equivalent data split for the ISR-free task. Replicating the ISR-free data of Experiment 2 (Figures 8D, 8E, and 8F), there is an increased tendency for heightened primacy with reduced recency when the first word is recalled first in Figure 12D, elevated recall in Figure 12E but extended recency with reduced primacy in Figure 12E when one of the last 4 words is first recalled.

One striking feature of these data is the similarity between the two tasks when the first response is controlled for. The data at list length 6 was examined across the two tasks using ISR-scoring on the data in Figures 12A and 12D in which recall initiated with serial position 1. A total of 17 participants from the ISR-free task and all 20 participants from the standard ISR task contributed to this analysis. This revealed a non-significant main effect of task, F (1, 35) = 0.13, MSE = .116, p > .05, a significant main effect of serial position, F (5, 175) = 37.03, MSE = .070, p < .001, and a non-significant interaction, F (5, 175) = 0.13, MSE = .070 p > .05. Thus, once equated for starting with serial position 1, there was similar performance on both tasks. Pair-wise comparisons revealed that there were extended primacy and 1-item recency.

When the same list length 6 data were compared using free recall scoring (Figures 12B and 12E) over serial positions 2 to 6, there were highly similar findings. The ANOVA revealed that there was no significant main effect of task, F (1, 35) = 0.07, MSE = .111, p > .05, a significant main effect of serial position, F (4, 140) = 6.85, MSE = .097, p < .001, and a non-significant interaction, F (5, 175) = 0.92, MSE = .097 p > .05. Pairwise comparisons revealed significant primacy and recency.

--Figure 12 about here--

Finally, when the list length 6 data were compared using ISR scoring (Figures 12C and 12F) over serial positions 1 to 6, was a significant main effect of task, F (1, 30) = 5.22, MSE = .180, $p < .05$, a significant main effect of serial position, F (5, 150) = 40.94, MSE = .074, $p < .001$, and a non-significant interaction, F (5, 150) = 1.28, MSE = .074 $p > .05$. Both tasks showed significant extended recency effects, but ISR was superior overall for ISR-free,

Lag analyses

Finally, we examined the extent to which the output orders showed evidence of forward ordered recall. Figure 13 plots the CRP for lag +1 responses as a function of list length for standard ISR and ISR-free. The CRP-values for lag+1 responses were analysed by a 2 (task: standard ISR and ISR-free) x 10 (List length: 2, 3, 4, 5, 6, 7, 8, 10, 12, 15) mixed ANOVA, which revealed a non-significant main effect of task, F (1, 37) = 0.09, MSE = .088, $p > .05$, a significant main effect of list length, F (9, 333) = 66.24, MSE = .020, $p < .001$, and a non-significant interaction, F (9, 333) = 1.50, MSE = .020, $p > .05$. Thus there is a similar tendency for forward ordered recall to decrease with increasing list length for both tasks, with a similar degree of forwards order on both tasks. Note that even at the longer list lengths there remains considerable forwards order recall (CRP values for lag +1 remain greater than .50) for both tasks, even though there is no formal requirement to output in forwards order in ISR-free.

--Figure 13 about here--

Discussion

Experiment 3 essentially replicated the main findings using ISR-free that were observed in Experiment 2. When the list was short, it was common to begin ISR-free with the first list item in short lists (even though they did not have to), but as the list length was increased so there was an increased tendency for ISR-free to begin with one of the last four list items. A similar albeit reduced tendency was also observed in the standard ISR task. The main differences between the tasks were that participants in the standard ISR task started their recall more often with the first item in the standard ISR condition (as they were instructed) than in the ISR-free task, and when they could not recall the first item in the standard ISR task, they started with an earlier item (often scored as an "other" serial position at longer list lengths) more often in the standard ISR task than in the ISR-free task. These data are consistent with the standard ISR task instructions, but show that when participants are free to output in any order, their initial recalls resemble standard ISR for shorter lists but at longer list length the tendency to output one of the last four words in the list becomes more exaggerated.

Interestingly, the effects of start position on serial position curves are very similar: greater primacy is found in both tasks when recall starts with the first word in the list, but greater recency is found when recall was from one of the last four words.

EXPERIMENT 4

Experiment 3 compared ISR with ISR-free and showed that in both tasks there was a reduction in initiating recall with the first list item as the length of the list increased. Experiment 4 examined whether a similar pattern of data could be observed using different variants of the reconstruction of order task. In reconstruction of order tasks, participants are presented with sequences of words for immediate recall as in the previous experiments, but at test all the list items are re-presented in a random order and participants must place each word into the correct serial position.

One advantage of the task is that it relies more heavily on order information (the items themselves are re-presented at test). A second advantage is that it is well suited for examining the effects of list length and output order. In Experiment 4, we examined performance at the standard version of the reconstruction of order task, in which the that participants must allocate the represented words into the grid starting with the word that is to be allocated to serial position 1 and then proceeding in a strict forwards order. We compared performance on this condition, with free reconstruction of order (see Lewandowsky, Brown & Thomas, 2009), in which participants are free to allocate the re-presented words into their positions in the grid in whatever order that they like. Through the positioning and order of responses, the task allows the recording of both serial position and output order without the need to ask participants to speak out loud during one or both conditions.

Method

Participants. Thirty volunteers participated in this experiment from the University of Essex. None had participated in any of the earlier three experiments.

Materials and Apparatus. The materials were the same as those used in Experiment 3.

Design. The experiment used a mixed design. The type of task was manipulated between subjects with two levels: standard Reconstruction of order and free Reconstruction of order. There were also two within-subjects independent variables: list length, with eleven levels (1, 2, 3, 4, 5, 6, 7, 8, 10, 12, and 15) and serial position (SP) with up to 15 levels. The dependent variable was the proportion of words recalled in the correct serial position.

Procedure. The method of presentation of the stimuli during study was identical to that used in Experiments 2 and 3. However, the method of testing differed. After the last word had been presented, the participants were presented with all the list items arranged in a new random order in a column on the left of the screen, and a response grid of the same type as presented in Experiments 1 to 3, with 2 columns and as many rows as words on the list. In the first column of the response grid in ascending value were the numbers 1 to list length. The second column of the response grid was left blank. The participants were required to fill the response grid by first clicking on the word to be entered (from the column on the left) and then clicking on the location in the response grid corresponding to that word's list position on the right.

The computer programme in the standard Reconstruction of order task only allowed participants to fill the response grid from top to bottom, that is, in strict serial order. By contrast, participants in the free Reconstruction of order task could click on any word and click on any location. It was not possible to change the responses once they had been entered. When the response grid was completed, participants pressed a computer button on the screen to continue.

Results

Figure 14 shows the proportion of words recalled in the correct serial position for the standard Reconstruction of order task (Figure 14A) and the free Reconstruction of order task (Figure 14B) at each of the 11 different list lengths. The graphs exclude the data from a small minority of trials (21 trials out of a total of 2244, less than 1%) in which due to computer software failure, participants were allowed to make inappropriate responses (responding with the same item in two locations). A 2 (task) x 11 (list length) mixed ANOVA was performed on the mean reconstruction of order scores for the two groups at each list length. This revealed a significant main effect of task, F (1, 32) = 18.61, MSE = .042, p <.001, a significant main effect of list length, F (10, 320) = 381.6, MSE = .008, p <.001, and a significant 2-way interaction, F (10, 320) = 7.17, MSE = .008, p <.001. The two-way interaction occurred because at list lengths 1 to 4 there were no significant differences in levels of recall for the two tasks, but at list lengths above five there were significant reconstruction of order advantages for the free conditions over the standard conditions.

--Figure 14 about here--

A 2 (task) x 4 (serial position) mixed ANOVA on the serial position curve at list length 4 revealed a non-significant main effect of task, F (1, 32) = 0.51, MSE = .039, p > .05, a significant main effect of serial position, F (3, 96) = 4.52, MSE = .007, p < .01, and a non-significant interaction, F (3, 96) = 0.32, MSE = .007, p > .05. For both tasks there was significant primacy and recency effects at list length 4, and there was little difference in the shapes of the serial position curves of the two tasks.

A 2 (task) x 8 (serial position) mixed ANOVA revealed a significant main effect of task, F (1, 32) = 14.31, MSE = .151, p < .001, a significant main effect of serial position, F (7, 224) = 25.93, MSE = .032, p < .001, and a significant interaction, F ((7, 224) = 5.84, MSE = .032, p < .01. There were significant primacy and recency effects with both tasks. However, although there was near equivalent performance on serial positions 1 to 3 on the two tasks, there were significant recall advantages at serial positions 4, 6, 7 and 8 for the free Reconstruction of Order task relative to the standard Reconstruction of Order task.

A 2 (task) x 15 (serial position) mixed ANOVA revealed a significant main effect of task, F (1, 32) = 15.40, MSE = .132, p < .001, a significant main effect of serial position, F (14, 448) = 29.88, MSE = .030, p < .001, and a significant interaction, F (14, 448) = 14.07, MSE = .030, p < .001. There were significant primacy and recency effects on both tasks, but whereas there were similar levels of primacy for the two tasks, there was significantly greater recency on the last three serial positions for the free reconstruction of order task.

Second, we examined which words from a list were the first to be recalled on each trial. Figure 15A (Standard Reconstruction of Order) and Figure 15B (Free Reconstruction of Order) show that in the vast majority of all trials with short list lengths, recall started with the first word in the list, but as the list length increases so this tendency decreased for both tasks, but more rapidly with Free Reconstruction of Order than with Standard Reconstruction of Order). This pattern was confirmed in a 2 (task: Standard Reconstruction of Order and Free Reconstruction of Order) x 10 (list lengths: 2-8, 10, 12, 15) mixed ANOVA performed on the proportion of trials starting with serial position 1. This revealed a significant main effect of task F (1, 32) = 25.37, MSE = .176, $p < .001$, a significant main effect of list length, F (9, 288) = 77.21, MSE = .032, $p < .001$, and a significant interaction, F (9, 288) = 5.81, MSE = .032, $p < .001$. There was no significant difference in the proportion of trials starting with serial position 1 at list lengths 2 t o 4, but from list length 5 there was a significant decrease in the proportion of trials starting with serial position 1 relative to the Standard reconstruction of order task.

Figures 15A and 15B also show that at longer list lengths there is an increased tendency to initiate recall with words from one of the last four serial positions (especially for the free reconstruction of order task) and from other serial positions (especially for the standard reconstruction of order task).

--Figure 15 about here--

Figure 16A shows the serial position curves for the free Reconstruction of order trials in which the first word recalled was from serial position 1, and Figure 16B shows the serial position curves for the free Reconstruction of order trials in which the first word recalled was from one of the other serial positions. As in earlier studies, the shapes of the serial position curves in Figures 16A and 16B are rather different: there is an increased tendency for heightened primacy when the first word is recalled first in Figure 16A, but more extended recency with reduced primacy in Figure 16B when one of the other words is first recalled.

For list length 6, the means for 17 participants in the free reconstruction of order task who initiated at least one trial with serial position 1 were .96, .77, .66, .48, .60, and .81 (ISR scoring), means which differed significantly, F (5, 80) =7.68, MSE = .064, p <.001. Pairwise comparisons revealed extended primacy and 1-item recency. The means for 15 participants in the free reconstruction of order task who initiated at least one trial with one of the last four serial positions were .55, .60, .37, .50, .73, and .98 (ISR scoring), means which differed significantly, F (5, 70) =7.70, MSE = .089, p <.001. Pairwise comparisons revealed significant recency effects and non-significant primacy effects.

--Figure 16 about here--

Lag analyses

Finally, we examined the extent to which the output orders showed evidence of forward ordered recall. Figure 17 plots the CRP for lag +1 responses as a function of list length for standard Reconstruction of order and free Reconstruction of order. The CRP-values for lag+1 responses were analysed by a 2 (task: standard Reconstruction of order and free Reconstruction of order) x 10 (List length: 2, 3, 4, 5, 6, 7, 8, 10, 12, 15) mixed ANOVA, which revealed a significant main effect of task, F (1, 32) = 19.08, MSE = .054, p < .001, a significant main effect of list length, F (9, 288) = 236.2, MSE = .014, p < .001, and a significant interaction, F (9, 288) = 3.85, MSE = .014, p < .001. An analyses of the interaction revealed that there was no significant difference in the CRP-values for Lag +1 between the two tasks at list lengths 2 to 5, but there was significantly heightened forwards ordered reconstruction in the free Reconstruction of order task for list lengths 6 and greater. Thus there is reduced forwards output with increasing list lengths for both tasks, but the reduction is steeper in the standard reconstruction of order task at longer list lengths.

--Figure 17 about here--

Discussion

Experiment 4 showed similar findings to Experiment 3, but used reconstruction of order tasks rather than serial recall tasks. At short list lengths participants tended to initiate recall with the first word in the list in both the standard reconstruction of order task and the free reconstruction of order task and this tendency decreased as the list length increased. As the list length increased, participants tended to start their reconstructions with one of the last 4 list items in the free reconstruction of order task, but in the standard reconstruction of order task, participants started with other list items.

We note that was less of a tendency to start with one of the last 4 conditions with increasing list length in the standard reconstruction of order task compared to the standard ISR task of Experiment 3, but this is likely to reflect participants' opportunity to omit early serial positions in the recall task of Experiment 3, whereas they were compelled to select items for all earlier serial positions before the last serial positions in the standard reconstruction of order task.

General Discussion

In all four experiments, participants were presented with lists of between 1 and 15 words for tests of immediate memory. In each experiment, participants tended to start their recall with the first word in the list for short lists, but as the list length was increased so there was a decreased tendency to start with the first list item. This tendency was found in the standard ISR task (Experiment 3) and the Standard Reconstruction of Order tasks (Experiment 4) where participants are instructed to start their recall with the first item. Critically, this tendency was also shown in the immediate free recall task (Experiments 1-2), the ISR-free task (Experiments 2-3), and the free reconstruction of order task (Experiment 4), even though these tasks do not require that the first word in the list should be output first.

As the list length increased so there was an increased tendency to start with one of the last serial positions when the participants were free to do so. Thus, in immediate free recall (Experiments 1-2), ISR-free (Experiments 2-3), and free reconstruction of order (Experiment 4), participants increasingly tended to recall one of the most recent four words first at increasing list lengths. This tendency was also observed (to a lesser extent) in standard ISR (Experiment 3), but although possible, such outputs necessarily resulted in all the early and middle list items being omitted. Such a strategy is not possible in the standard reconstruction of order task (Experiment 4), where participants are compelled to select the word corresponding to serial position 1 first.

In all tasks, the list length and the start position strongly influenced the shape of the resultant serial position curves: when recall started at serial position 1, there were elevated recall of early list items; when recall started towards the end of the list, there were extended recency effects. There was also a high tendency to recall in forwards serial order in all tasks, a tendency that decreased with increasing list length.

The similarities in these findings across the different tasks suggest that at least some of the differences observed between free recall and ISR result from the differences in the list lengths that are typically used in the two tasks. Thus, although free recall is typically associated with extended recency effects and recall starting with one of the last few items, such findings are specific to the longer list lengths that are typically used. Similarly, ISR is typically associated with extended primacy effects and recall starting with the first item, this finding is observed only at shorter list lengths.

Rather, our data across these different tasks suggest that there is a "natural" order in which to recall short lists of words, regardless of the exact task instructions: participants will tend to start recall with the first list item, even when this is not strictly necessarily, and continue to recall early list items. As the list length increases, so participants show reduced ability to recall (or reconstruct) the first list item when instructed, resulting in reduced primacy and elevated recency effects.

These findings offer challenges to models of free recall that assume that free recall is heavily recency-based (e.g., Davelaar et al, 2005; Howard & Kahana, 2002; Tan & Ward, 2000). The prevalence of recency in the free recall of longer lists (which are typically used) has perhaps overly focused theorising on explaining recency effects; primacy effects being relegated to control processes such as rehearsal or attention. Although, one could argue that the early list items benefit from rehearsal or attention in this study, we suspect that we would obtain similar forwards order recall with shorter list lengths under revised methodology (such as with speeded presentation rates, articulatory suppression or under divided attention, although these experiments have not as yet been performed).

We also believe that the findings also offer challenges to models of ISR that are distinct from free recall (Baddeley & Hitch, 1974; Baddeley, 1986, 2000; Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1992, 1999, 2006; Farrell & Lewandowsky, 2002; Henson, 1998; Lewandowsky & Farrell, 2008; Nairne, 1988; Oberauer and Lewandowsky, 2008; Page & Norris, 1998), especially given recent findings of similar effects of diagnostic variables such as presentation rate, word length and articulatory suppression on the two tasks in lists of length 6 and 8 (Bhatarah et al, 2009).

Rather, these findings suggest that there needs to be a closer relationship between theories of free recall and ISR. One interpretation of our findings is that there are distinct primacy-based mechanisms (typically associated with ISR) and recency-based mechanisms (typically associated with free recall), both of which operate at both tasks, but whose relative importance varies with increasing list length. According to this interpretation, the primacy-based mechanisms are increasingly dominant for shorter list lengths, but are supplanted at longer list lengths by recency-based mechanisms. According to this interpretation, theoretical integration between free recall is valuable in helping to explain the free recall and ISR of list lengths 4-9.

An alternative perspective is to see whether theories of free recall could be adjusted to take into account this heightened tendency to initiate recall with the first list item with short lists. If one assumed that the first list items were far more accessible in short lists due to increases in temporal distinctiveness, contextual discrimination, increased attention, or some other mechanism, then a complete theory of free recall may arguably go a long way to also explain ISR data. Evidence consistent with a common account of free recall and ISR includes the fact that the encoding strategy and the degree of forwards ordered recall can be identical in the two tasks (Bhatarah Ward, & Tan, 2008), and the patterns of rehearsal can be very similar on the two tasks (Bhatarah, Ward, Smith, & Hayes, 2009).

The data also goes some way to explain the relationship between free recall and rehearsal. We have for some time been interested in whether the processes underpinning rehearsal could also be the same as those underpinning recall (see also Laming, 2006, 2008). In a number of earlier studies, we have presented participants with lists of 16 or 20 words for free recall, and examined the patterns of rehearsals during study by asking the participants to rehearse out loud whatever earlier list items they were thinking about during the inter-stimulus intervals. One potential problem is why recall at the end of the list appears to be dominated by recency, but the rehearsal order early during study tends to be forwards-ordered, and show extended primacy effects (Tan & Ward, 2000; Ward, Woodward, Stevens, & Stinson, 2003). The data from the current experiment partially explains the differences between the primacy-based output orders for the early patterns of rehearsal and later recency-based patterns of recall. The patterns of rehearsal early in the study can be considered to be analogous to the recall of words from a short list: forwards serial recall, starting with the first list item. By contrast, the patterns of recall at the end of the list are from a longer list, where recall tends to be dominated by recalling first a recently experienced list item. It would appear that the patterns of rehearsals observed early in study in free recall, closely resemble the patterns of recall of lists of words of similar length, providing evidence supporting the relationship between recall and rehearsal.

In summary, this study represents an initial investigation into output order effects and serial position curves in an immediate memory task using a wide range of list lengths. The data suggest that many theories of free recall currently underestimate the role of primacy at shorter list lengths, that theories of ISR underestimate the role of recency at longer list lengths, and that a greater knowledge of these effects might strengthen our understanding of the relationship between rehearsal and recall and free recall and ISR.

Footnote

[1] We wish to thank Mike Page for his persistence in pointing this out to us.

References

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38*, 341-380.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American, 225*, 82-90.

Baddeley, A. D. (1976). *The Psychology of memory.* New York: Basic Books.

Baddeley, A.D. (1986). *Working memory*. Oxford: Clarendon Press.

Baddeley, A.D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *Recent advances in learning and motivation, Vol. 8.* (pp. 47-90). London: Academic Press.

Baddeley, A.D., & Hitch, G. J. (1977). Recency re-examined. In S. Dornic, (Ed.), *Attention and Performance VI*. (pp. 647-667). Hillsdale, N.J.: Erlbaum.

Bhatarah, P., Ward, G., & Tan, L. (2006). Examining the relationship between immediate serial recall and free recall: the effect of concurrent task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 215-229.

Bhatarah, P., Ward, G., & Tan, L. (2008). Examining the relationship between free recall and immediate serial recall: The serial nature of recall and the effect of test expectancy. *Memory & Cognition, 36*, 20-34.

Bhatarah, P., Ward, G., Smith, J., & Hayes, L. (2009). Examining the relationship between free recall and immediate serial recall: Similar patterns of rehearsal, and similar effects of word length, presentation rate, and articulatory suppression. *Memory & Cognition*, *36*, 20-34.

Botvinick, M. M., & Plaut, D. C. (2006). Short-term memory for serial order: A recurrent neural network model. *Psychological Review*, *113*, 201-233.

Brown, G. D. A., Chater, N., & Neath, I. (2008). Serial and free recall: Common effects and common mechanisms? A reply to Murdock (2008). *Psychological Review, 115,* 781-785.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*, 539-576.

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

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

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

Burgess, N. and G. J. Hitch (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language. 55*, 627-652.

Crowder, R. G. (1969). Behavioral strategies in immediate memory. *Journal of Verbal Learning & Verbal Behavior*. *8*, 524-528.

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effects. *Psychological Review, 112*, 3 - 42.

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

Farrell, S. & Lewandowsky, S. (2008). Empirical and theoretical limits on lag-recency in free recall. *Psychonomic Bulletin & Review, 15*, 1236-1250.

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). Norms for the Toronto Word Pool: Norms for imagery, concreteness, orthographic variables and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation, 14*, 375-399.

Glanzer, M. (1972). Storage mechanisms in recall. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory, Vol. V* (pp. 129-193). New York: Academic Press.

Golomb, J. D., Peelle, J. E., Addis, K. M., Kahana, M. J., & Wingfield, A. (2008). Effects of adult aging on utilization of temporal and semantic associations during free and serial recall. *Memory & Cognition, 36*, 947–956.

Grossberg, S. & Pearson, L. R. (2008). Laminar cortical dynamics of cognitive and motor working memory, sequence learning and performance: Toward a unified theory of how the cerebral cortex works. *Psychological Review, 115*, 677-732.

Henson, R.N.A. (1998). Short-term memory for serial order: The Start-End Model of serial recall. *Cognitive Psychology, 36*, 73-137.

Hogan, R. M. (1975). Interitem encoding and directed search in free recall. *Memory & Cognition, 3*, 197-209.

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1-19.

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology, 46*, 269-299.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition, 24*, 103-109.

Laming, D. (1999). Testing the idea of distinct storage mechanisms in memory. *International Journal of Psychology, 34*, 419-426.

Laming, D. (2006). Predicting free recalls. *Journal of Experimental Psychology: Learning,* *Memory, and Cognition, 32*, 1146-1163.

Laming, D. (2008). An improved algorithm for predicting free recalls. *Cognitive Psychology, 57*, 179-219.

Laming, D. (2009). Failure to recall. *Psychoological Review, 116*, 157-186.

Nairne, J. S. (1988). A framework for interpreting recency effects in immediate serial recall. *Memory & Cognition, 16*, 343-352.

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

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review, 115*, 544-576.

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

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review, 116*, 129–156.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of Associative Memory. *Psychological Review, 88*, 93 - 134.

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. *Psychological Review, 115*, 893–912.

Tan, L, & Ward, G. (2000). A recency-based account of primacy effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1589-1625.

Tan, L., & Ward, G. (2007). Output order in immediate serial recall. *Memory and Cognition, 35*, 1093-1106.

Ward, G. (2001). A critique of the working memory model. In J. Andrade, (Ed.), *Working Memory in Perspective,* 219-239. Hove, UK: Psychology Press.

Ward, G., Woodward, G., Stevens, A., & Stinson, C. (2003). Using overt rehearsals to explain word frequency effects in free recall. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *29,* 186-210.

Acknowledgements

We would like to thank Lydia Allen, Chung Ting Tse, Jennifer Townsend, and XXX for the collection of the data in this study. Rachel Grenfell-Essam is supported by a 1+3 competitive postgraduate studentship granted by the Economic & Social Research Council, UK under the supervision of the first author.

Table Captions

Table 1. The distribution of the first words recalled on each trial, as a function of the list length and the words' serial position. The italicised values represent the frequency of trials in which the first word recalled was from serial position 1, and the bold values represent the frequency of trials in which the first word recalled was from one of the last four serial positions.

Table 1.

Serial Position	List Length														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	*164*	*146*	*146*	*116*	*93*	*60*	*45*	*38*	*21*	*23*	*23*	*14*	*22*	*8*	*18*
2		**15**	**13**	**15**	**14**	11	9	8	10	6	7	6	3	3	5
3			**6**	**15**	**15**	**11**	7	7	3	2	8	1	3	4	1
4				**18**	**18**	**7**	**8**	6	9	1	2	4	4	4	3
5					**25**	**35**	**18**	**17**	6	3	4	1	2	1	1
6						**41**	**35**	**24**	**11**	7	4	2	5	1	2
7							**43**	**27**	**22**	**6**	8	2	2	3	
8								**38**	**34**	**21**	**8**	2	4	5	
9									**48**	**44**	**18**	**18**	4	8	3
10										**52**	**35**	**20**	**7**	2	5
11											**48**	**44**	**27**	**7**	8
12												**51**	**33**	**20**	**9**
13													**48**	**42**	**25**
14														**56**	**30**
15															**55**
Incorrect or No response	1	4	0	1	0	0	0	0	1	0	0	0	1	1	0

Figure Captions

Figure 1. Data from the free recall task of Experiment 1 showing the serial position curves: the mean proportion of words recalled from each list length as a function of the serial position.

Figure 2. Data from the free recall task of Experiment 1 showing the proportion of trials at each list length in which recall initiated with the first word in the list, one of the last four words in the list, or all other words in the list.

Figure 3. Data from the free recall task of Experiment 1. Serial position curves for trials which began with serial position 1 using ISR scoring (Figure 3A) and free recall scoring (Figure 3B), and for trials which began with one of the last four serial positions using free recall scoring (Figure 3C). In Figure 3C, the serial positions have been recency-justified.

Figure 4. Data from the free recall task of Experiment 1. Lag-CRP (Conditionalised response probabilities) curves for each list length. The lag refers to the difference in serial position between successive words recalled, such that smaller lags reflect the successive recall of words that were presented closer to each other on the list, and that positive values reflect pairs of words recalled in the same relative order as at presentation. The CRP represents the mean probability that a word of a particular lag was recalled. It is calculated by dividing the frequency of observed lag transitions by the number of legitimate opportunities in which words at each lag could be recalled.

Figure 5. Data from Experiment 1. The proportion of lag+1 responses as a function of list length.

Figure 6. Data from Experiment 2 showing the serial position curves for the free recall task (Figure 6A) and the ISR-free task (Figure 6B) for each list length.

Figure 7. Data from the free recall task (Figure 7A) and the ISR-free task (Figure 7B) of Experiment 2 showing the proportion of trials at each list length in which recall initiated with the first word in the list, one of the last four words in the list, or all other words in the list.

Figure 8. Data from Experiment 2. Serial position curves for free recall trials which began with serial position 1 using ISR scoring (Figure 8A) and free recall scoring (Figure 8B), and for trials which began with one of the last four serial positions using free recall scoring (Figure 8C). Figure 8D-8F show serial position curves for the ISR-free data for trials which began with serial position 1 using ISR scoring (Figure 8D) and free recall scoring (Figure 8E), and for trials which began with one of the last four serial positions using ISR scoring (Figure 8F). In Figures 8C and 8F, the serial positions have been recency-justified.

Figure 9. Data from Experiment 2. The proportion of lag+1 responses as a function of list length and task.

Figure 10. Data from Experiment 3 showing the serial position curves for the standard ISR task (Figure 10A) and the ISR-free task (Figure 10B) for each list length.

Figure 11. Data from the standard ISR task (Figure 11A) and the ISR-free task (Figure 11B) of Experiment 3 showing the proportion of trials at each list length in which recall initiated with the first word in the list, one of the last four words in the list, or all other words in the list.

Figure 12. Data from Experiment 3. Serial position curves for standard ISR trials which began with serial position 1 using ISR scoring (Figure 12A) and free recall scoring (Figure 12B), and for trials which began with one of the last four serial positions using free recall scoring (Figure 12C). Figure 12D-8F show serial position curves for the ISR-free data for trials which began with serial position 1 using ISR scoring

(Figure 12D) and free recall scoring (Figure 12E), and for trials which began with one of the last four serial positions using ISR scoring (Figure 12F). In Figures 12C and 12F, the serial positions have been recency-justified.

Figure 13. Data from Experiment 3. The proportion of lag+1 responses as a function of list length and task.

Figure 14. Data from Experiment 4 showing the serial position curves for the standard Reconstruction of Order task (Figure 14A) and the free Reconstruction of Order task (Figure 14B) for each list length.

Figure 15. Data from the standard Reconstruction of order task (Figure 15A) and the free Reconstruction of order task (Figure 15B) of Experiment 4 showing the proportion of trials at each list length in which recall initiated with the first word in the list, one of the last four words in the list, or all other words in the list.

Figure 16. Data from Experiment 4. Serial position curves for the Free reconstruction of Order task for trials which began with serial position 1 using ISR scoring (Figure 16A) and for trials which began with one of the last four serial positions using free recall scoring (Figure 16B). In Figure 16B, the serial positions have been recency-justified.

Figure 17. Data from Experiment 4. The proportion of lag+1 responses as a function of list length and task.

Figure 1

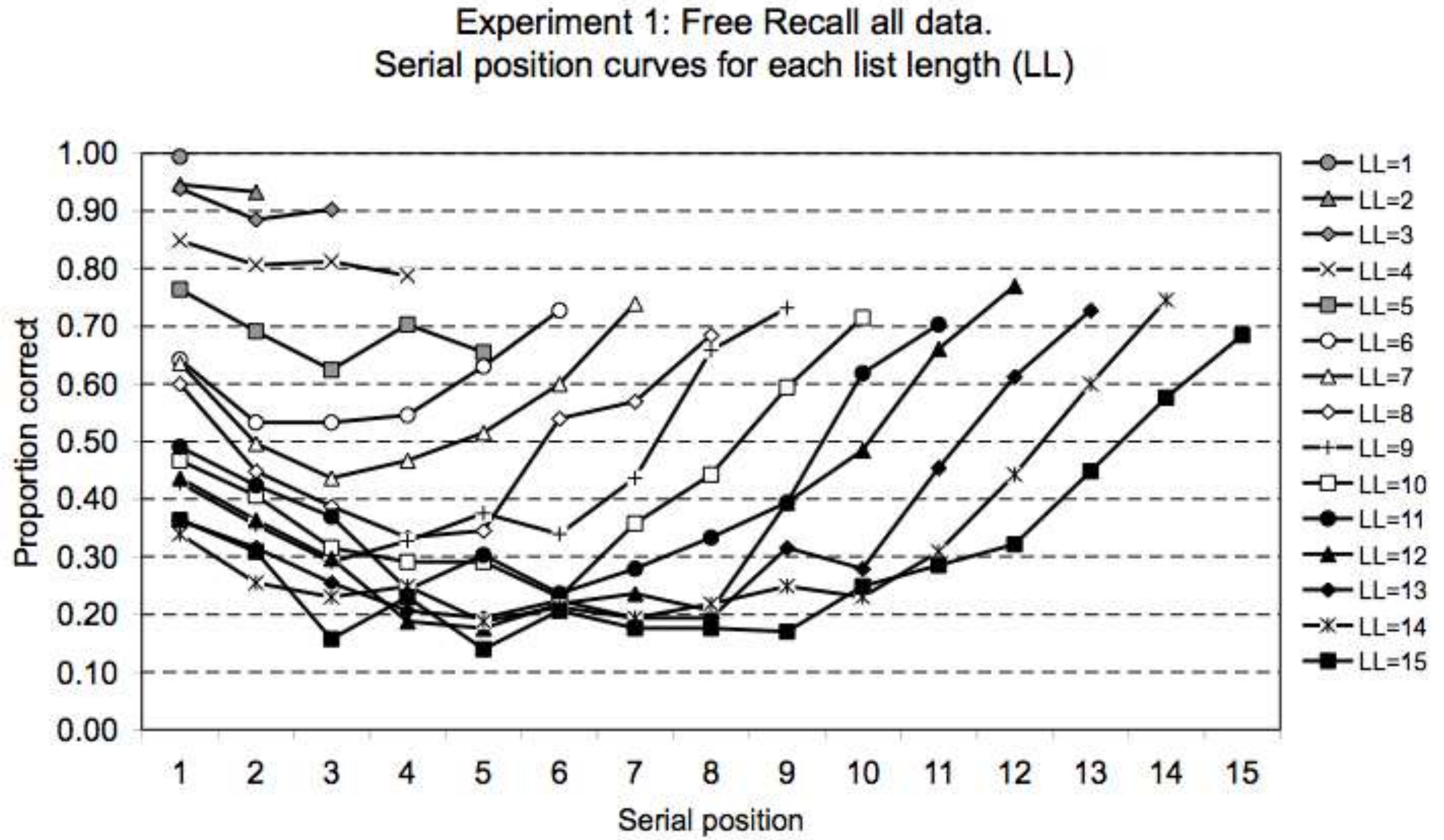

Experiment 1: Free Recall all data.
Serial position curves for each list length (LL)
Proportion correct
Serial position
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=9
LL=10
LL=11
LL=12
LL=13
LL=14
LL=15

Figure 2

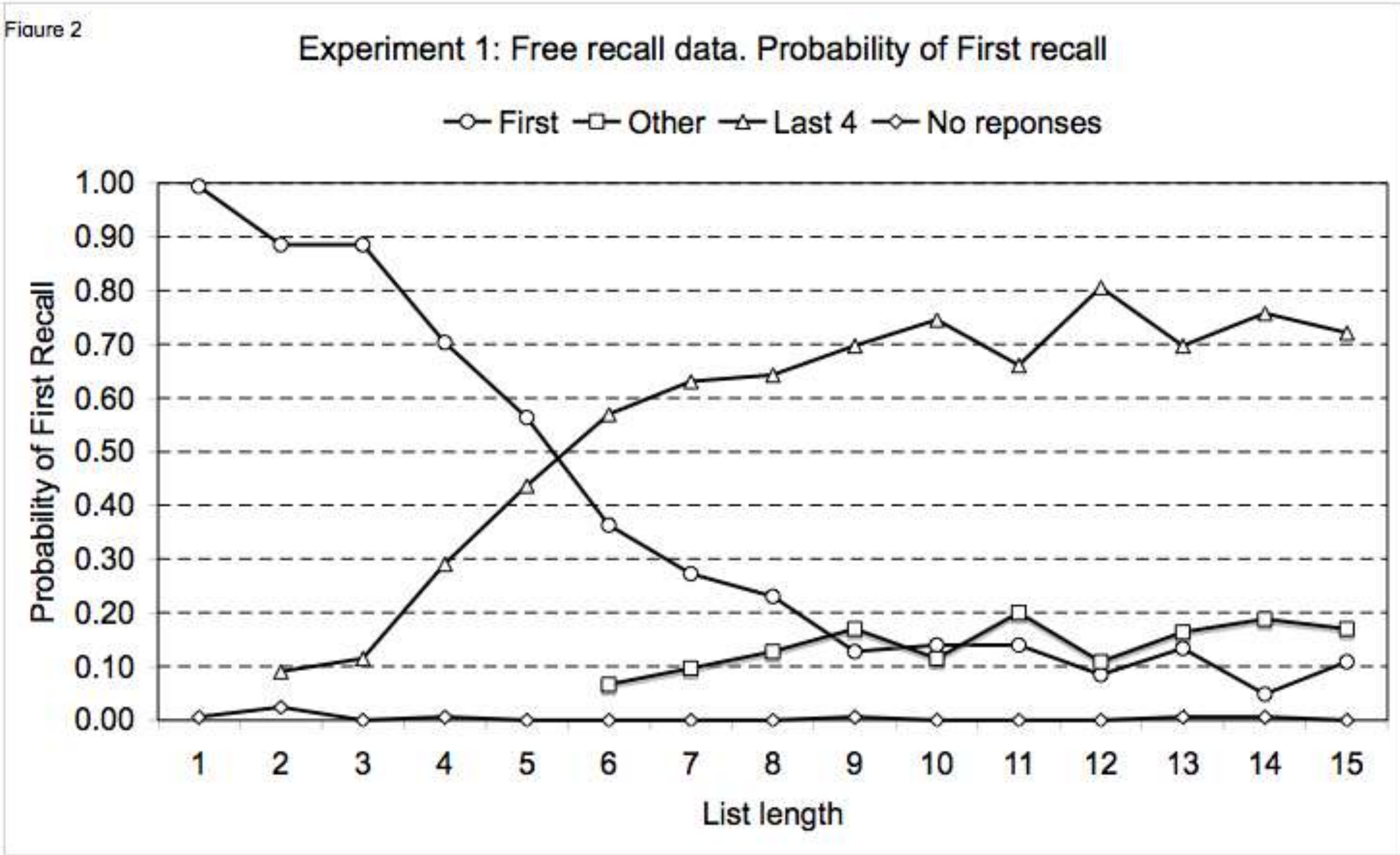
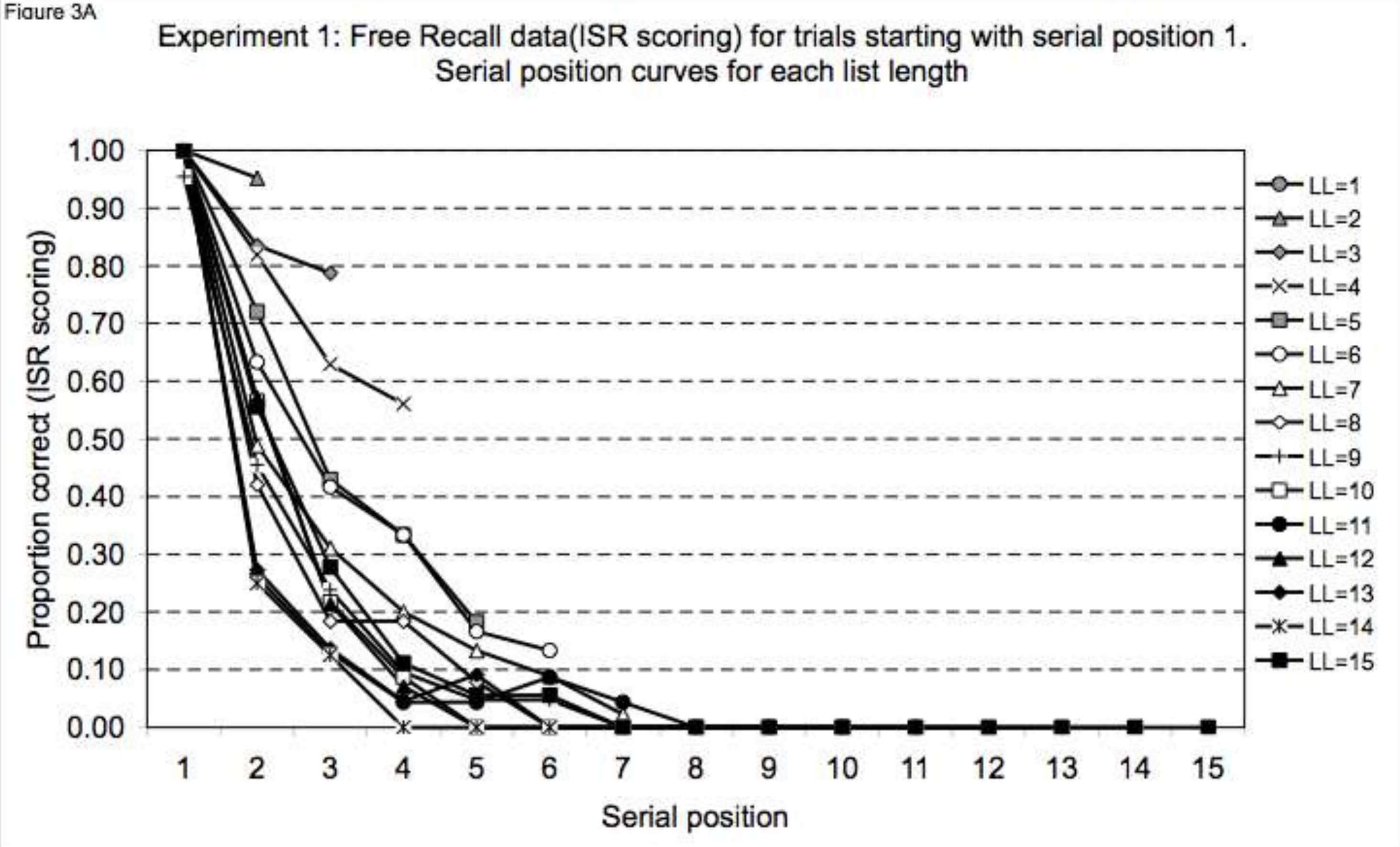

Experiment 1: Free recall data. Probability of First recall
First
Other
Last 4
No reponses
Probability of First Recall
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
List length

Figure 3A


Experiment 1: Free Recall data(ISR scoring) for trials starting with serial position 1.
Serial position curves for each list length
Proportion correct (ISR scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=9
LL=10
LL=11
LL=12
LL=13
LL=14
LL=15

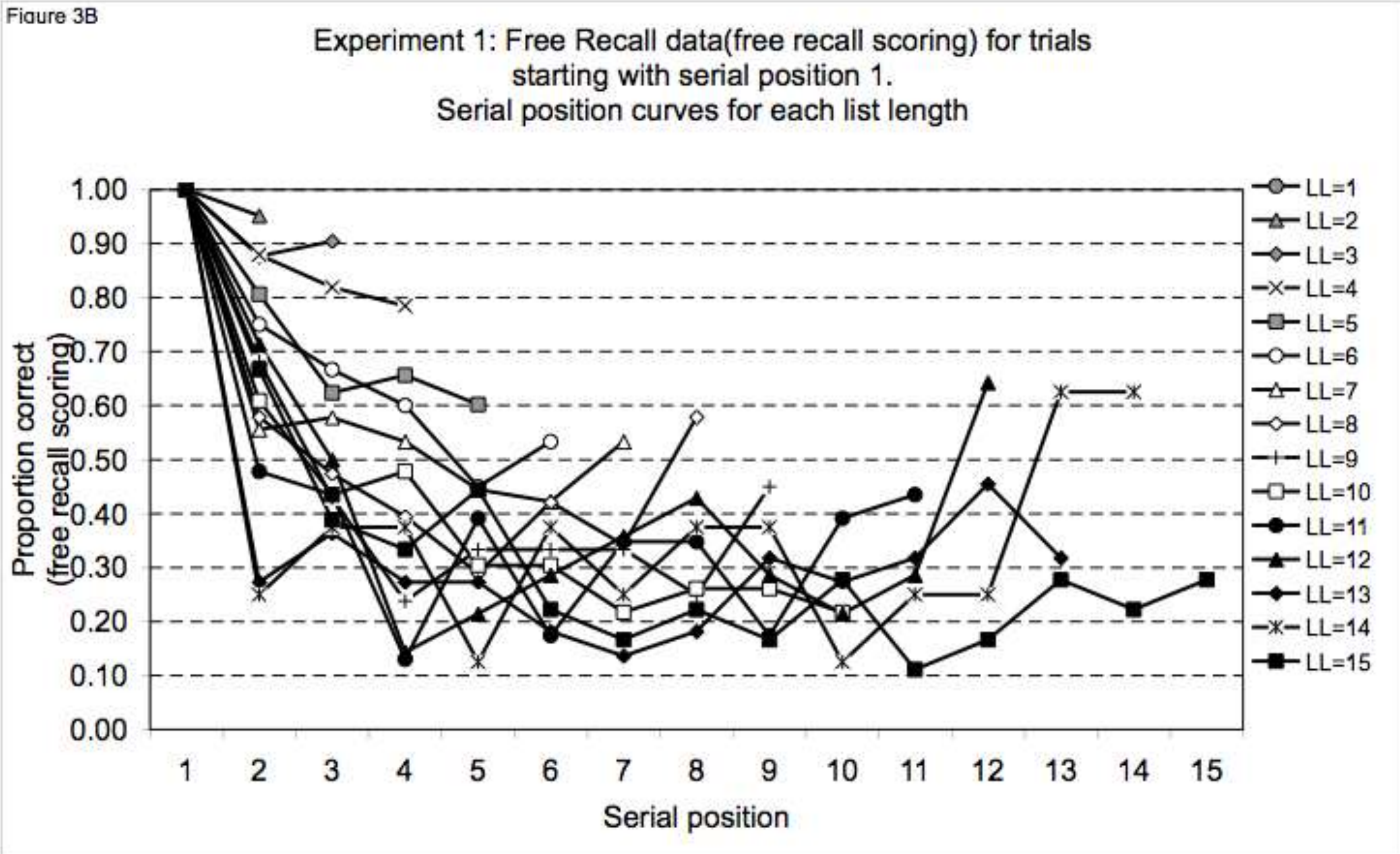


Figure 3B
Experiment 1: Free Recall data(free recall scoring) for trials starting with serial position 1.
Serial position curves for each list length
Proportion correct (free recall scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=9
LL=10
LL=11
LL=12
LL=13
LL=14
LL=15

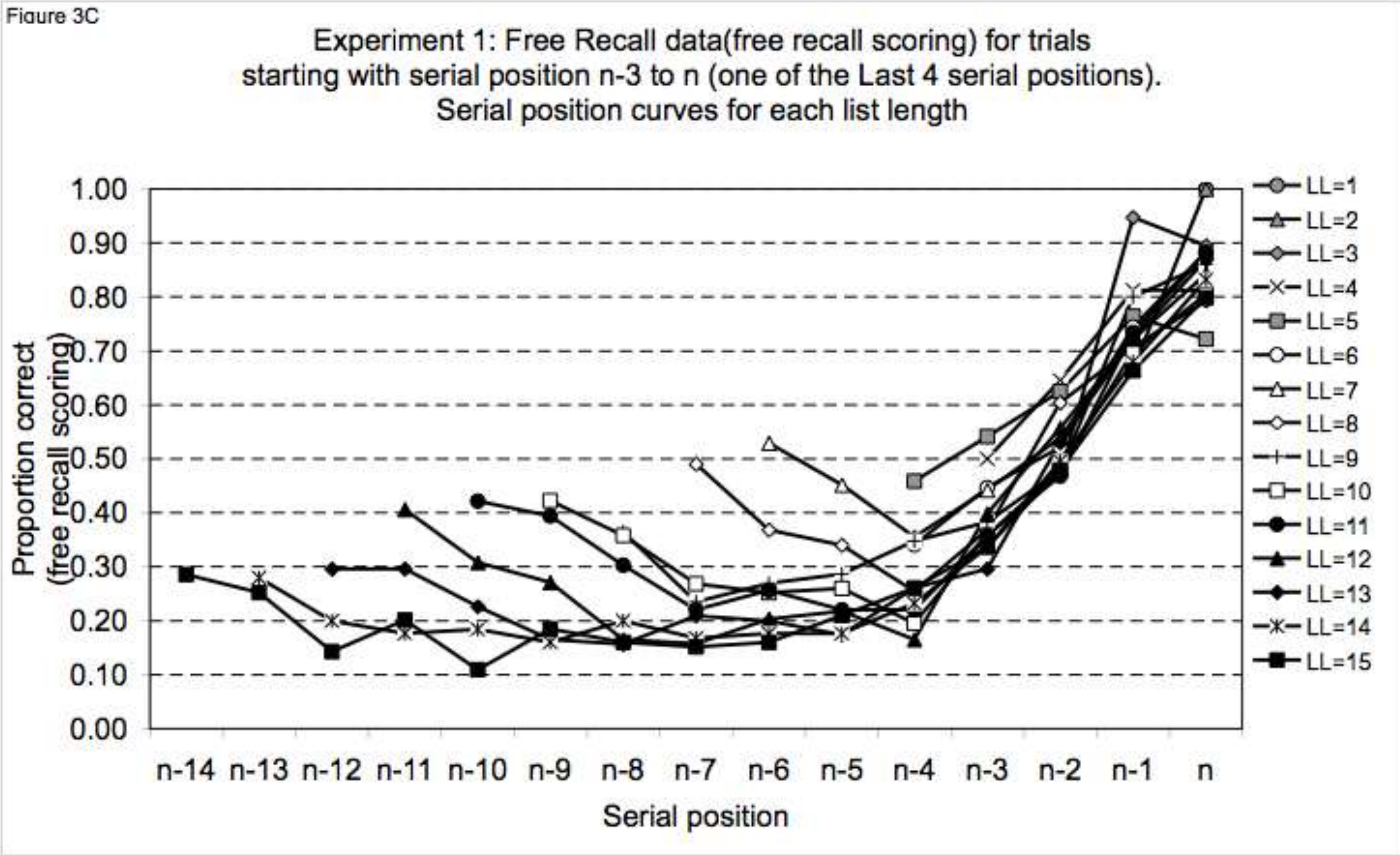


Figure 3C
Experiment 1: Free Recall data(free recall scoring) for trials starting with serial position n-3 to n (one of the Last 4 serial positions). Serial position curves for each list length
Proportion correct (free recall scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
n-14 n-13 n-12 n-11 n-10 n-9 n-8 n-7 n-6 n-5 n-4 n-3 n-2 n-1 n
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=9
LL=10
LL=11
LL=12
LL=13
LL=14
LL=15

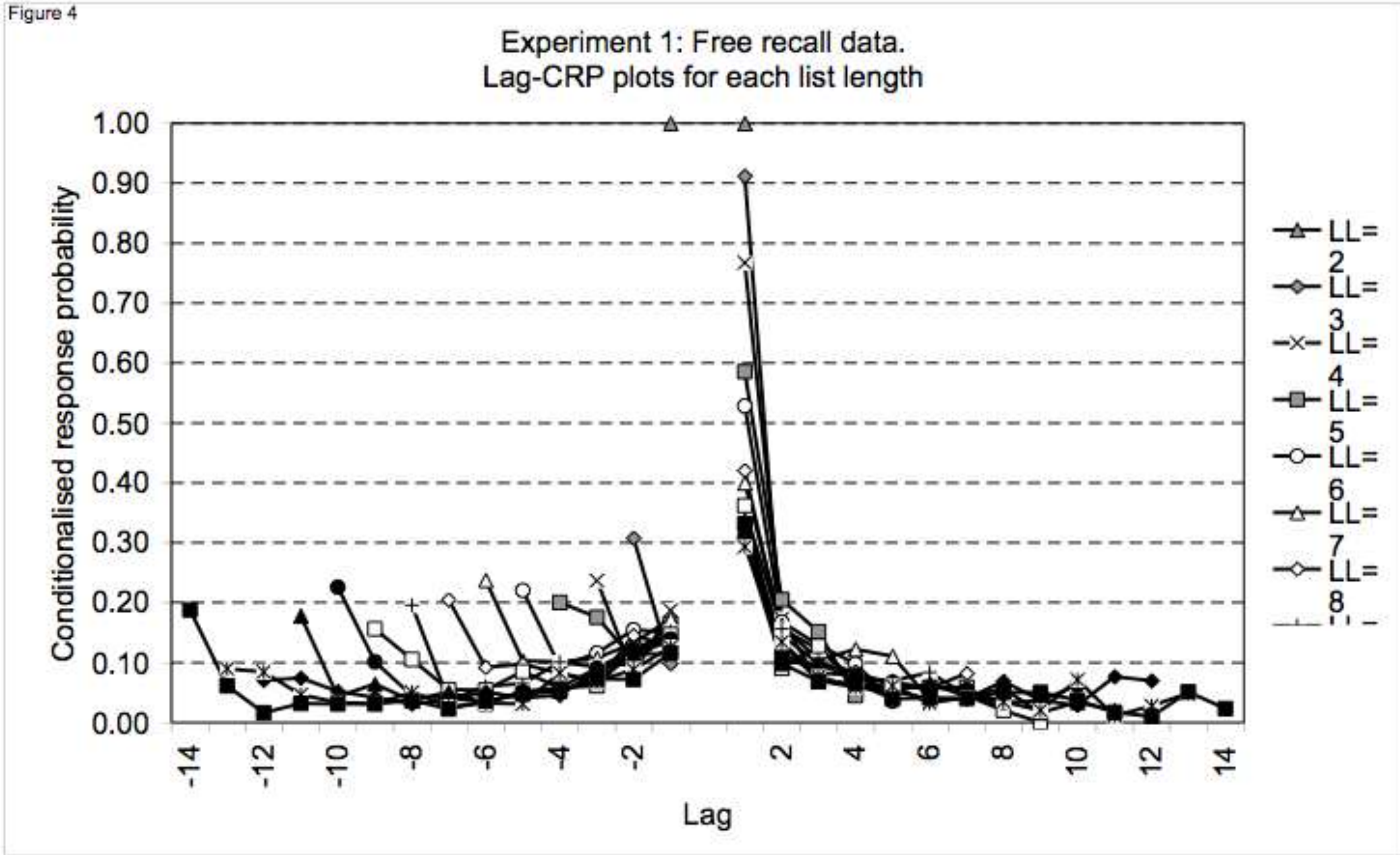


Figure 4
Experiment 1: Free recall data.
Lag-CRP plots for each list length
Conditionalised response probability
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
-14
-12
-10
-8
-6
-4
-2
2
4
6
8
10
12
14
Lag
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8

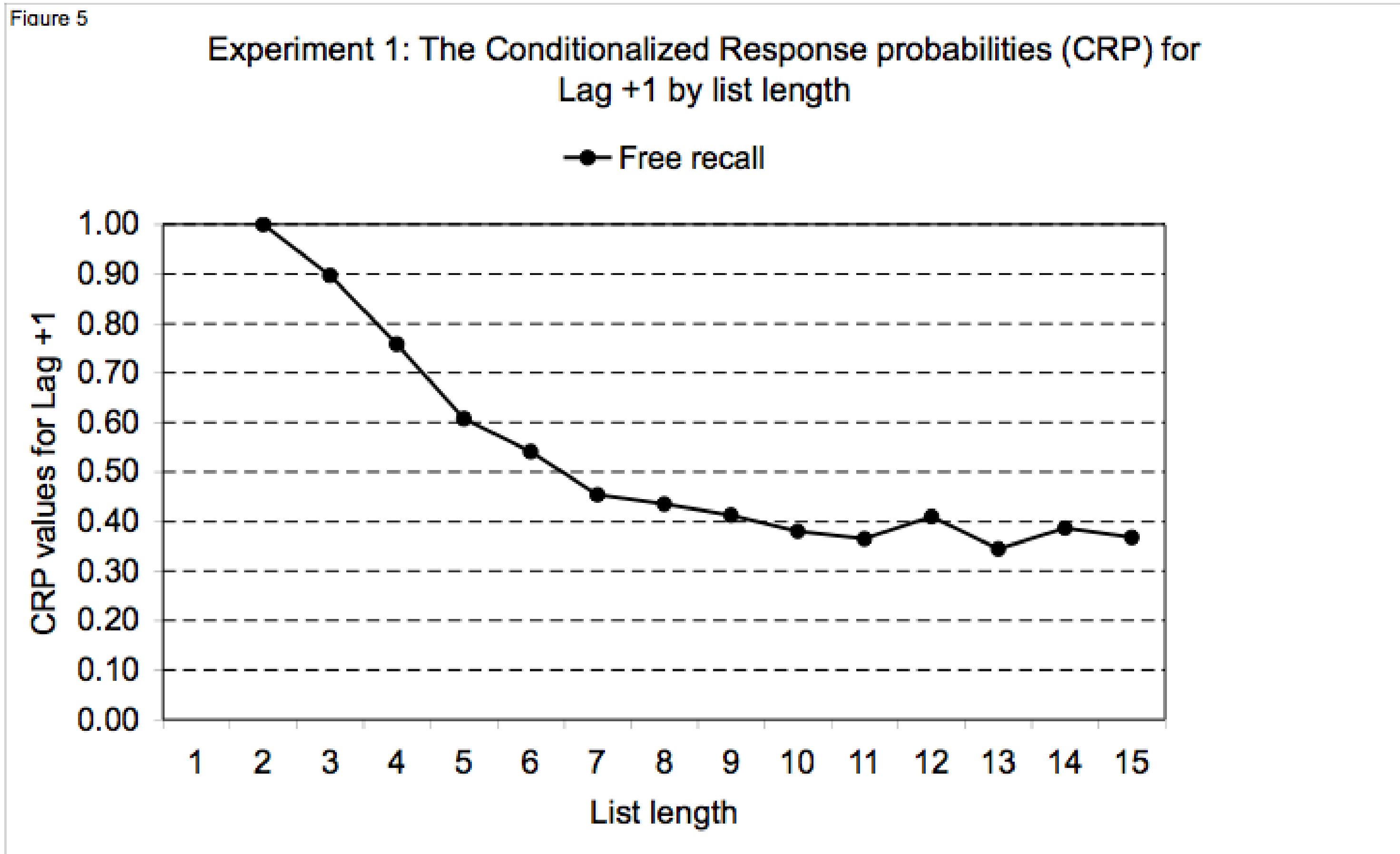


Figure 5
Experiment 1: The Conditionalized Response probabilities (CRP) for Lag +1 by list length
Free recall
CRP values for Lag +1
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
List length

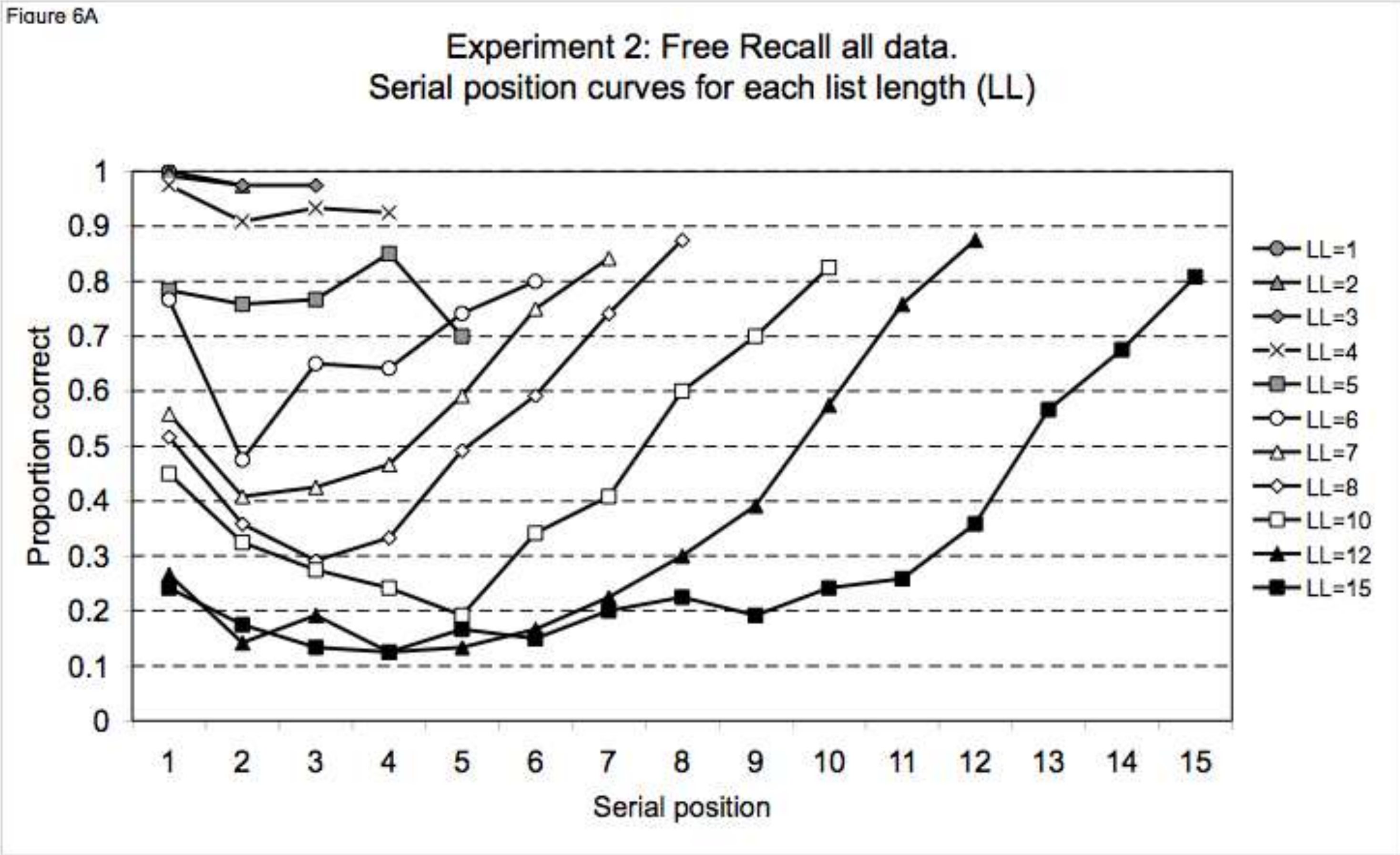


Figure 6A
Experiment 2: Free Recall all data.
Serial position curves for each list length (LL)
Proportion correct
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

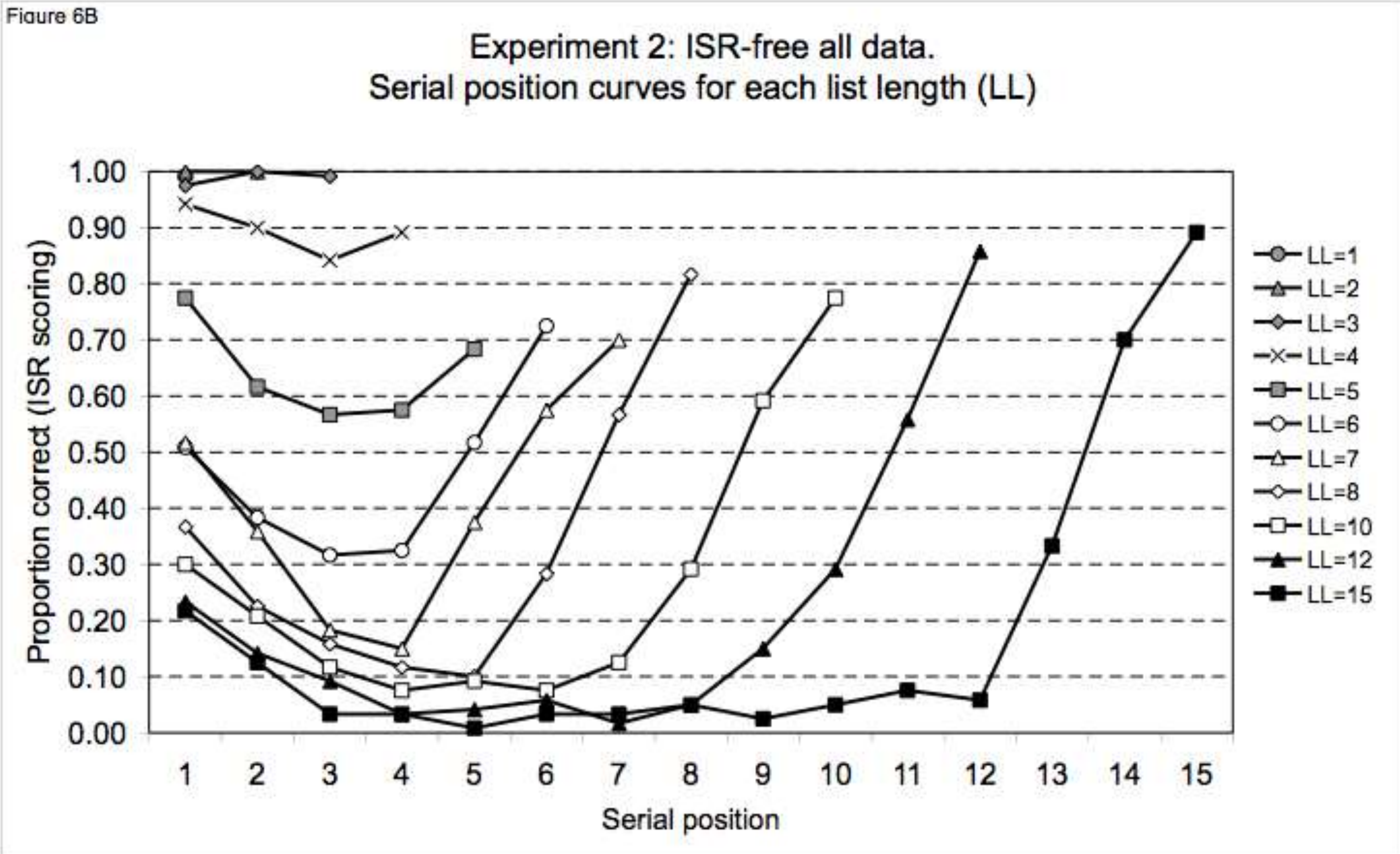


Figure 6B
Experiment 2: ISR-free all data.
Serial position curves for each list length (LL)
Proportion correct (ISR scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

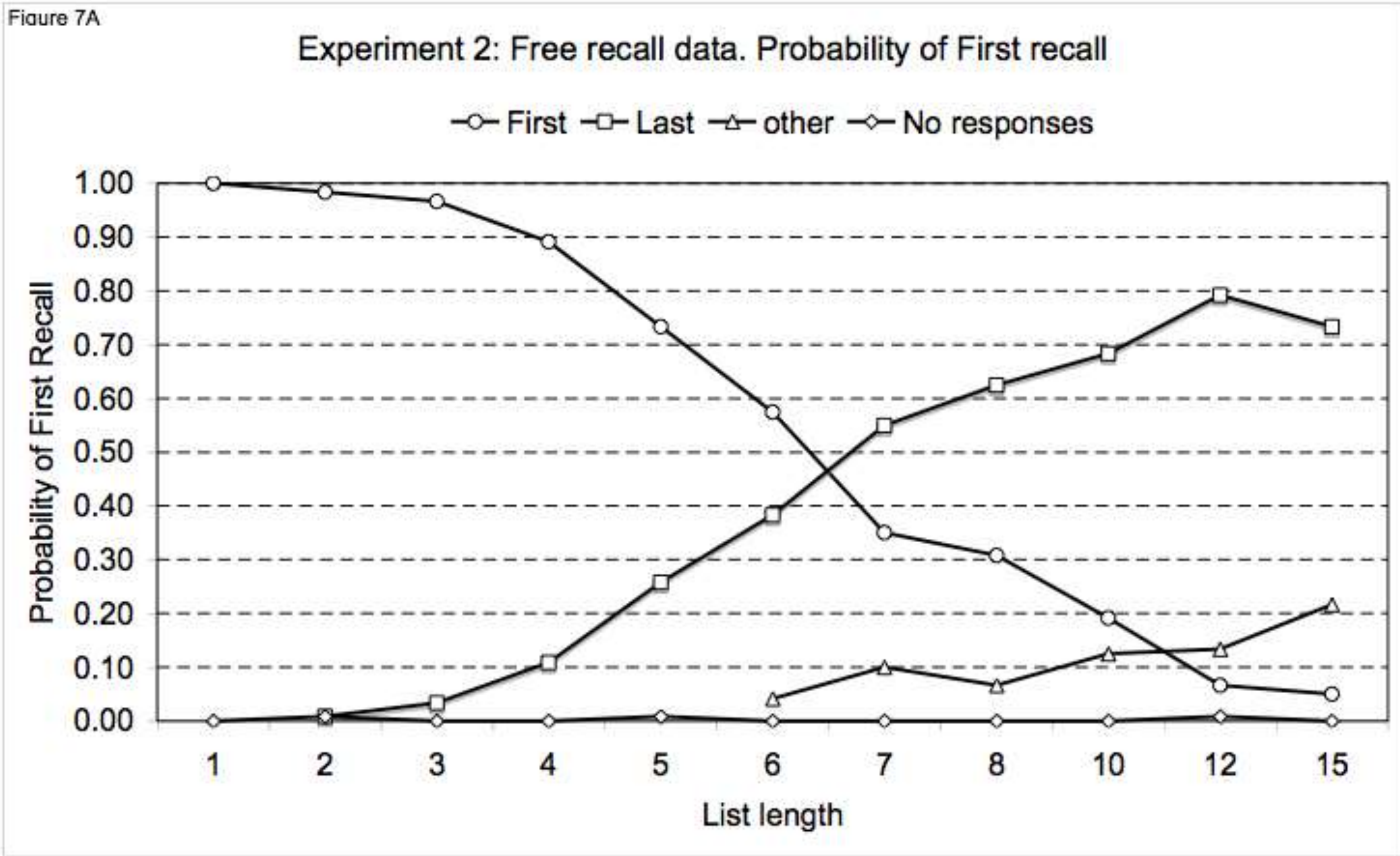


Figure 7A
Experiment 2: Free recall data. Probability of First recall
First
Last
other
No responses
Probability of First Recall
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
10
12
15
List length

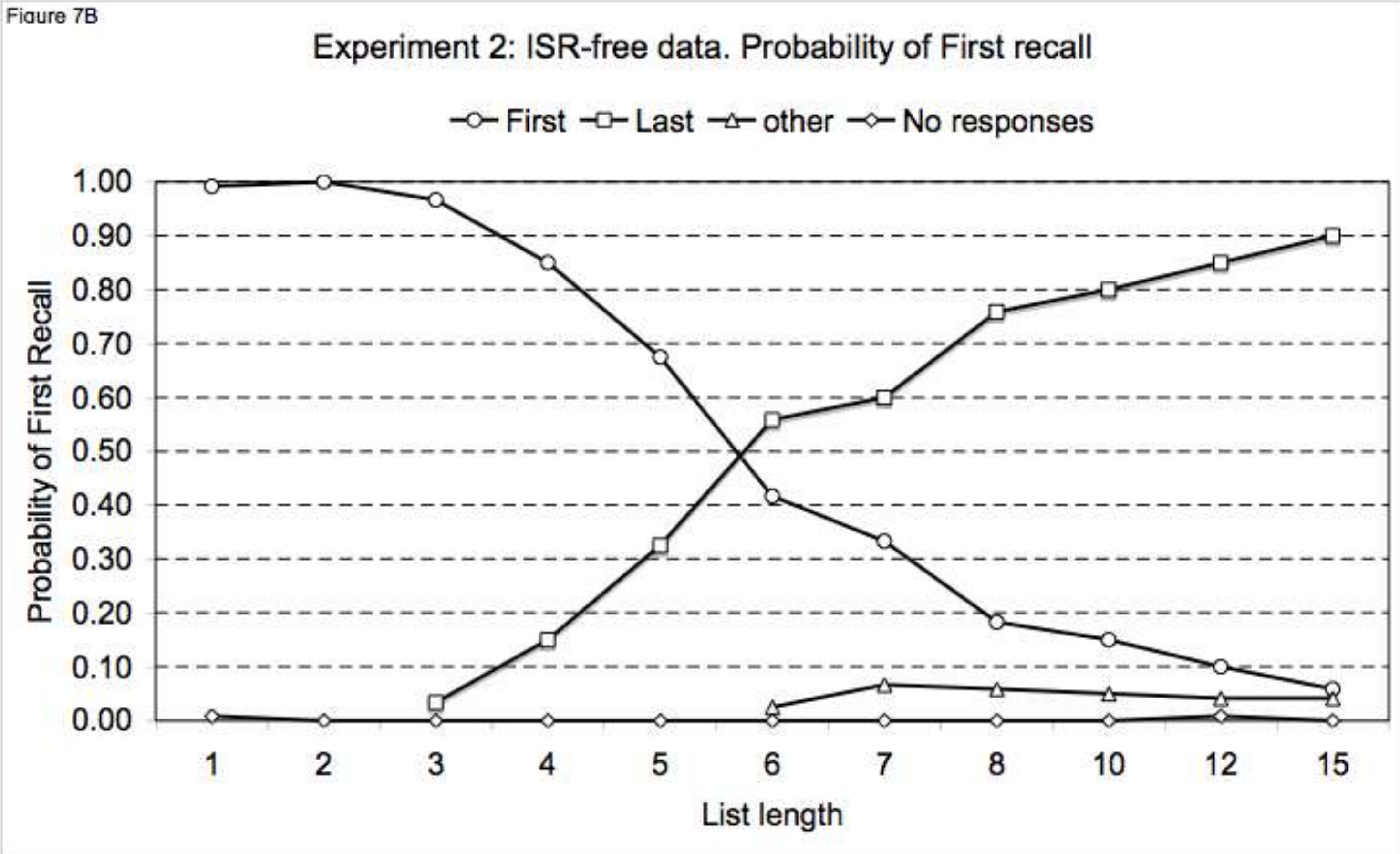
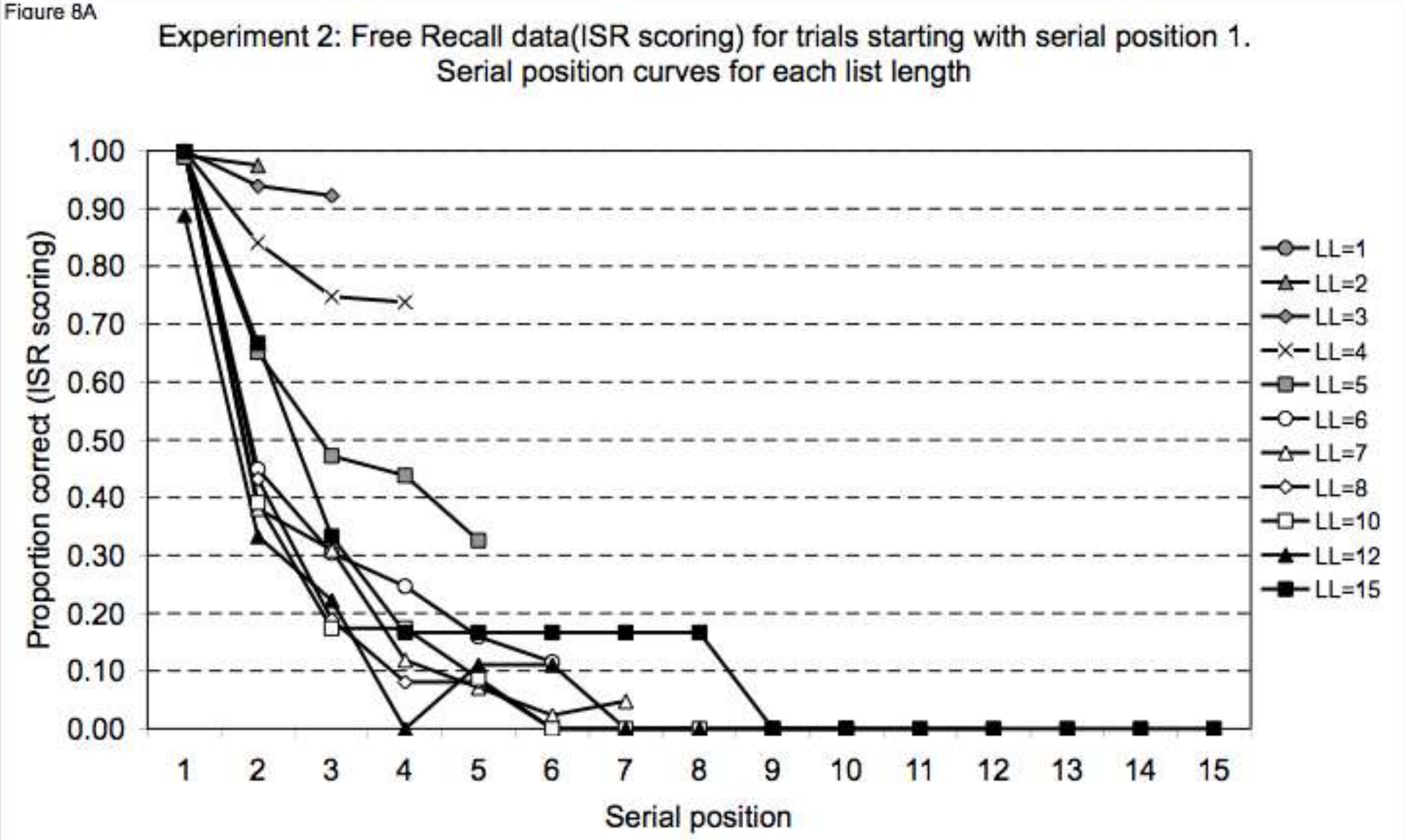


Figure 7B
Experiment 2: ISR-free data. Probability of First recall
First
Last
other
No responses
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Probability of First Recall
1
2
3
4
5
6
7
8
10
12
15
List length

Figure 8A


Experiment 2: Free Recall data(ISR scoring) for trials starting with serial position 1.
Serial position curves for each list length
Proportion correct (ISR scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

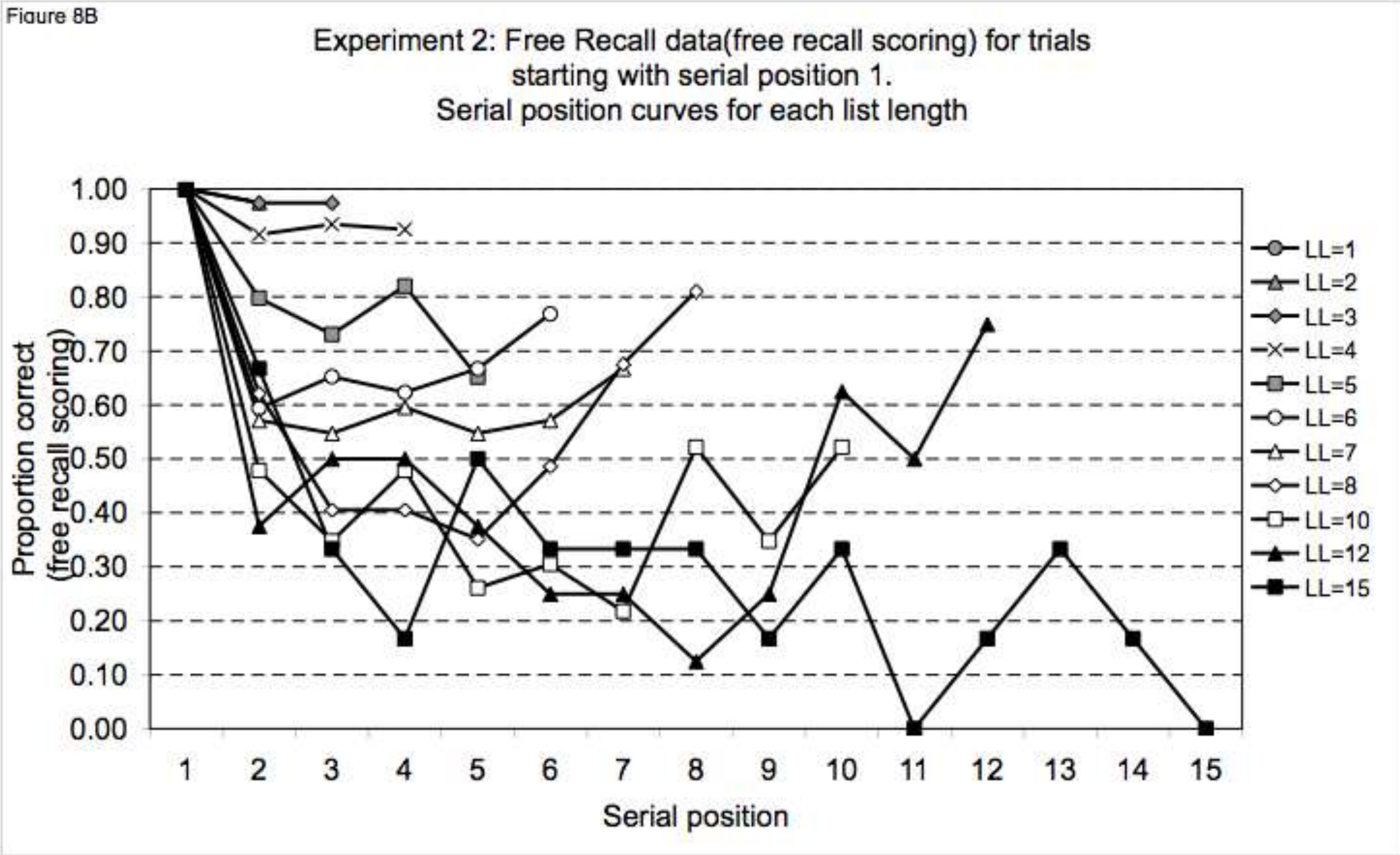
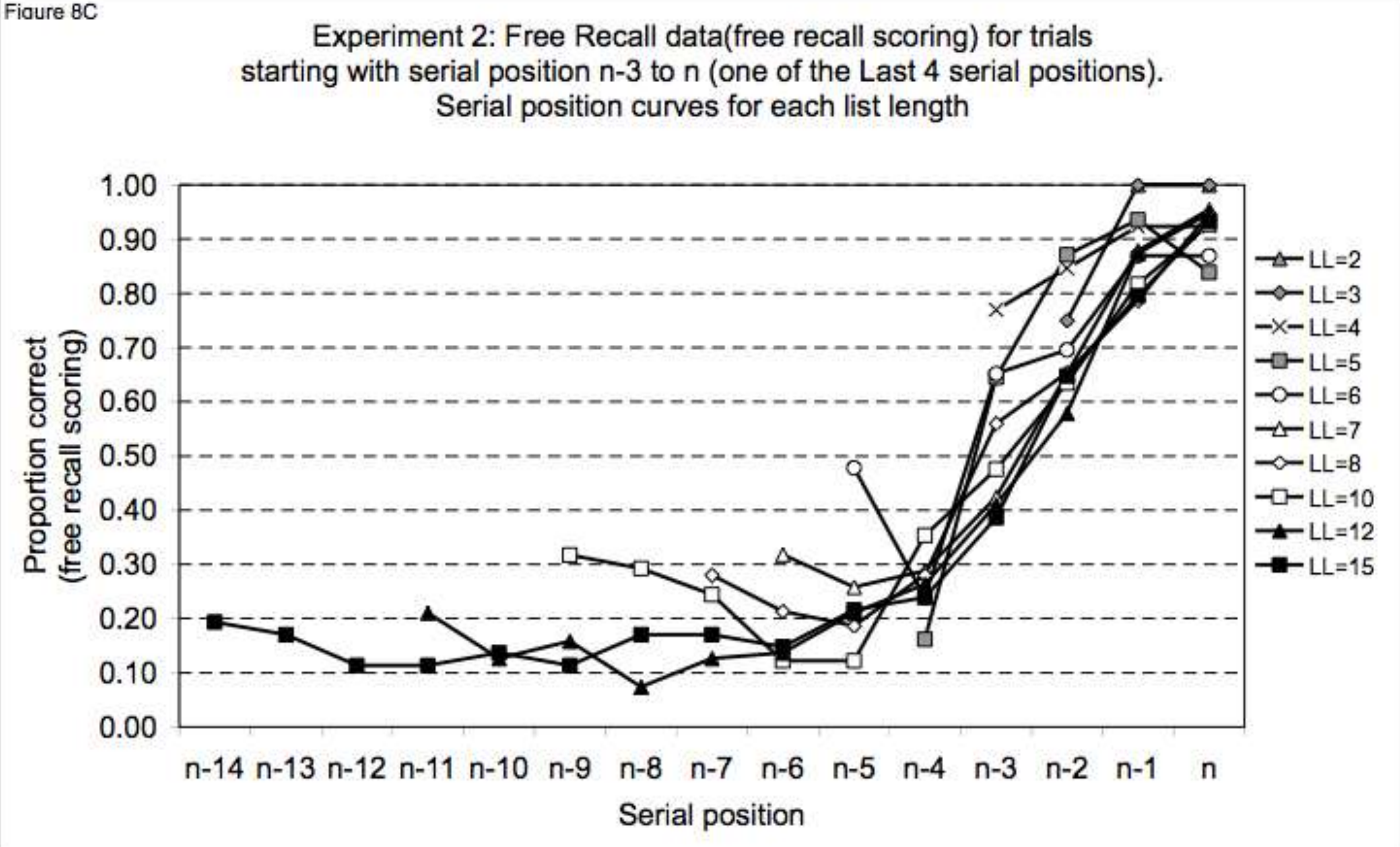


Figure 8B
Experiment 2: Free Recall data(free recall scoring) for trials starting with serial position 1.
Serial position curves for each list length
Proportion correct (free recall scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

Figure 8C


Experiment 2: Free Recall data(free recall scoring) for trials starting with serial position n-3 to n (one of the Last 4 serial positions). Serial position curves for each list length
Proportion correct (free recall scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
n-14 n-13 n-12 n-11 n-10 n-9 n-8 n-7 n-6 n-5 n-4 n-3 n-2 n-1 n
Serial position
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

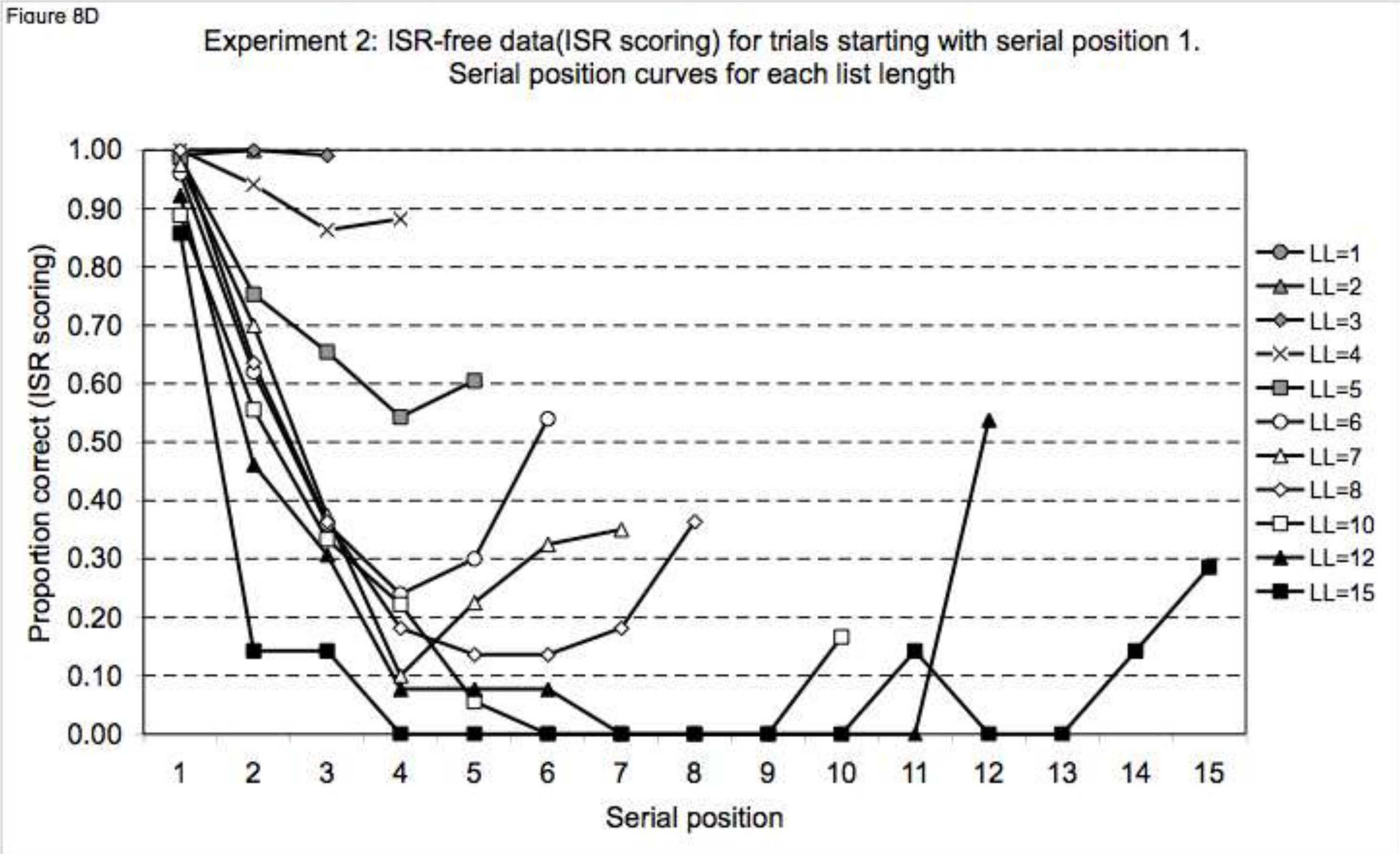


Figure 8D
Experiment 2: ISR-free data(ISR scoring) for trials starting with serial position 1.
Serial position curves for each list length
Proportion correct (ISR scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

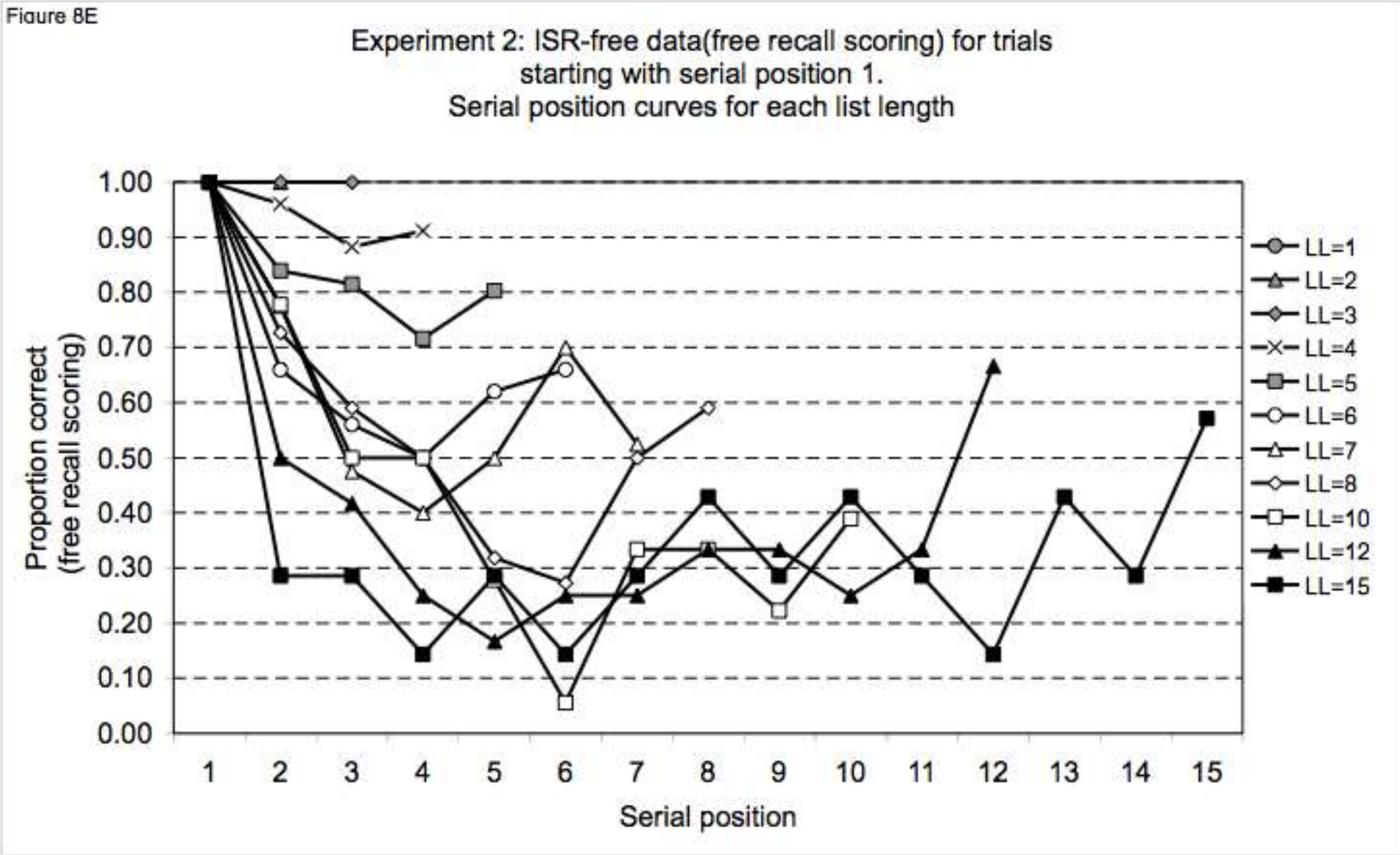
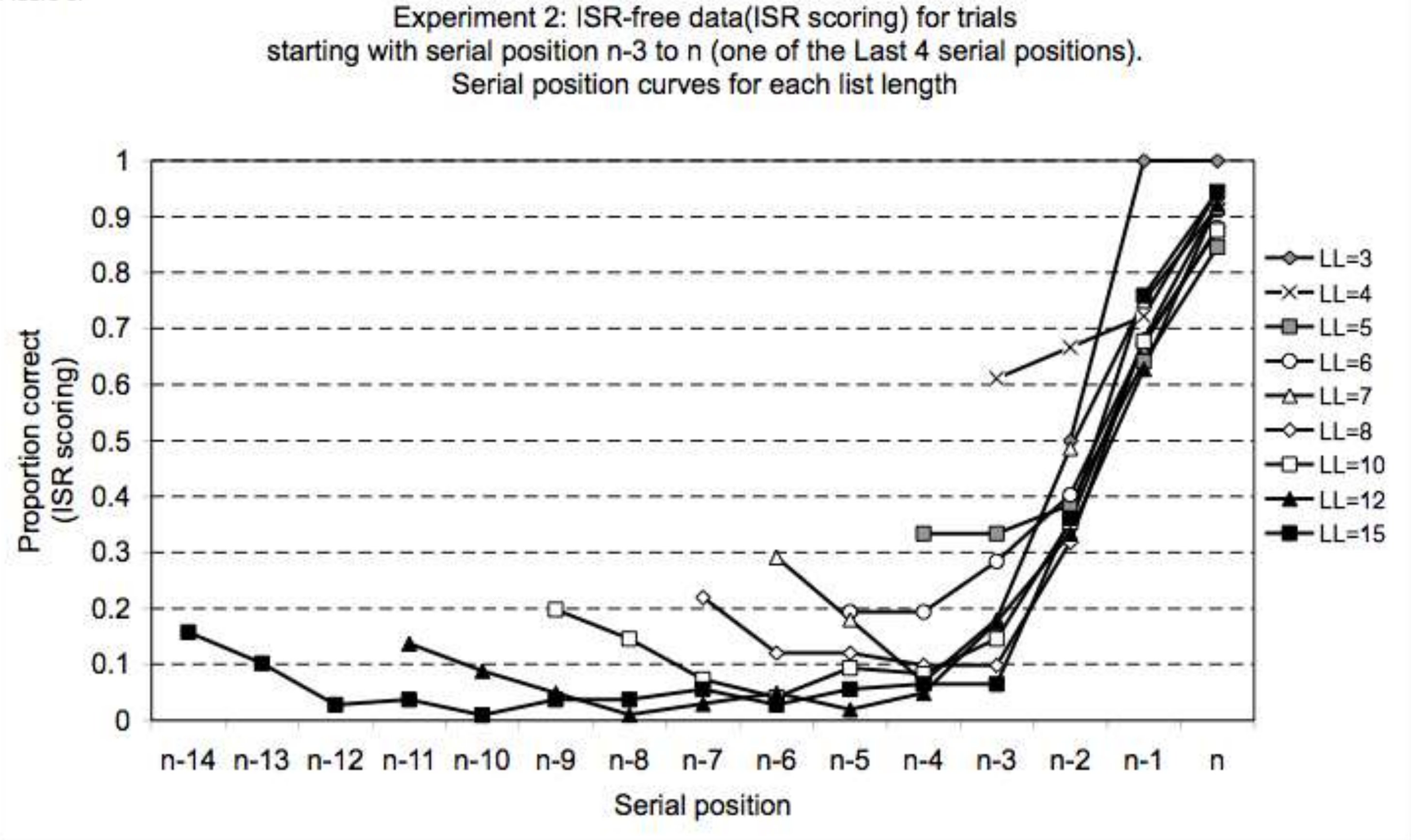


Figure 8E
Experiment 2: ISR-free data(free recall scoring) for trials
starting with serial position 1.
Serial position curves for each list length
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Proportion correct
(free recall scoring)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

Figure 8F


Experiment 2: ISR-free data(ISR scoring) for trials
starting with serial position n-3 to n (one of the Last 4 serial positions).
Serial position curves for each list length
Proportion correct
(ISR scoring)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
n-14 n-13 n-12 n-11 n-10 n-9 n-8 n-7 n-6 n-5 n-4 n-3 n-2 n-1 n
Serial position
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

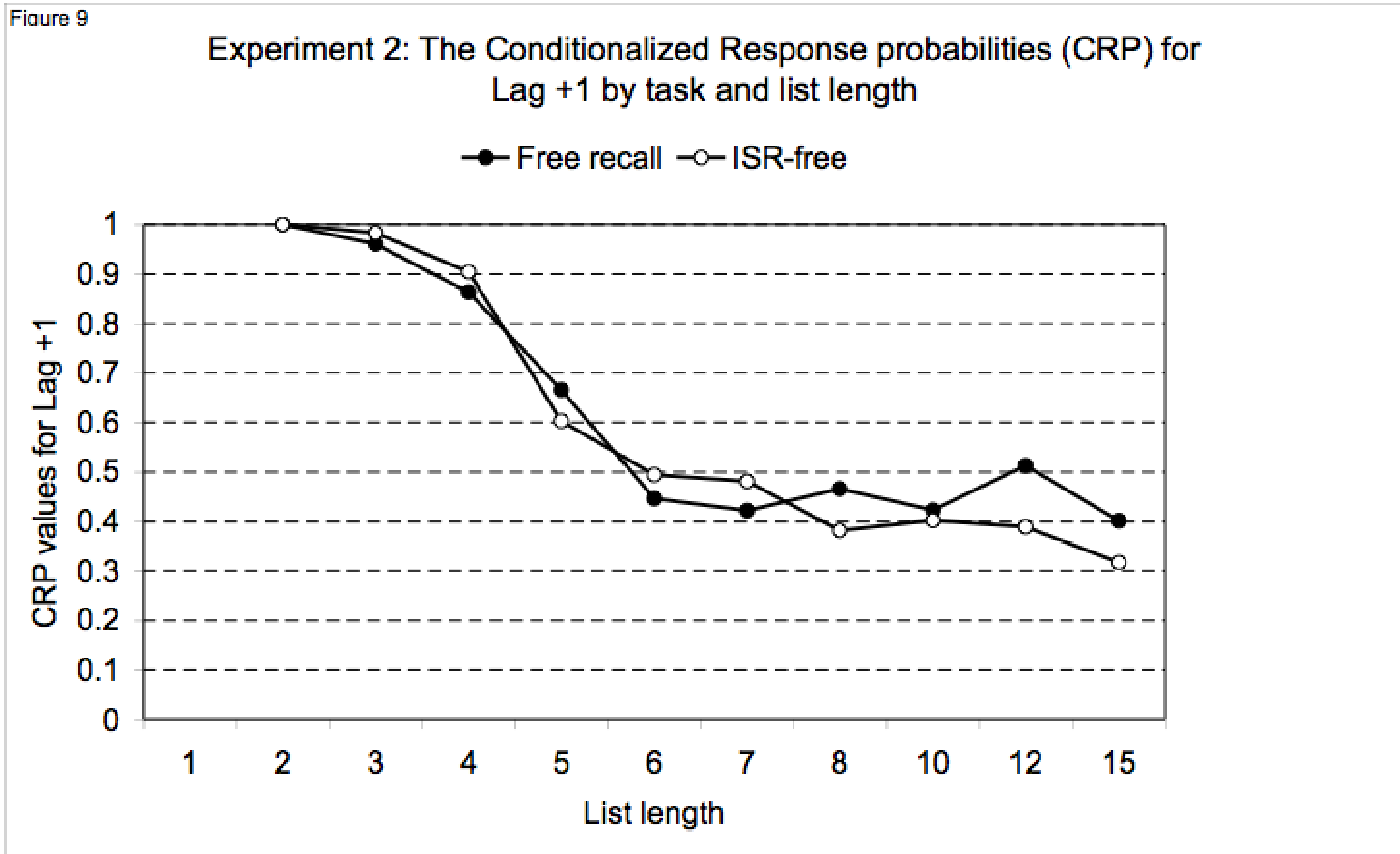


Figure 9
Experiment 2: The Conditionalized Response probabilities (CRP) for Lag +1 by task and list length
Free recall
ISR-free
CRP values for Lag +1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7
8
10
12
15
List length

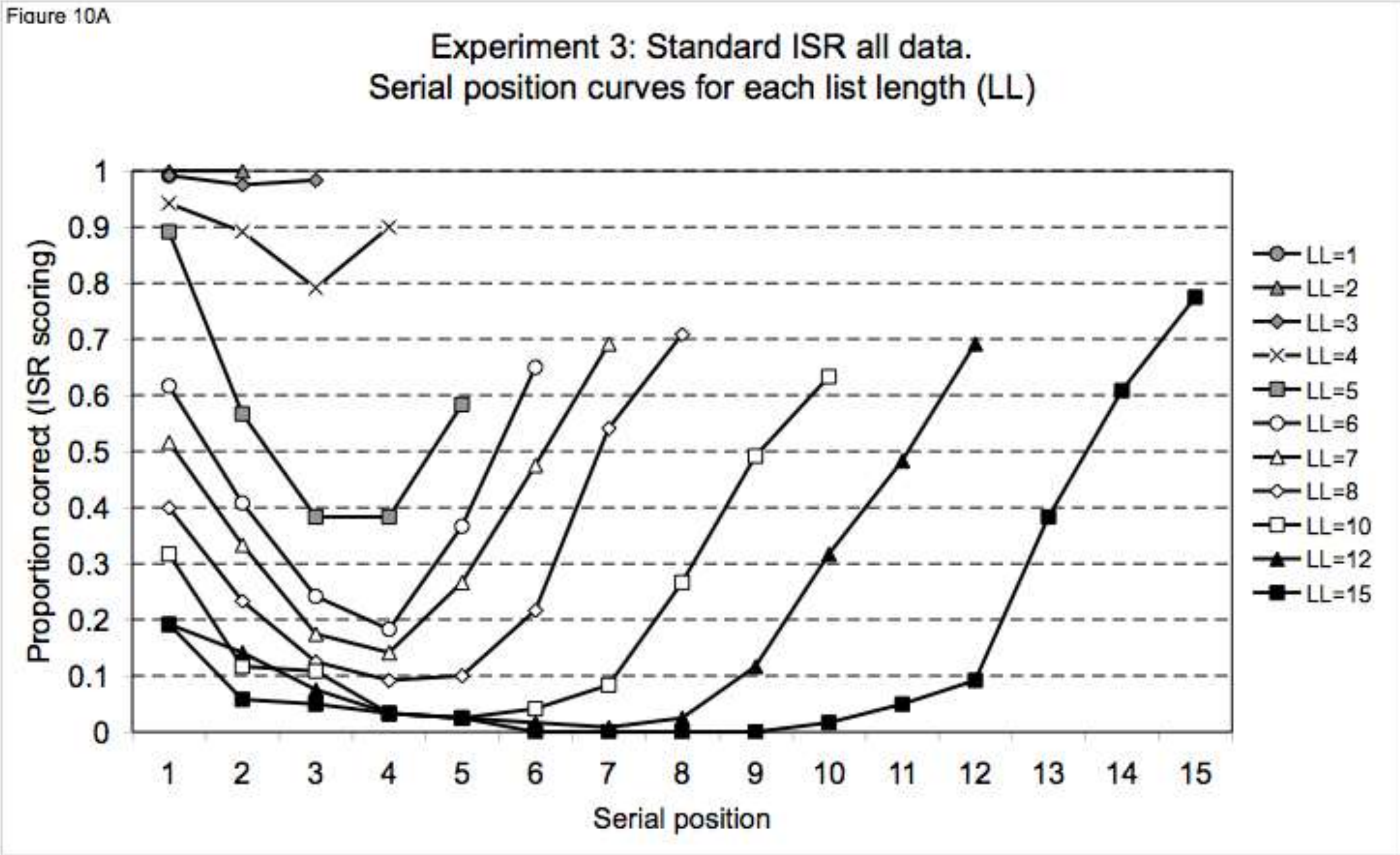


Figure 10A
Experiment 3: Standard ISR all data.
Serial position curves for each list length (LL)
Proportion correct (ISR scoring)
Serial position
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

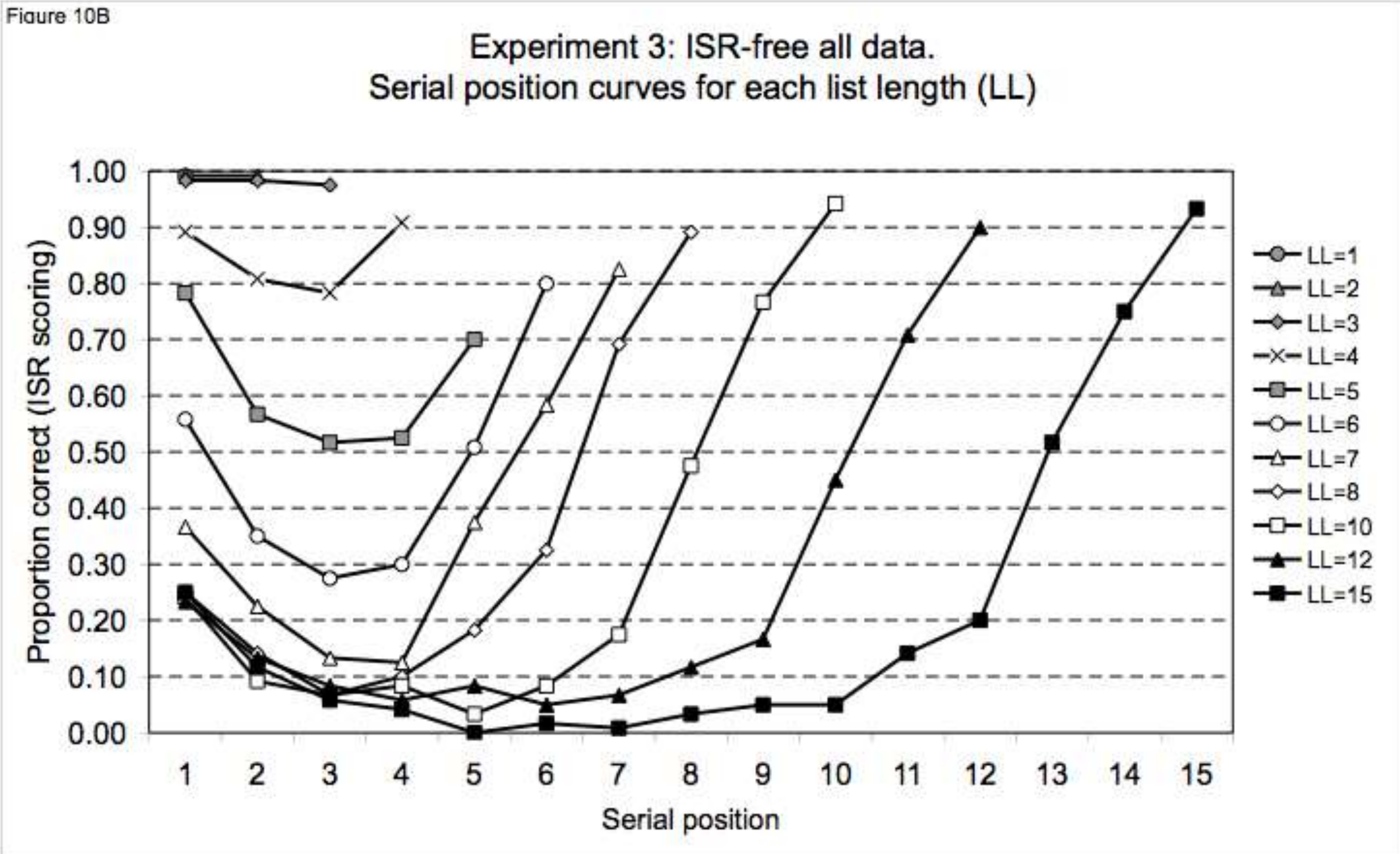
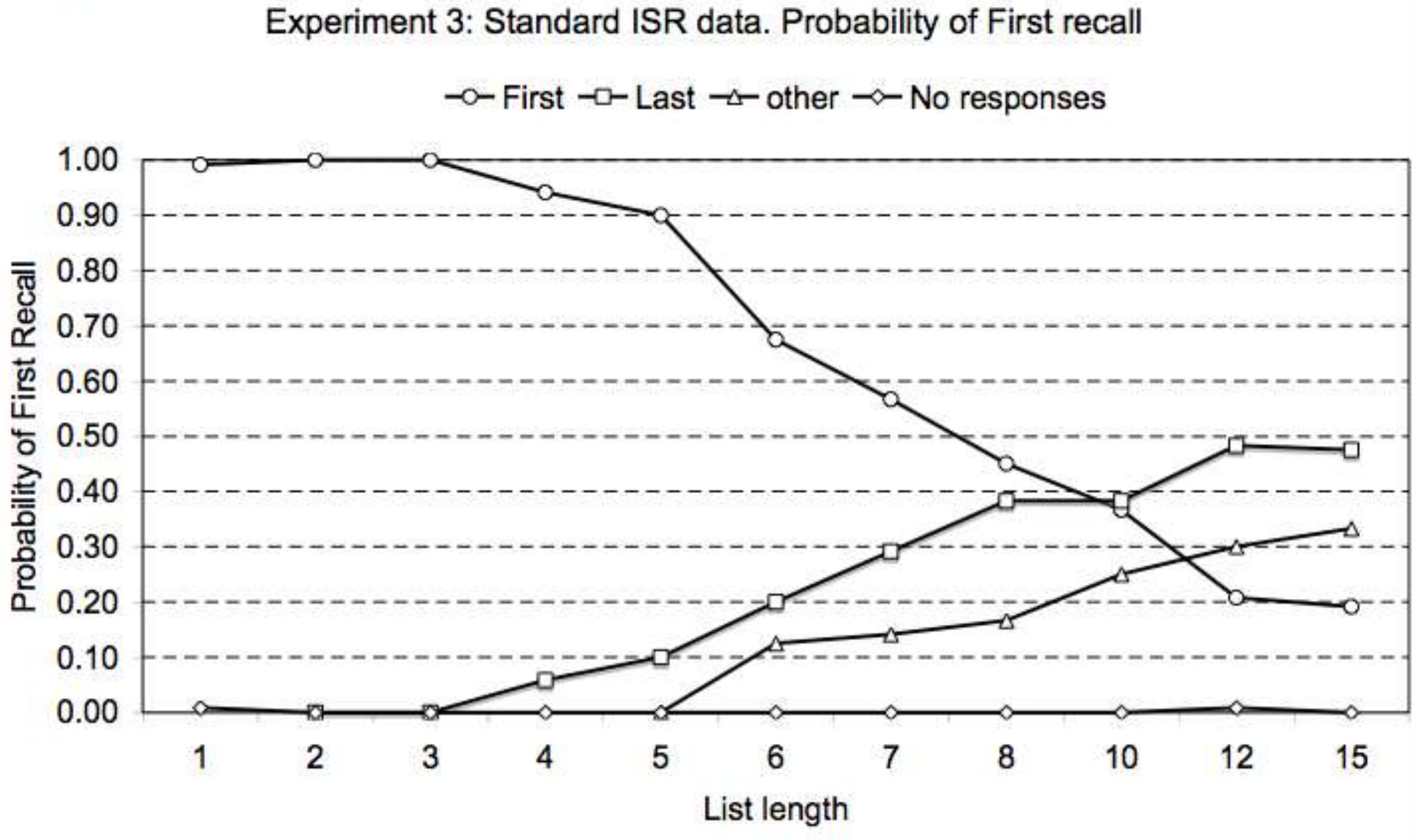


Figure 10B
Experiment 3: ISR-free all data.
Serial position curves for each list length (LL)
Proportion correct (ISR scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

Figure 11A



Experiment 3: Standard ISR data. Probability of First recall
First
Last
other
No responses
Probability of First Recall
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
10
12
15
List length

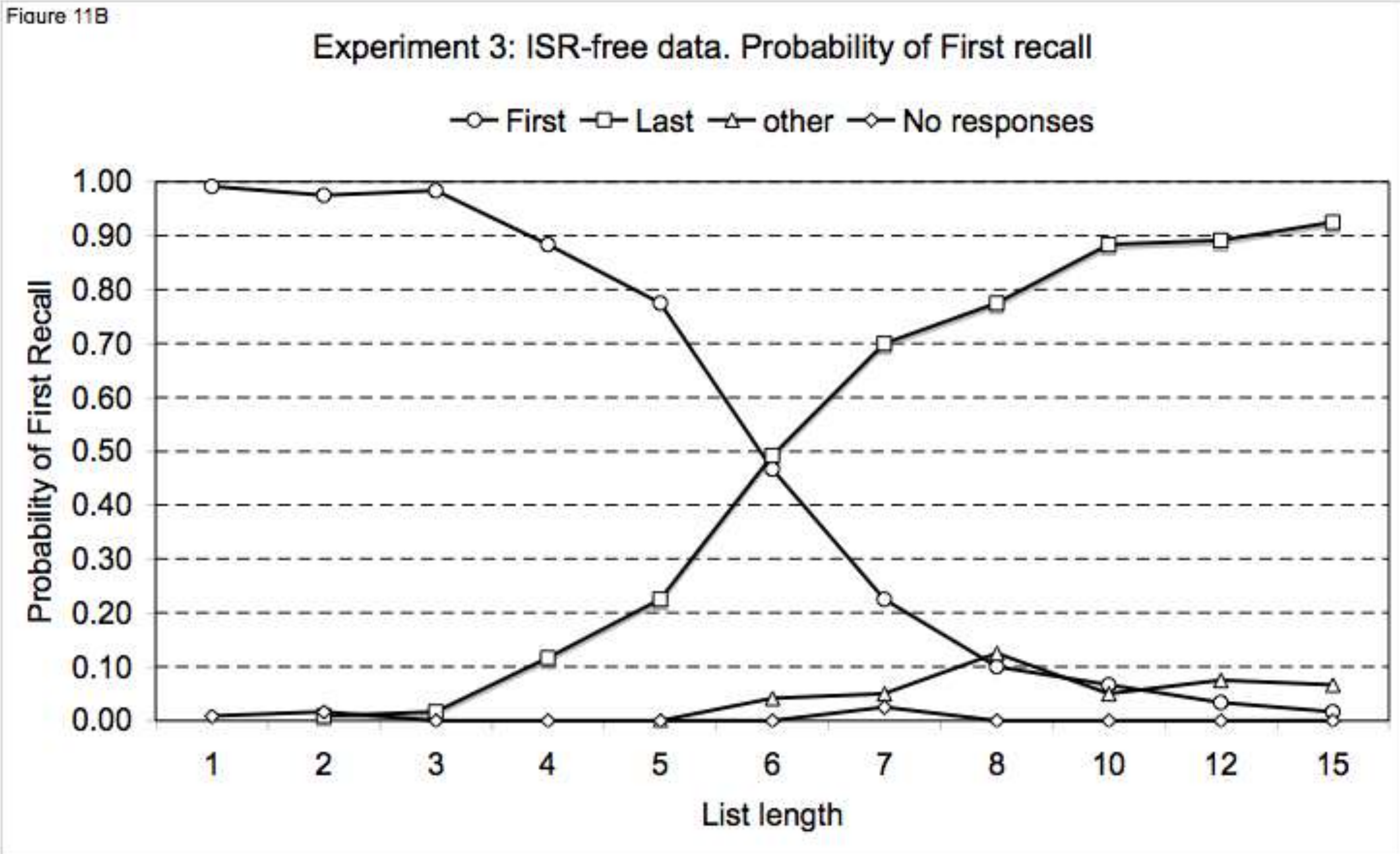


Figure 11B
Experiment 3: ISR-free data. Probability of First recall
First
Last
other
No responses
Probability of First Recall
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
10
12
15
List length

Figure 12A

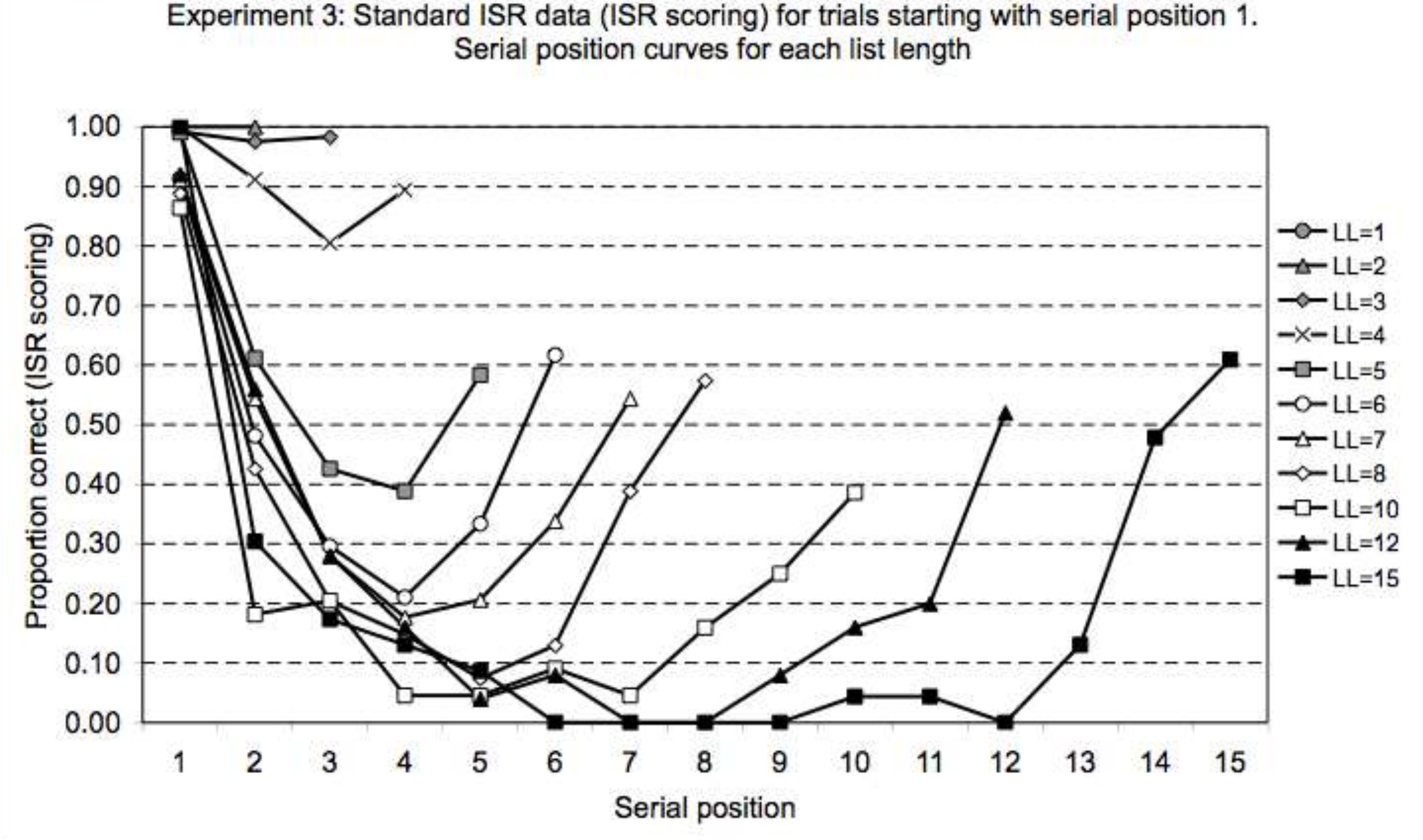

Experiment 3: Standard ISR data (ISR scoring) for trials starting with serial position 1.
Serial position curves for each list length
Proportion correct (ISR scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

Figure 12B

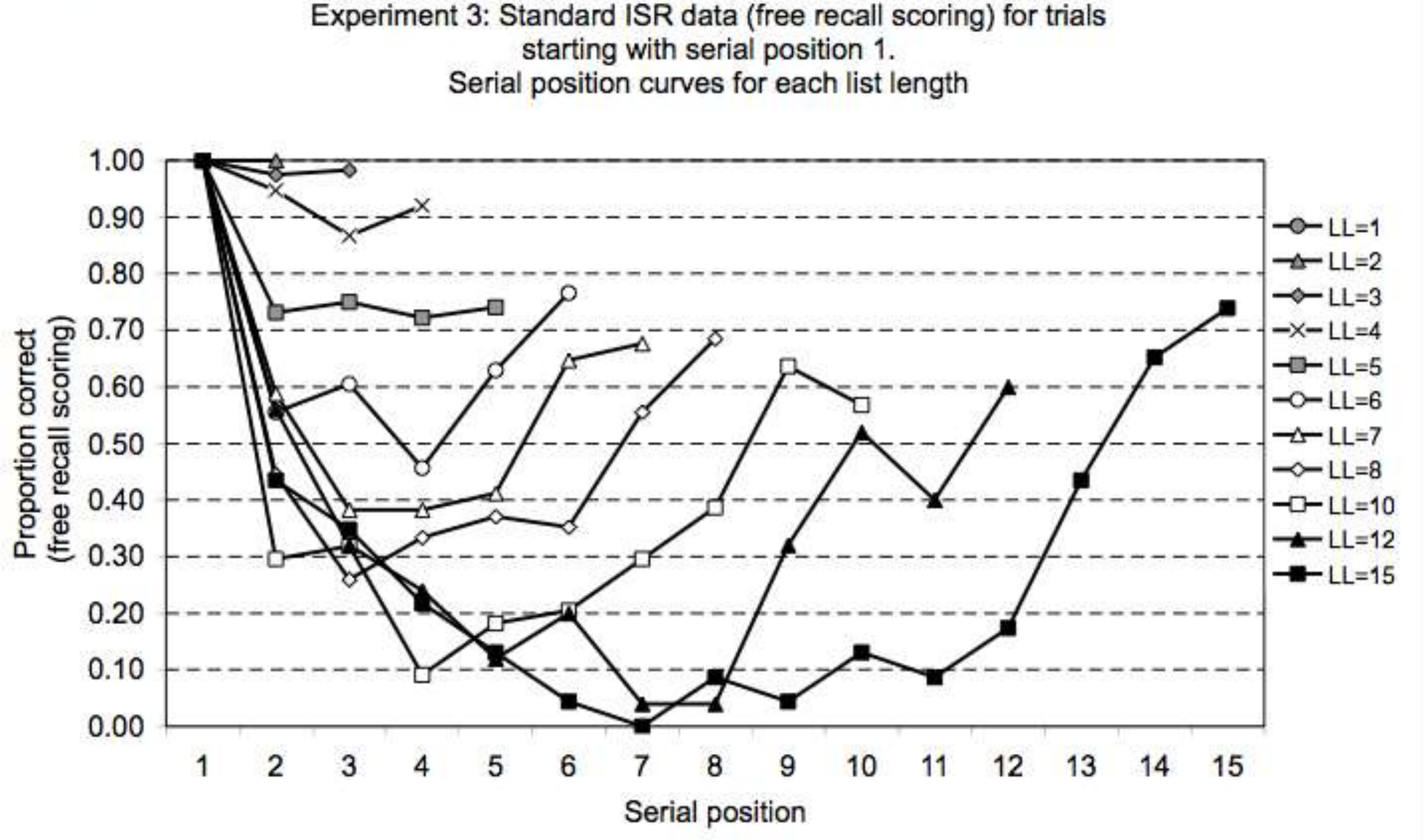
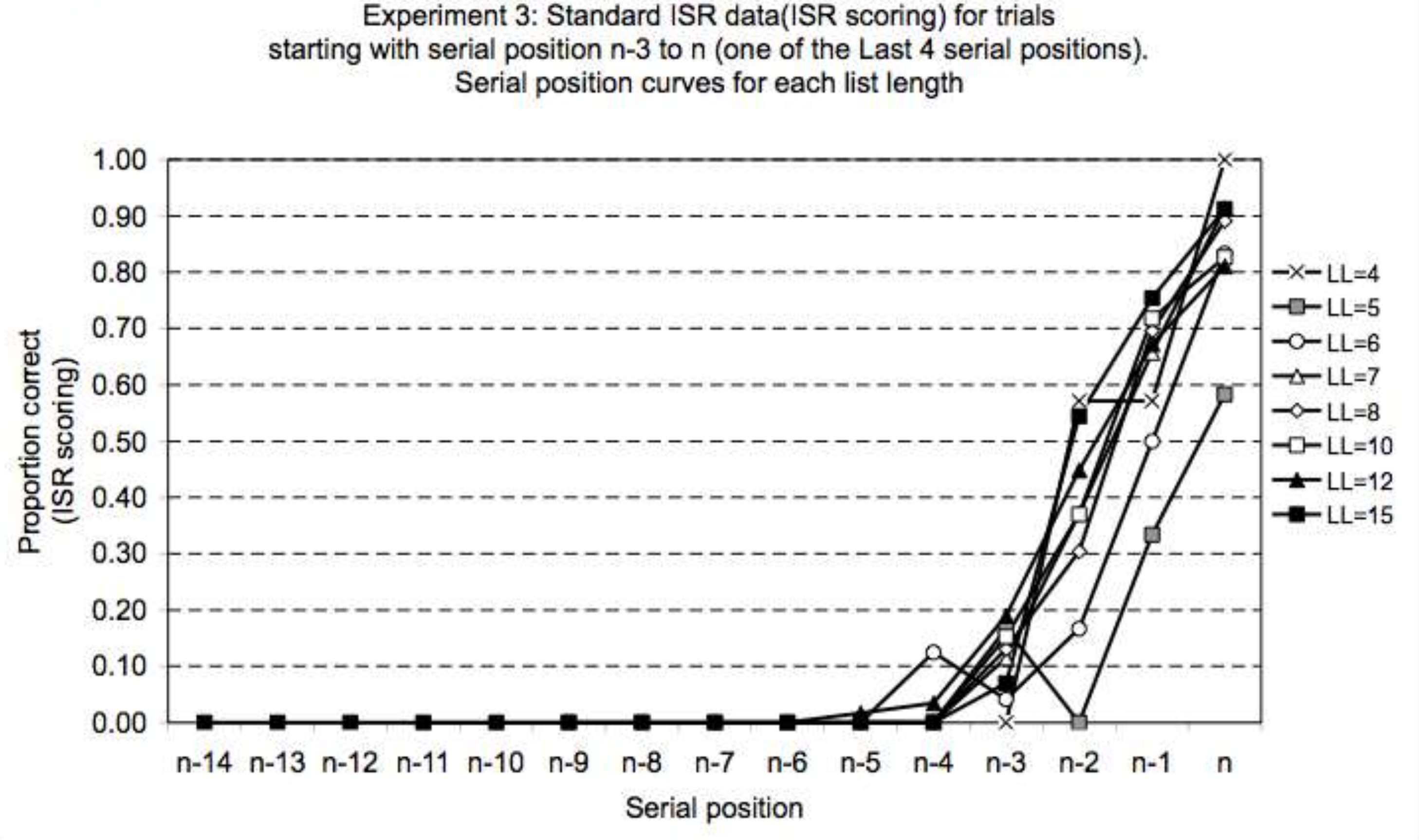
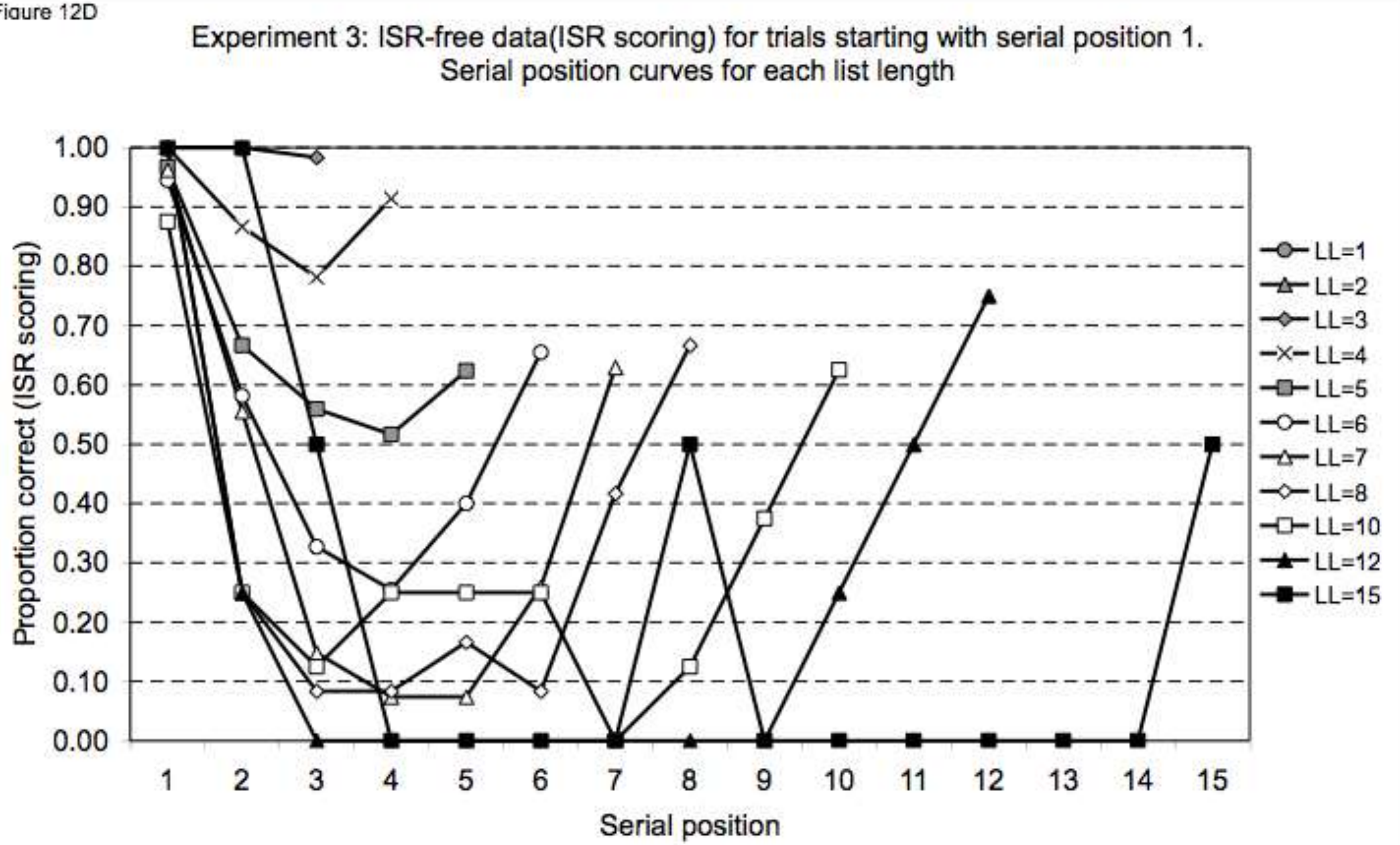

Experiment 3: Standard ISR data (free recall scoring) for trials starting with serial position 1.
Serial position curves for each list length
Proportion correct (free recall scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

Figure 12C



Experiment 3: Standard ISR data(ISR scoring) for trials starting with serial position n-3 to n (one of the Last 4 serial positions). Serial position curves for each list length
Proportion correct (ISR scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
n-14
n-13
n-12
n-11
n-10
n-9
n-8
n-7
n-6
n-5
n-4
n-3
n-2
n-1
n
Serial position
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

Figure 12D


Experiment 3: ISR-free data(ISR scoring) for trials starting with serial position 1.
Serial position curves for each list length
Proportion correct (ISR scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

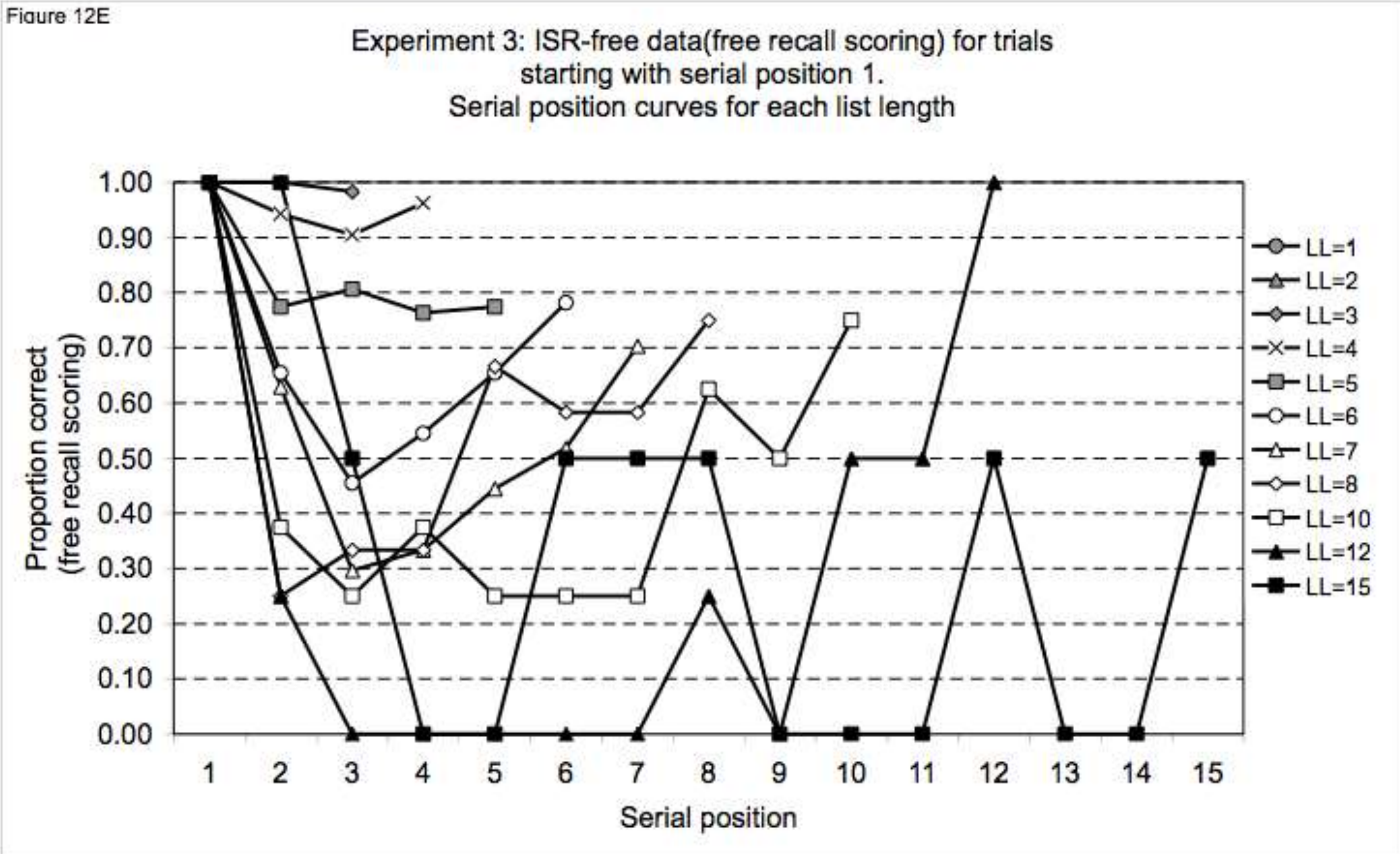


Figure 12E
Experiment 3: ISR-free data(free recall scoring) for trials starting with serial position 1.
Serial position curves for each list length
Proportion correct (free recall scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

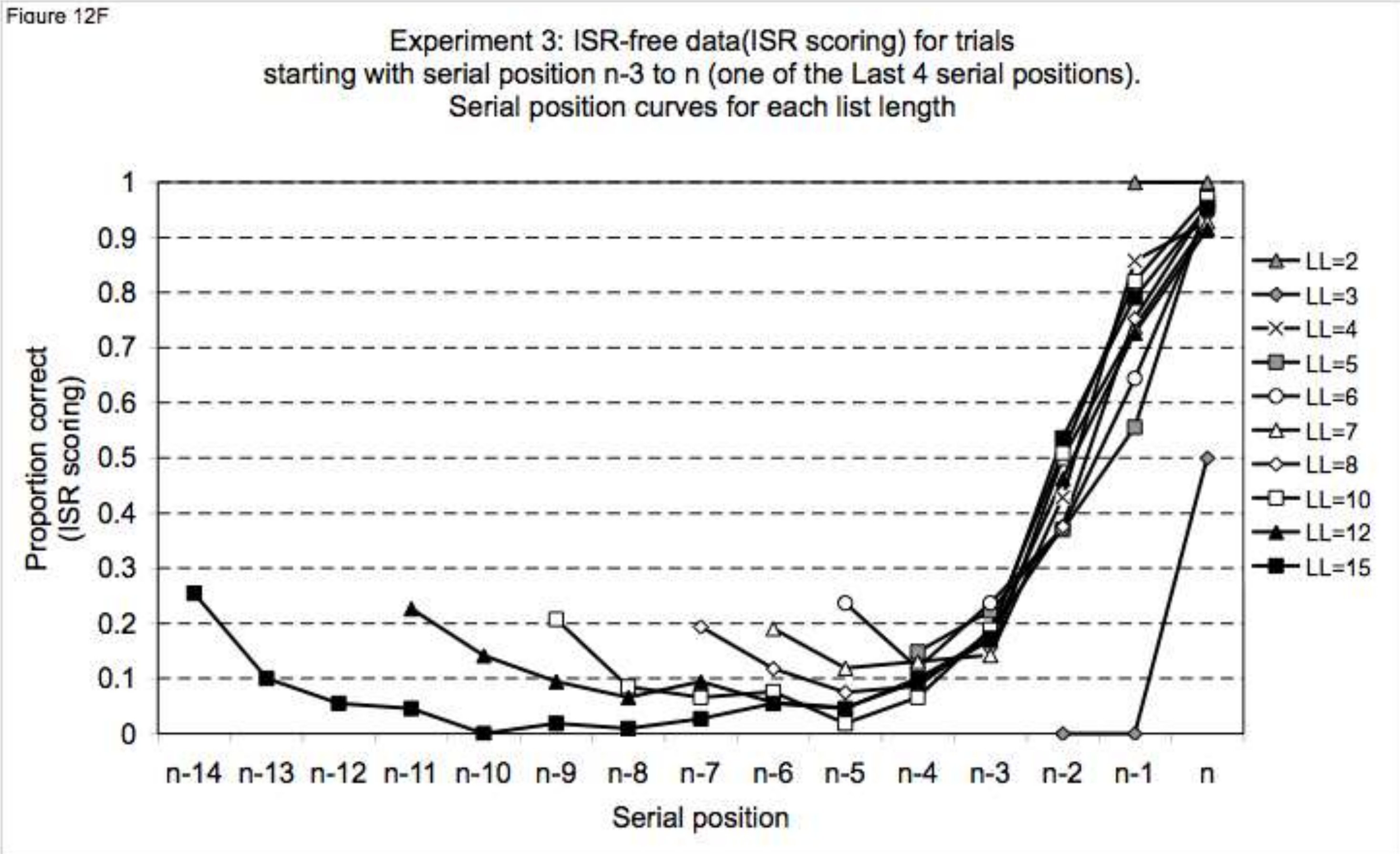


Figure 12F
Experiment 3: ISR-free data(ISR scoring) for trials
starting with serial position n-3 to n (one of the Last 4 serial positions).
Serial position curves for each list length
Proportion correct
(ISR scoring)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
n-14 n-13 n-12 n-11 n-10 n-9 n-8 n-7 n-6 n-5 n-4 n-3 n-2 n-1 n
Serial position
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

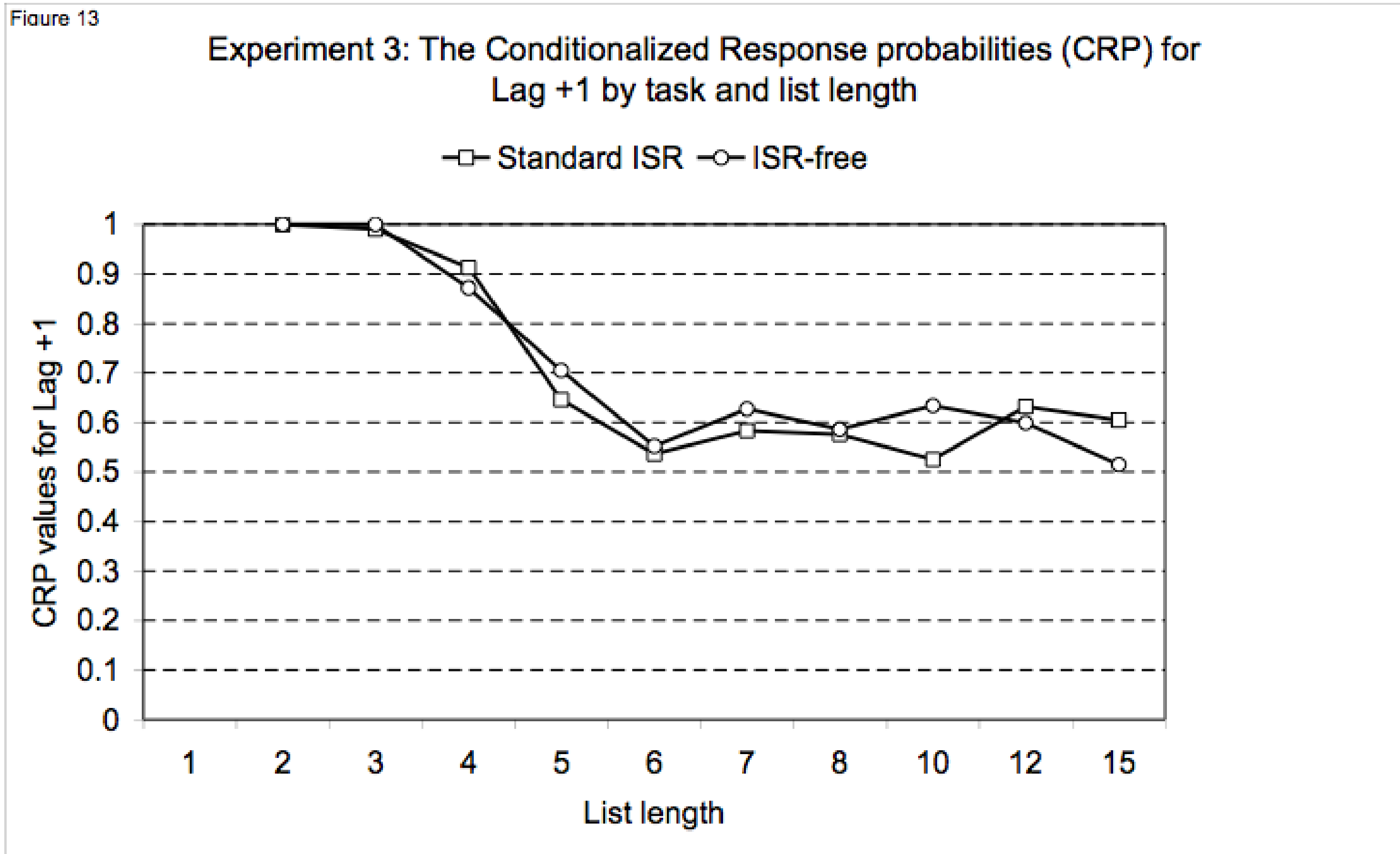


Figure 13
Experiment 3: The Conditionalized Response probabilities (CRP) for Lag +1 by task and list length
Standard ISR
ISR-free
CRP values for Lag +1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7
8
10
12
15
List length

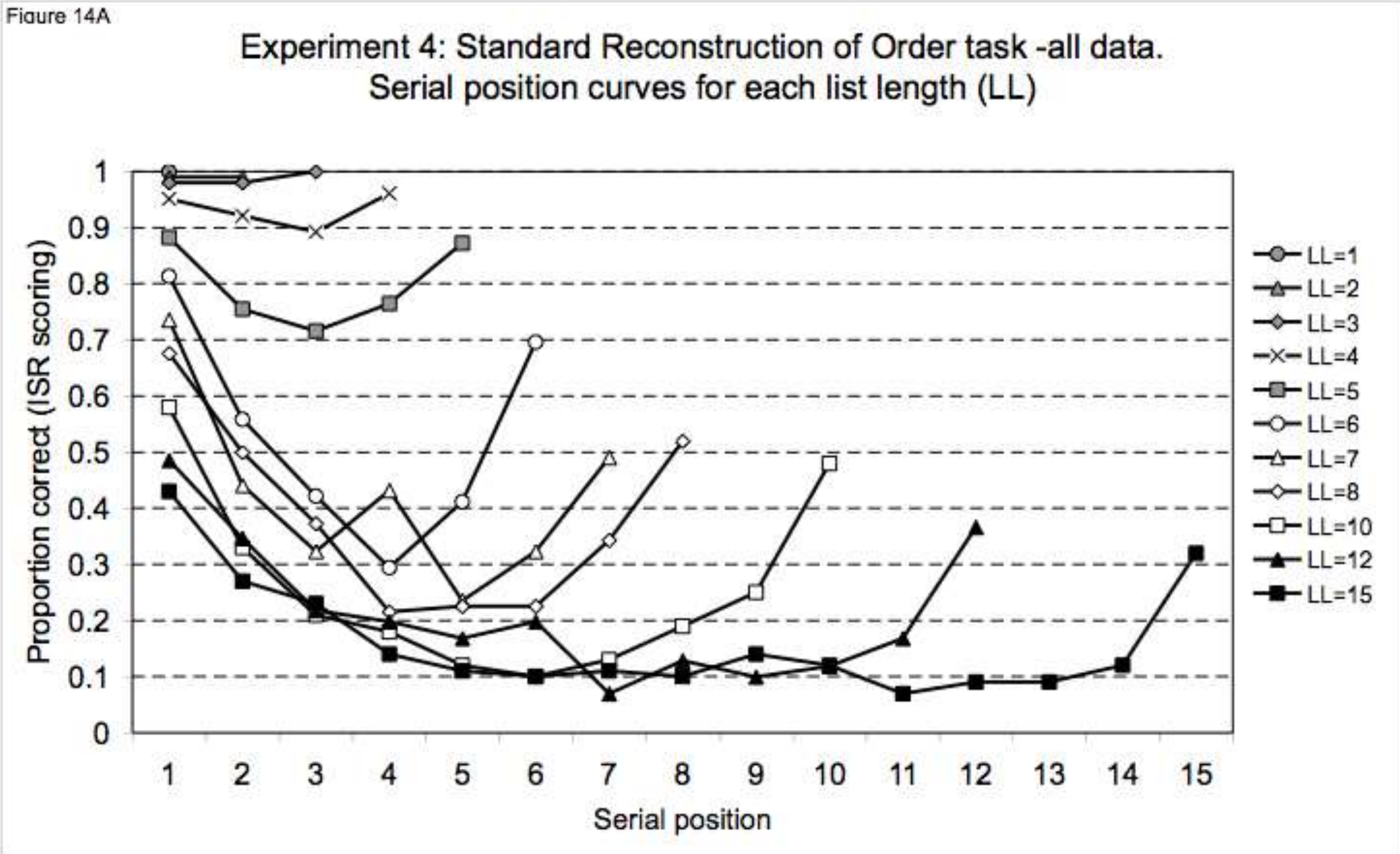
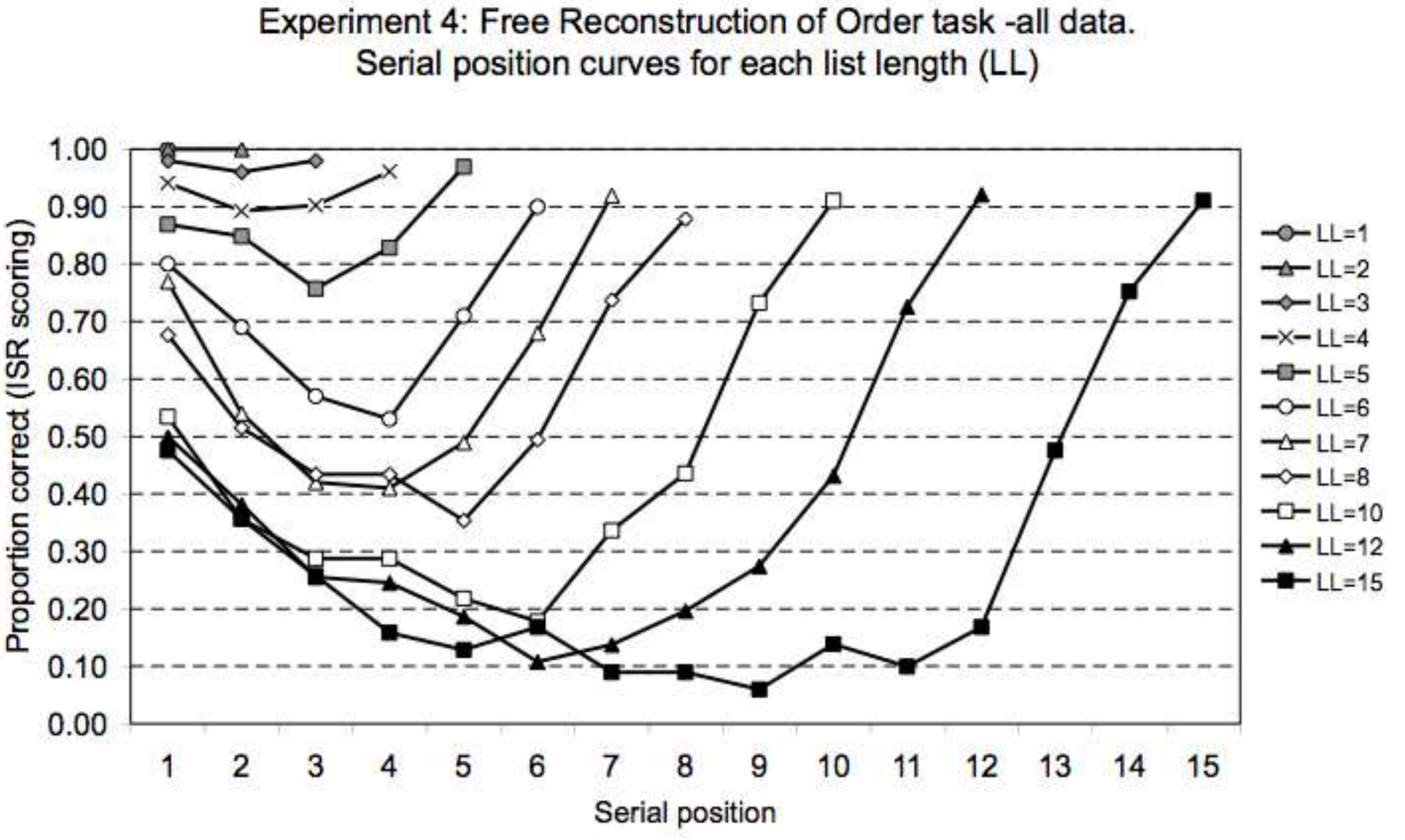


Figure 14A
Experiment 4: Standard Reconstruction of Order task -all data.
Serial position curves for each list length (LL)
Proportion correct (ISR scoring)
Serial position
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

Figure 14B


Experiment 4: Free Reconstruction of Order task -all data.
Serial position curves for each list length (LL)
Proportion correct (ISR scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

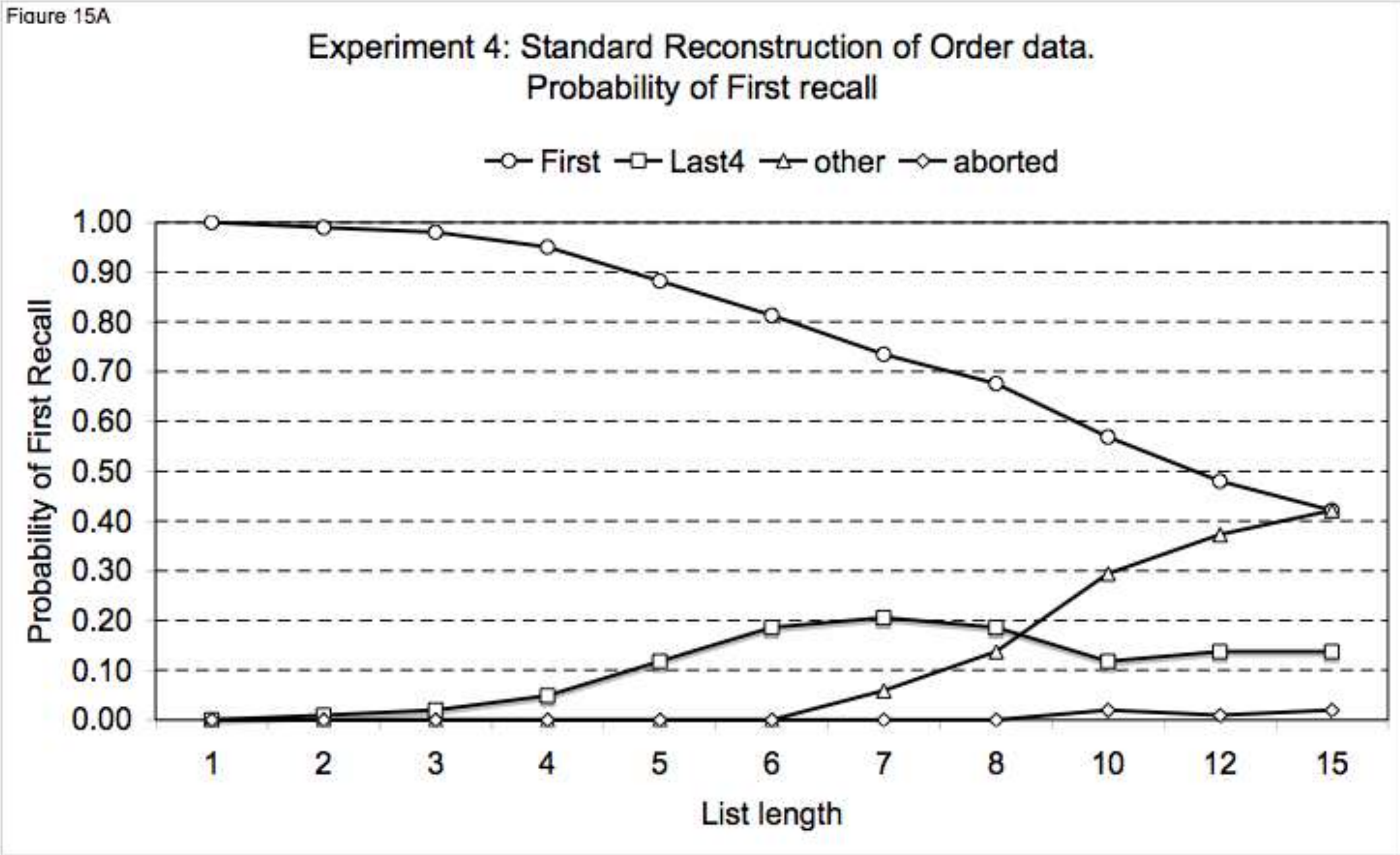


Figure 15A
Experiment 4: Standard Reconstruction of Order data.
Probability of First recall
First
Last4
other
aborted
Probability of First Recall
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
10
12
15
List length

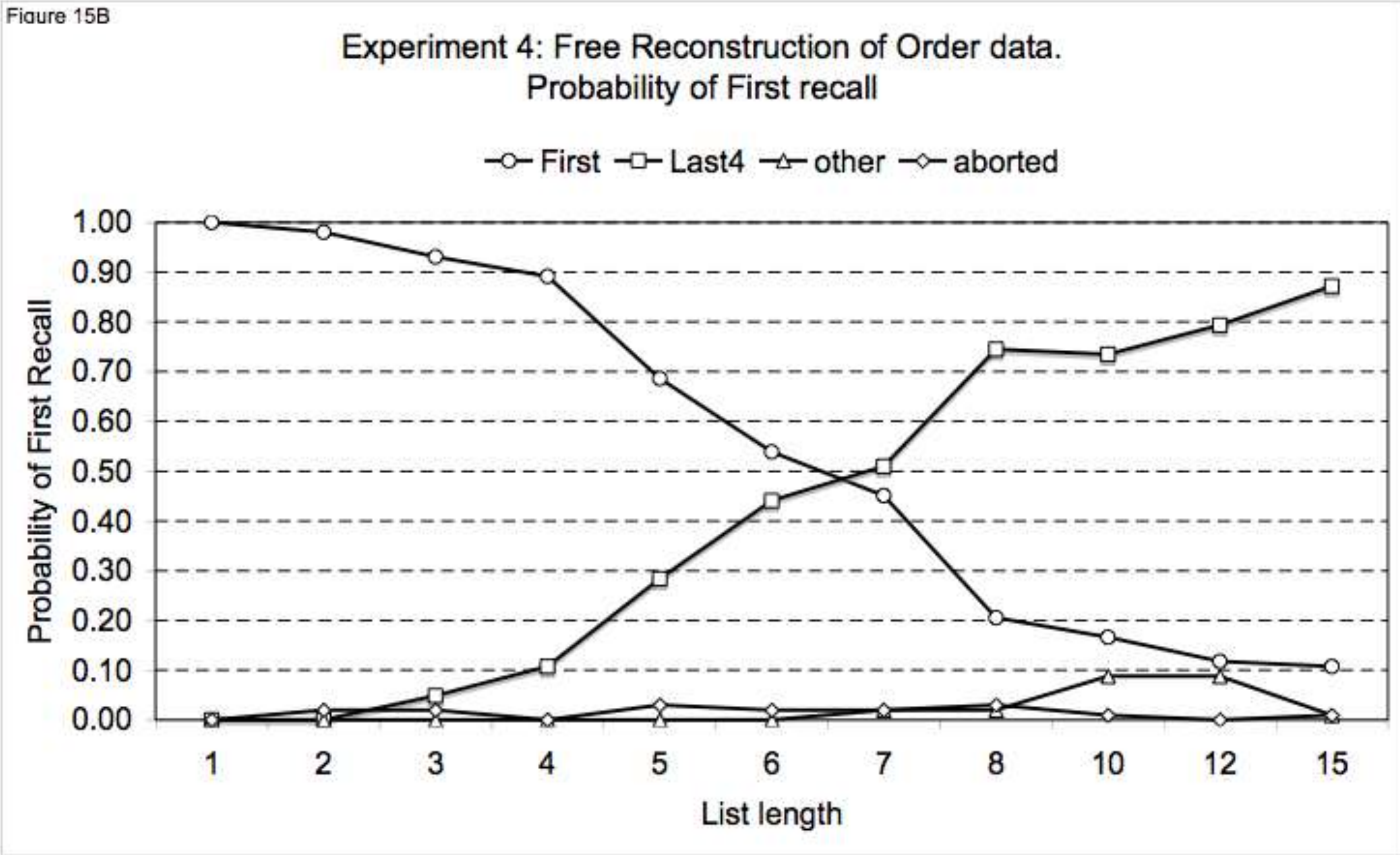


Figure 15B
Experiment 4: Free Reconstruction of Order data.
Probability of First recall
First
Last4
other
aborted
Probability of First Recall
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
10
12
15
List length

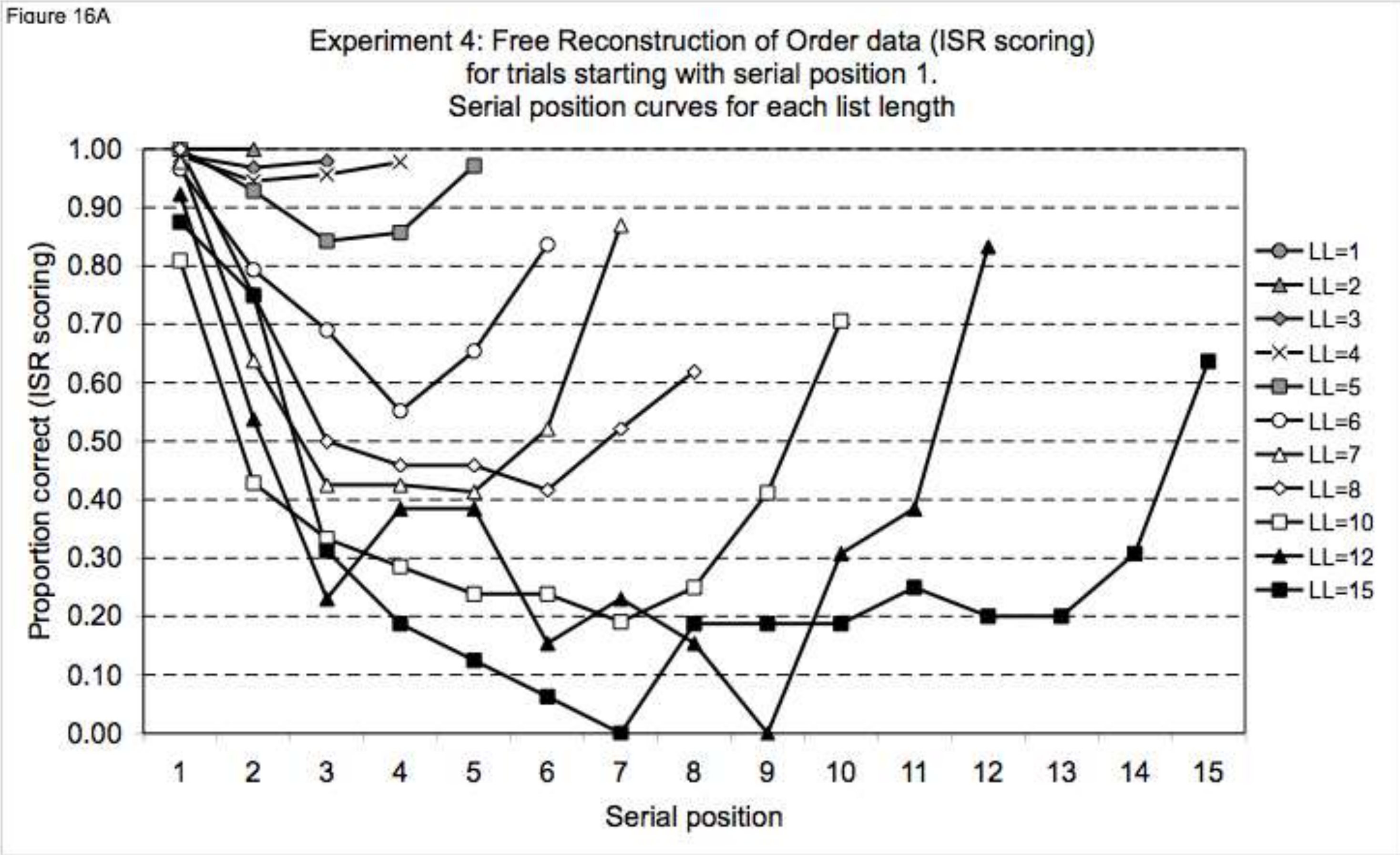


Figure 16A
Experiment 4: Free Reconstruction of Order data (ISR scoring)
for trials starting with serial position 1.
Serial position curves for each list length
Proportion correct (ISR scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

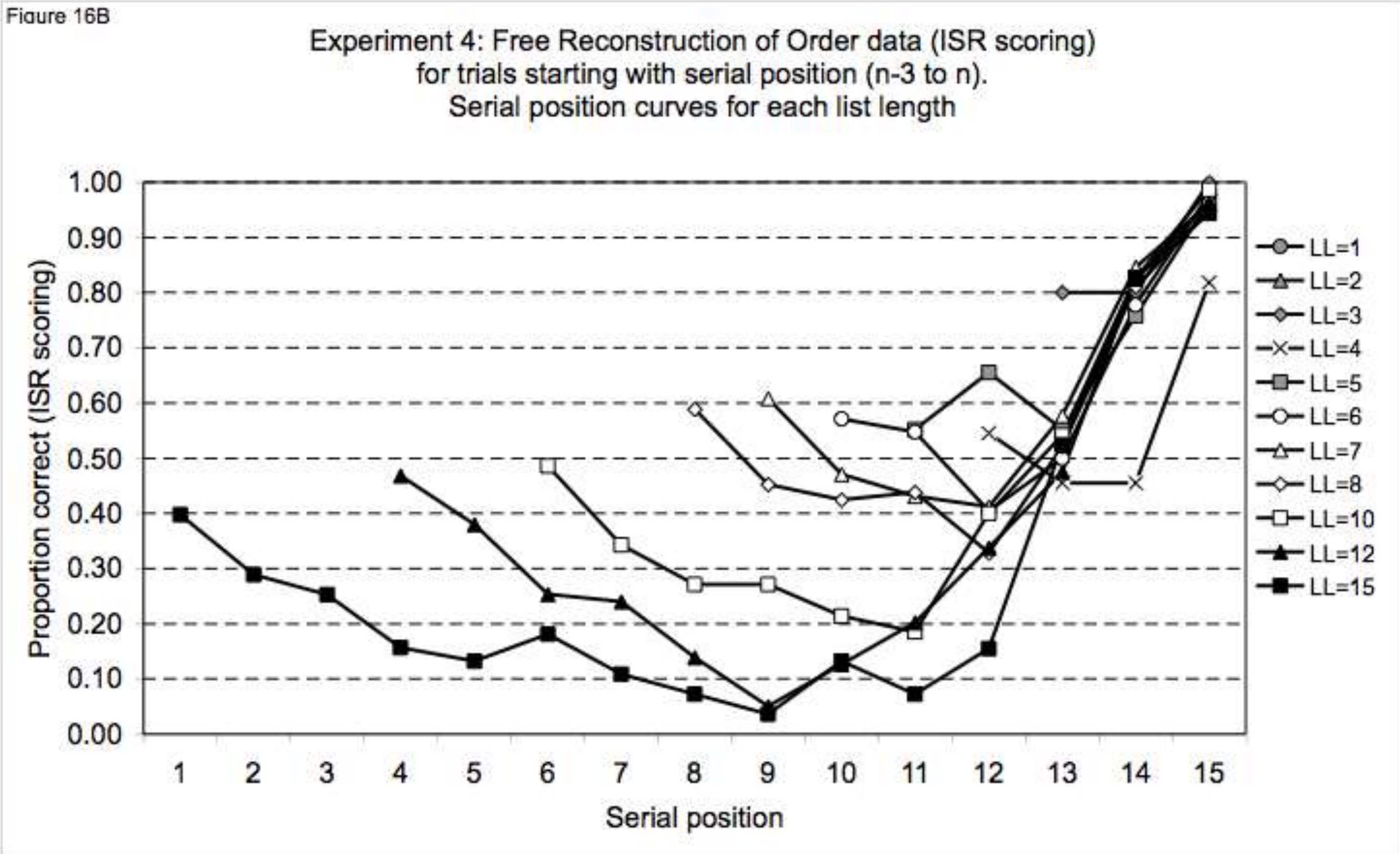
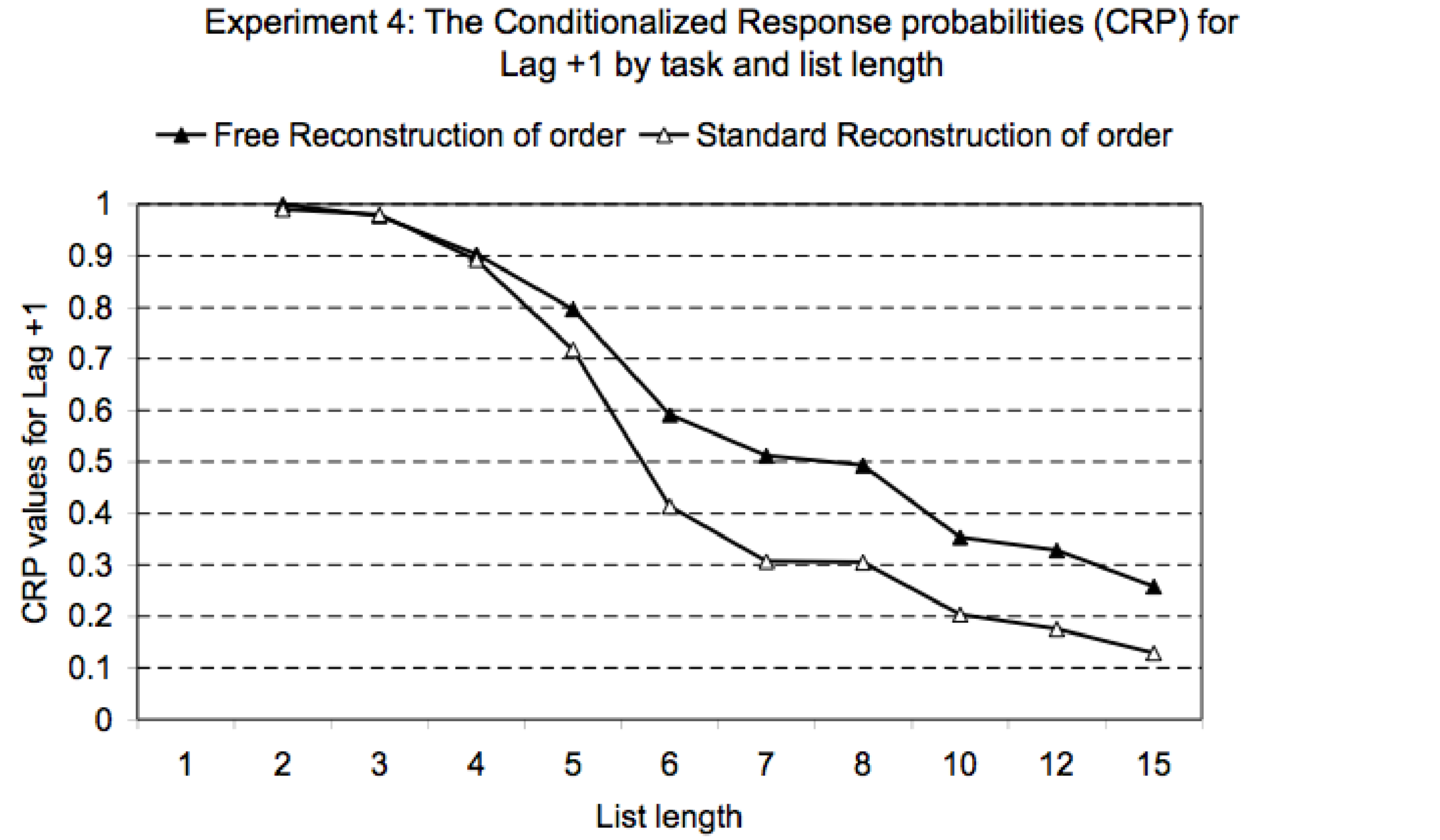


Figure 16B
Experiment 4: Free Reconstruction of Order data (ISR scoring)
for trials starting with serial position (n-3 to n).
Serial position curves for each list length
Proportion correct (ISR scoring)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Serial position
LL=1
LL=2
LL=3
LL=4
LL=5
LL=6
LL=7
LL=8
LL=10
LL=12
LL=15

Figure 17


Experiment 4: The Conditionalized Response probabilities (CRP) for Lag +1 by task and list length
Free Reconstruction of order
Standard Reconstruction of order
CRP values for Lag +1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7
8
10
12
15
List length